



Universidade de Évora
Departamento de Física

Um Estudo Biometeorológico em São Paulo: Elaboração de
um Índice de Risco Meteorotrópico (IRM) sobre a
Morbilidade por Doenças Respiratórias na Cidade de São
Paulo.

Fábio Cunha Conde

Dissertação apresentada à Universidade de Évora
para obtenção do grau de Doutor em Física.

Orientador: Prof. João Corte-Real

Observação: Esta tese não inclui as críticas e sugestões feitas pelo júri.

Évora – 2006.



Universidade de Évora
Departamento de Física

Um Estudo Biometeorológico em São Paulo: Elaboração de
um Índice de Risco Meteorotrópico (IRM) sobre a
Morbilidade por Doenças Respiratórias na Cidade de São
Paulo.



159 507

Fábio Cunha Conde

Dissertação apresentada à Universidade de Évora
para obtenção do grau de Doutor em Física.

Orientador: Prof. João Corte-Real

Observação: Esta tese não inclui as críticas e sugestões feitas pelo júri.

Évora - 2006.

Um Estudo Biometeorológico em São Paulo...

por Fábio Cunha Conde

Discussão sobre os Fundamentos da Dissertação

por José Félix Costa

Departamento de Matemática, Instituto Superior Técnico

Lisboa, Portugal

`fgc@math.ist.utl.pt`

20 de Novembro de 2006

Índice

1	Introdução	2
2	A qualidade do texto em análise	3
3	Questões gerais	3
4	Detalhes e provas	4
5	Crítica ao capítulo fundamental — Capítulo V	7
6	Do apêndice	7
7	Conclusão	12



1 Introdução

Na sua dissertação, o candidato propõe um modelo de morbidade respiratória centrado na informação meteorológica e concentração de poluentes na atmosfera.

O modelo proposto é *conexionista*, isto é, o modelo matemático é uma rede neuronal (construída pelo candidato).

O algoritmo genérico de carregamento de uma rede neuronal com informação necessária e suficiente de modo a conferir-lhe poder de generalização é difícil e, sobretudo, moroso. Um trabalho desta natureza pode levar anos a um investigador que, após a aquisição do conhecimento técnico, esteja disposto a mergulhar nos domínios de uma “religião” denominada RETROPROPAGAÇÃO DE ERROS.

Repito, reiterando o conteúdo da minha apreciação preliminar, para uma boa compreensão deste meu relatório: reconheço que o uso das ferramentas a que o autor recorre, para o objectivo pretendido supramencionado, não é trivial e pode exigir sensibilidades que demoram anos a adquirir. Veja-se o caso da estrutura (secundária, terciária) de proteínas usando as mesmas ferramentas.

Faço notar que o modelo de Fábio Conde não está disponível para experimentação. Não ponho em dúvida a honestidade e a qualidade dos resultados, mas deixo este aspecto à consideração de outros membros deste júri, os quais se terão debruçado sobre a aplicação biometeorológica.

Vou considerar para discussão a fundamentação teórica da tese, a qual se desenvolve essencialmente nos capítulos segundo e terceiro. Constam de 60 páginas de um total de 180 (das quais 80 páginas são constituídas por tabelas e gráficos, com as respectivas legendas).



2 A qualidade do texto em análise

Embora a minha apreciação global do mérito seja positiva, a avaliação que fiz da fundamentação teórica da dissertação é negativa, devendo salientar os seguintes aspectos:

- Os meus bons alunos de *cursos de licenciatura*, através dos anos, e em disciplinas como “Neurocomputação” e “Novos Paradigmas da Computação”, escrevem melhor, desenvolvem mais rigor técnico e apresentam relatórios com mais qualidade.

O texto está mal escrito, não se entende a sua estrutura, o autor repete amiúde conceitos e ideias sendo, as mais das vezes, pouco coerente e, sobretudo, pouco interessante.

- O rigor expositivo é muito pequeno para provas de doutoramento (em Física).
- Como mostrarei, não creio que a alteração pontual de partes da dissertação venha a modificar substancialmente a qualidade do texto.
- Não há neste trabalho qualquer esforço de natureza formal, seja de Física-Matemática ou de Matemática Aplicada.
- A fundamentação teórica da dissertação não aproveita à aprendizagem de conceitos fundamentais que o autor tanto se preocupa em elucidar.
- Não se encontra nesta dissertação justificação da correlação entre variáveis físico-químicas e o risco biometeorológico.

[Quem nos garante que não se trata de uma *impostura intelectual*?]



3 Questões gerais

Que conhecimento de Meteorologia é necessário para construir, com êxito, o modelo que propõe?

Por que não se debruça também sobre modelos de poluição atmosférica e modelos de difusão e retenção de poluentes? Em que medida estes podem determinar a qualidade de um modelo conexionista? Não serão relevantes tais processos na tão vasta área da cidade de S. Paulo?

E.g., alguns poluentes têm distribuição espacial conhecida em redor de focos.

E.g., há situações atmosféricas particularmente notáveis como a inversão de temperatura.

Onde entra toda esta complexa informação no seu modelo? Ou, antes, onde é intencional e justificadamente removida?



4 Detalhes e provas

- O autor promete descrever (página 24, linha 3), com grande detalhe, *os algoritmos de retropropagação de erros (sic)*. No entanto, tal descrição — a meu ver bastante relevante — não se encontra no texto, mas sim uma explicação sumária errada (página 52, linha -5), deixando no ar a dúvida sobre se o autor, para além de dominar uma ferramenta informática, entende, de facto, o modelo com que trabalha. Aliás, tendo argumentado (página 44, linha -10; página 51, linha 12; ...) que *as funções precisam ser diferenciáveis para que o gradiente possa ser calculado (sic)*, não se compreende onde esta diferenciabilidade intervém nas correcções dos pesos da página 52.

O autor devia ter incluído um capítulo sobre a fundamentação matemática da retropropagação de erros.

- Ocorre-nos perguntar em que consiste, afinal, o método do gradiente? (Apresentado sumariamente no Anexo A.)
- O autor refere (página 30, linha -5) que *redes MCP com apenas uma camada conseguem implementar funções linearmente separáveis (sic)*. O que são funções linearmente separáveis? (Embora o conceito ocorra inúmeras vezes, o autor não esclarece do que se trata, embora discorra sobre questões *de lana caprina*, tais como (a) gráfico de função linear, (b) gráfico de função linear por troços, (c) composição de aplicações lineares (curiosidade: *vide* página 43).) É a função booleana *ou exclusivo* linearmente separável? Sim e não!
- Que pretende, de facto, dizer (página 42, linha 1) com *conforme citado na secção anterior, as redes de uma só camada resolvem apenas problemas linearmente separáveis. A solução de problemas não linearmente separáveis passa pelo uso de redes com uma ou mais camadas intermediárias, ou escondidas (sic)*. (De facto, como pode falar de separabilidade linear no contexto das redes neuronais sem se referir primeiro a uma teoria de aprendizagem?)
- O autor não explica, nem no capítulo segundo nem no capítulo terceiro, por que razão a aprendizagem supervisionada foi preferida à aprendizagem não supervisionada.
- O autor refere (página 35, linha -7) que *esta expressão aparece tanto no algoritmo de treino do perceptrão quanto no algoritmo de treino do ADALINE e posterior generalização para o algoritmo backpropagation (sic)*. A lei de plasticidade sináptica está errada. Com a devida

correção, em que se assemelha e em que difere da regra delta generalizada? De novo, não se compreende a questão da diferenciabilidade da função de activação. Afirma (página 50, linha -4) que *o backpropagation é baseado na regra proposta por Widrow e Hoff*. A que regra se refere?

- O autor menciona (página 36, linha -10) *os resultados obtidos por Hebb (sic)*. Que resultados são estes? Por que há-de haver sincronismo entre *input* e *output*?

Que leis gerais pode deduzir destes fragmentos relevantes dos tratados de Donald Hebb e William James:

When an axon of cell A is near enough to excite a cell B and repeatedly or persistently takes part in firing it, some growth process or metabolic change takes place in one or both cells such that A's efficiency as one of the cells firing B, is increased.

Donald Hebb, *Organization of Behavior*, 1949, p.62.

When two elementary brain processes have been active together or in immediate succession one of them on recurring, tends to propagate its excitement into the other... The amount of activity at any given point in the brain-cortex is the sum of the tendencies of all other points to discharge into it...

William James, *Psychology: Briefer Course*, 1892, p.226.

- Acerca de aprendizagem não supervisionada, escreve o candidato (página 38, linha 12): *a unidade mais forte fica ainda mais forte e seu efeito inibidor sobre as outras unidades de saída torna-se dominante (sic)*. Relativamente à arquitectura ART, que quer isto dizer?

Tomemos como exemplo a MAXNET com dinâmica

$$x_i^+ = f(x_i - \epsilon \sum_{k \neq i} x_k),$$

e função de activação

$$f(x) = \begin{cases} x & x > 0 \\ 0 & x \leq 0 \end{cases}$$

Escolha-se $\epsilon = -0.2$ e activações iniciais $x_1(0) = 0.2$, $x_2(0) = 0.4$, $x_3(0) = 0.6$ e $x_4(0) = 0.8$.

Eis a dinâmica:

$$\begin{bmatrix} t = 0 \\ t = 1 \\ t = 2 \\ t = 3 \\ t = 4 \\ t = 5 \end{bmatrix} = \begin{bmatrix} 0.200 & 0.400 & 0.600 & 0.800 \\ 0.000 & 0.080 & 0.320 & 0.560 \\ 0.000 & 0.000 & 0.192 & 0.480 \\ 0.000 & 0.000 & 0.096 & 0.442 \\ 0.000 & 0.000 & 0.008 & 0.422 \\ 0.000 & 0.000 & 0.000 & 0.421 \end{bmatrix}.$$

Comente.

- Na minha opinião, o candidato não conhece nem os autores nem as referências que cita. Dou um exemplo: o célebre artigo de McCulloch e Pitts de 1943:

– Página 30: *redes de MCP com apenas uma camada (sic)*. As redes de McCulloch e Pitts realizam fórmulas temporais e são cíclicas, recorrentes!

– Página 30: *McCulloch e Pitts simplificaram o seu modelo assumindo que os nodos em cada camada da rede disparam sincronamente (sic)*. As redes de McCulloch e Pitts não se estruturam em camadas!

Veja-se este exemplo de uma rede de McCulloch e Pitts — adaptada a pesos.

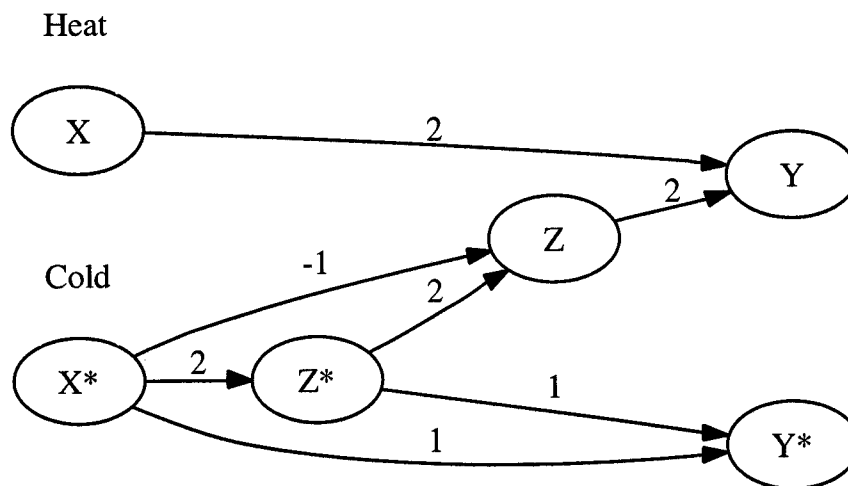


Figure 1: Modelo de percepção de *frio* e *quente*.

- Página 40: *o perceptrão foi proposto por Rosenblatt e originou-se do trabalho de McCulloch e Pitts (sic)*. Kleene veio a demonstrar que as redes de McCulloch e Pitts correspondem exactamente aos autómatos finitos, que as redes são precisamente autómatos finitos. O perceptrão de Rosenblatt, embora seja uma rede

neuronal, inspira-se numa lei de aprendizagem alheia ao trabalho de McCulloch e Pitts.

- Página 40: *o modelo proposto por Rosenblatt, conhecido como perceptrão, era composto por uma estrutura de rede tendo como unidades básicas nodos MCP e uma regra de aprendizado (sic).* Mais uma vez se evidencia que o autor crê que as redes de McCulloch e Pitts envolvem plasticidade sináptica.
- Página 41: *alguns anos mais tarde, Rosenblatt demonstrou o teorema de convergência do perceptrão, o qual mostra que um nodo MCP treinado com o algoritmo de aprendizado do perceptrão sempre converge, caso o problema seja linearmente separável (sic).* Que significa esta asserção? O enunciado do Teorema de Rosenblatt, que podia muito bem ter sido incluído, para evitar estas asserções ambíguas, é o seguinte: seja $\{\vec{x}(p) : t(p)\}$, $p = 1, \dots, n$, um conjunto de padrões vectoriais a classificar e f a função de activação do perceptrão; se existe um vector de pesos \vec{u} tal que $f(\vec{x}(p) \cdot \vec{u}) = t(p)$, para todo o $p = 1, \dots, n$, então o algoritmo de aprendizagem do perceptrão converge num número finito de passos, obtendo-se um vector de pesos \vec{w} (possivelmente diferente de \vec{u}) tal que $f(\vec{x}(p) \cdot \vec{w}) = t(p)$, para todo o $p = 1, \dots, n$. O que converge é o algoritmo de aprendizagem, não os *nodos* do perceptrão!



5 Crítica ao capítulo fundamental — Capítulo V

O autor não se refere à aplicação informática que utiliza. Não se sabe se é produto comercial, ou antes, se se trata de ferramenta computacional desenvolvida pelo autor. Esta informação é importante.

Tal como o autor indica (página 47, linha 14 e seguintes), a conexidade da rede pode não ser total, quando os valores dos pesos forem abaixo de certo limiar de aceitabilidade. Ora, um bom método de treino de redes neuronais com vista a conexidade parcial é o algoritmo genético. Que lhe apraz dizer?



6 Do apêndice

A linguagem do apêndice é quase incompreensível quando comparada com a linguagem do corpo da tese. Usando a matemática do autor (*vide* capítulo 4), vou apresentar uma alternativa à sua descrição, para ilustrar um dos casos tratados em apêndice e, assim, desmistificar o método do autor (versão simplificada).

A ADALINE (ADAPtive LInear NEuron), foi desenhada em 1959 por Bernard Widrow na Universidade de Stanford e desenvolvida mais tarde por Widrow e seus colegas. Trata-se de um neurónio de *output* bipolar. Os *inputs* são, geralmente, também bipolares, mas não há restrição: podem ser binários ou bipolares, ou mesmo números reais.

Considerando m padrões $\vec{x}^{(i)} : y^{(i)}$, e tomando w_0, w_1, \dots, w_n para denotar os coeficientes (pesos) de um elemento (neurónio de Widrow-Hoff), obtemos as equações:

$$\begin{aligned} w_0 + w_1 x_1^{(1)} + \dots + w_n x_n^{(1)} - y^{(1)} &= \epsilon^{(1)} \\ &\vdots \\ w_0 + w_1 x_1^{(m)} + \dots + w_n x_n^{(m)} - y^{(m)} &= \epsilon^{(m)} \end{aligned}$$

onde o i -ésimo erro $\epsilon^{(i)}$ é a diferença entre o estímulo total e a resposta desejada (e não a diferença entre o valor da função de activação e a resposta desejada).

Este sistema de equações pode representar-se matricialmente por

$$\begin{bmatrix} 1 & x_1^{(1)} & \dots & x_n^{(1)} \\ \vdots & \vdots & \vdots & \vdots \\ 1 & x_1^{(m)} & \dots & x_n^{(m)} \end{bmatrix} \begin{bmatrix} w_0 \\ \vdots \\ w_n \end{bmatrix} - \begin{bmatrix} y^{(1)} \\ \vdots \\ y^{(m)} \end{bmatrix} = \begin{bmatrix} \epsilon^{(1)} \\ \vdots \\ \epsilon^{(m)} \end{bmatrix},$$

ou seja

$$\begin{bmatrix} -y^{(1)} & 1 & x_1^{(1)} & \dots & x_n^{(1)} \\ \vdots & \vdots & \vdots & \vdots & \vdots \\ -y^{(m)} & 1 & x_1^{(m)} & \dots & x_n^{(m)} \end{bmatrix} \begin{bmatrix} 1 \\ w_0 \\ \vdots \\ w_n \end{bmatrix} = \begin{bmatrix} \epsilon^{(1)} \\ \vdots \\ \epsilon^{(m)} \end{bmatrix}.$$

Resulta que o i -ésimo erro $\epsilon^{(i)}$ é dado por

$$\epsilon_i = \begin{bmatrix} -y^{(i)} & 1 & x_1^{(i)} & \dots & x_n^{(i)} \end{bmatrix} \begin{bmatrix} 1 \\ w_0 \\ \vdots \\ w_n \end{bmatrix} = \begin{bmatrix} 1 & w_0 & \dots & w_n \end{bmatrix} \begin{bmatrix} -y^{(i)} \\ 1 \\ x_1^{(i)} \\ \vdots \\ x_n^{(i)} \end{bmatrix}.$$

Um processo de resolver este sistema (sobredeterminado) é submetê-lo ao critério dos mínimos quadrados, procurando a solução χ do sistema que

corresponde ao mínimo valor do erro (neste caso existe um mínimo global, porquê?), $\bar{\epsilon}$, definido como a soma dos quadrados dos erros, $\epsilon^{(i)}$, $i = 1, \dots, m$, que correspondem aos m padrões. Toma-se:

$$\begin{aligned} \bar{\epsilon} &= \sum_{i=1}^m \left[1 \quad w_0 \quad w_1 \quad \dots \quad w_n \right] \begin{bmatrix} -y^{(i)} \\ 1 \\ x_1^{(i)} \\ \vdots \\ x_n^{(i)} \end{bmatrix} \begin{bmatrix} -y^{(i)} & 1 & x_1^{(i)} & \dots & x_n^{(i)} \end{bmatrix} \begin{bmatrix} 1 \\ w_0 \\ w_1 \\ \vdots \\ w_n \end{bmatrix} \\ &= \left[1 \quad w_0 \quad w_1 \quad \dots \quad w_n \right] \sum_{i=1}^m \left\{ \begin{bmatrix} -y^{(i)} \\ 1 \\ x_1^{(i)} \\ \vdots \\ x_n^{(i)} \end{bmatrix} \begin{bmatrix} -y^{(i)} & 1 & x_1^{(i)} & \dots & x_n^{(i)} \end{bmatrix} \right\} \begin{bmatrix} 1 \\ w_0 \\ w_1 \\ \vdots \\ w_n \end{bmatrix} \end{aligned}$$

em que a matriz central é simétrica. Encontramos a solução χ para o mínimo valor de $\bar{\epsilon}$ requerendo

$$\frac{\partial \bar{\epsilon}}{\partial w_0} = \frac{\partial \bar{\epsilon}}{\partial w_1} = \dots = \frac{\partial \bar{\epsilon}}{\partial w_n} = 0.$$

Estas $n + 1$ igualdades dar-nos-ão $n + 1$ equações a $n + 1$ incógnitas w_0, w_1, \dots, w_n :

$$\begin{aligned} 0 = \frac{\partial \bar{\epsilon}}{\partial w_0} &= \begin{bmatrix} 0 & 1 & 0 & \dots & 0 \end{bmatrix} R\chi + \chi^T R \begin{bmatrix} 0 \\ 1 \\ 0 \\ \vdots \\ 0 \end{bmatrix} \\ 0 = \frac{\partial \bar{\epsilon}}{\partial w_n} &= \begin{bmatrix} 0 & 0 & 0 & \dots & 1 \end{bmatrix} R\chi + \chi^T R \begin{bmatrix} 0 \\ 0 \\ 0 \\ \vdots \\ 1 \end{bmatrix}, \end{aligned}$$

onde χ é o vector

$$\chi = \begin{bmatrix} 1 \\ w_0 \\ w_1 \\ \vdots \\ w_n \end{bmatrix}.$$

Como $r_{ij} = r_{ji}$, ambos os termos de cada uma destas $n + 1$ igualdades representam a mesma quantidade, pelo que podemos escrever

$$0 = \frac{\partial \bar{\epsilon}}{\partial w_0} = 2 \begin{bmatrix} r_{10} & r_{11} & r_{12} & \cdots & r_{1(n+1)} \\ \vdots & & & & \\ r_{(n+1)0} & r_{(n+1)1} & r_{(n+1)2} & \cdots & r_{(n+1)(n+1)} \end{bmatrix} \chi$$

Seguidamente, reunimos estas equações na equação matricial

$$\begin{bmatrix} r_{10} & r_{11} & r_{12} & \cdots & r_{1(n+1)} \\ \vdots & \vdots & \vdots & \vdots & \vdots \\ r_{(n+1)0} & r_{(n+1)1} & r_{(n+1)2} & \cdots & r_{(n+1)(n+1)} \end{bmatrix} \begin{bmatrix} 1 \\ w_0 \\ \vdots \\ w_n \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \\ \vdots \\ 0 \end{bmatrix},$$

ou seja

$$\begin{bmatrix} r_{11} & r_{12} & \cdots & r_{1(n+1)} \\ \vdots & \vdots & \vdots & \vdots \\ r_{(n+1)1} & r_{(n+1)2} & \cdots & r_{(n+1)(n+1)} \end{bmatrix} \begin{bmatrix} w_0 \\ \vdots \\ w_n \end{bmatrix} = - \begin{bmatrix} r_{10} \\ \vdots \\ r_{(n+1)0} \end{bmatrix}$$

cuja solução é

$$\begin{bmatrix} w_0 \\ \vdots \\ w_n \end{bmatrix} = - \begin{bmatrix} r_{11} & r_{12} & \cdots & r_{1(n+1)} \\ \vdots & \vdots & \vdots & \vdots \\ r_{(n+1)1} & r_{(n+1)2} & \cdots & r_{(n+1)(n+1)} \end{bmatrix}^{-1} \begin{bmatrix} r_{10} \\ \vdots \\ r_{(n+1)0} \end{bmatrix}.$$

Vamos calcular cada um dos elementos r_{ij} da matriz R . Recordando a definição da matriz R e desenvolvendo os cálculos encontramos:

$$R = \sum_{i=1}^m \begin{bmatrix} -y^{(i)} \\ 1 \\ x_1^{(i)} \\ \vdots \\ x_n^{(i)} \end{bmatrix} \begin{bmatrix} -y^{(i)} & 1 & x_1^{(i)} & \cdots & x_n^{(i)} \end{bmatrix}$$

$$\begin{aligned}
&= \sum_{i=1}^m \begin{bmatrix} y^{(i)2} & -y^{(i)} & -y^{(i)}x_1^{(i)} & \dots & -y^{(i)}x_n^{(i)} \\ -y^{(i)} & 1 & x_1^{(i)} & \dots & x_n^{(i)} \\ \vdots & \vdots & \vdots & \vdots & \vdots \\ -y^{(i)}x_n^{(i)} & x_n^{(i)} & x_1^{(i)}x_n^{(i)} & \dots & x_n^{(i)2} \end{bmatrix} \\
&= \begin{bmatrix} \sum_{i=1}^m y^{(i)2} & -\sum_{i=1}^m y^{(i)} & -\sum_{i=1}^m y^{(i)}x_1^{(i)} & \dots & -\sum_{i=1}^m y^{(i)}x_n^{(i)} \\ -\sum_{i=1}^m y^{(i)} & n & \sum_{i=1}^m x_1^{(i)} & \dots & \sum_{i=1}^m x_n^{(i)} \\ \vdots & \vdots & \vdots & \vdots & \vdots \\ -\sum_{i=1}^m y^{(i)}x_n^{(i)} & \sum_{i=1}^m x_n^{(i)} & \sum_{i=1}^m x_1^{(i)}x_n^{(i)} & \dots & \sum_{i=1}^m x_n^{(i)2} \end{bmatrix}
\end{aligned}$$

Conclui-se que

$$\begin{bmatrix} w_0 \\ w_1 \\ \vdots \\ w_n \end{bmatrix} = \begin{bmatrix} n & \sum_{i=1}^m x_1^{(i)} & \dots & \sum_{i=1}^m x_n^{(i)} \\ \sum_{i=1}^m x_1^{(i)} & \sum_{i=1}^m x_1^{(i)2} & \dots & \sum_{i=1}^m x_n^{(1)}x_n^{(i)} \\ \vdots & \vdots & \vdots & \vdots \\ \sum_{i=1}^m x_n^{(i)} & \sum_{i=1}^m x_1^{(i)}x_n^{(i)} & \dots & \sum_{i=1}^m x_n^{(i)2} \end{bmatrix}^{-1} \begin{bmatrix} \sum_{i=1}^m y^{(i)} \\ \sum_{i=1}^m y^{(i)}x_1^{(i)} \\ \vdots \\ \sum_{i=1}^m y^{(i)}x_n^{(i)} \end{bmatrix}$$

que corresponde ao caso linear tratado pelo autor no seu apêndice.

Considerando de novo o erro $\bar{\epsilon}$, temos que

$$\begin{aligned}
\frac{\partial \bar{\epsilon}}{\partial w_i} &= 2 \begin{bmatrix} r_{i0} & r_{i1} & \dots & r_{i(n+1)} \end{bmatrix} \chi \\
&= 2 \begin{bmatrix} -\sum_{j=1}^m y^{(j)}x_i^{(j)} & \sum_{j=1}^m x_i^{(j)} & \sum_{j=1}^m x_1^{(j)}x_i^{(j)} & \dots & \sum_{j=1}^m x_i^{(j)}x_n^{(j)} \end{bmatrix} \chi.
\end{aligned}$$

Para um só padrão,

$$\frac{\partial \bar{\epsilon}}{\partial w_i} = 2 \begin{bmatrix} -yx_i & x_i & x_1x_i & \dots & x_nx_i \end{bmatrix} \chi,$$

ou seja

$$\begin{aligned}
\frac{\partial \bar{\epsilon}}{\partial w_i} &= -2yx_i + 2x_iw_0 + 2x_ix_1w_1 + \dots + 2x_ix_nw_n \\
&= -2x_i(y - (w_0 + w_1x_1 + \dots + w_nx_n)) = -2(y - e)x_i.
\end{aligned}$$

Em cada passo de aprendizagem, o erro pode ser reduzido ajustando os pesos na proporção do gradiente descendente, ou seja,

$$\Delta w_i = -\eta' \frac{\partial \bar{e}}{\partial w_i} = \eta(y - e)x_i,$$

o que conduz à lei de aprendizagem $w_i^+ = w_i + \eta(y - e)x_i$.

Ficava bem ao autor um estudo exaustivo, desta natureza, na sua dissertação. De igual modo, sente-se a falta de uma boa secção sobre pre-processamento do *input*.

¶

7 Conclusão

Afigura-se-me que uns tantos meses bastariam para redigir, sem grande esforço, um texto sucinto e elegante sobre as ferramentas de que Fábio Conde faz tanto uso.

A minha conclusão é a seguinte: independentemente do valor intrínseco das correlações demonstradas na sua dissertação, o autor navega em águas pouco profundas de uma teoria matemática que ele parece não conhecer.

Évora, 20 de Novembro de 2006

José Félix Costa



(Professor Associado c/ Agregação)

AGRADECIMENTOS

Primeiramente, eu agradeço a Deus pelo equilíbrio proporcionado nos momentos difíceis.

Agradeço ao meu orientador, Prof. Dr. João Alexandre Medina Corte-Real, pela orientação e apoio dado durante o desenvolvimento deste trabalho.

À Andréa Ramos, sempre presente em todos os momentos de minha vida, proporcionando confiança, companheirismo e tranquilidade.

Ao Dr. Luiz Alberto Pereira pelo suporte dos dados de doenças respiratórias, assim como, pela disponibilidade e orientações no tratamento dos códigos de doenças respiratórias e suporte na área da medicina.

À Diretoria de Hidrografia e Navegação (DHN) pela aquisição das cartas sinópticas, especialmente, os Srs. Alberto e Helton pela cordialidade, eficiência e rapidez no tratamento e envio das cartas.

Ao Departamento de Física e Centro de Geofísica de Évora por me oferecer as condições adequadas para a realização deste trabalho.

Aos Professores Rui Namorado Rosa, António Heitor Reis e Ana Maria Silva pela confiança em mim depositada.

Agradeço aos meus amigos, Mário Drumond, Gérard e, especialmente, Paulo Lucio pela amizade, brincadeiras e companheirismo.

Ao Dr. Waldimir Santana pela convivência, conselhos e dedicação depositada em mim, durante a realização da pesquisa.

Aos meus pais, Constantino Conde da Silva e Maria Lúcia Cunha Conde, pelo apoio sempre prestado: Meu eterno agradecimento.

Resumo

O meio ambiente atmosférico inclui uma série de factores cuja relação com organismos deve ser feita de início em separado, ainda que existam inter-relações que multiplicam ou reduzem os efeitos observáveis. Do mesmo modo, verificou-se que em muitas ocasiões, tendem a aparecer no homem epidemias ou doenças agudas com acessos quase periódicos. Uma doença generalizada, que aparece ou desaparece consoante as mudanças atmosféricas, denomina-se doença meteorotrópica. Este trabalho tem como objectivo o desenvolvimento de uma metodologia que viabilize a criação de um Índice de Risco Meteorotrópico (IRM), considerando as consequências das variáveis ambientais (poluentes e meteorológicas) sobre a variabilidade da morbidade de todas as doenças do aparelho respiratório (conforme, CID 9 e CID 10), na Cidade de São Paulo durante o período de 1993 a 2002. A elaboração do IRM foi efectuada através do emprego das Rede Neurais Artificiais, considerando a classificação etária e o sexo do paciente. Além das RNA's, foram utilizadas técnicas de análises multivariadas de Análise de Cluster (AC) e da Análise em Componentes Principais (ACP). A AC foi utilizado em carácter exploratório, para verificar a similaridade das variáveis envolvidas e a ACP em carácter confirmatório. Essas duas técnicas corroboraram com as informações obtidas pela rede neuronal. A metodologia desenvolvida nesta pesquisa viabilizou a criação do Índice de Risco Meteorotrópico, assim como a previsão das doenças respiratórias conforme a classificação etária e sexo, utilizando uma rede de multicamadas perceptrão com o algoritmo de aprendizagem "backpropagation". Os prognósticos da morbidade respiratória, quando envolvia todos os grupos etários e o sexo [RESP6 - F, M e MF] da Cidade de São Paulo, apresentou uma performance mais robusta para a solução do problema com um grau de acerto em torno de 82% quando comparado com os valores observados. Os padrões de sazonalidade e ciclo semanais das doenças respiratórias foram detectados durante o aprendizado da rede.

Abstract

The atmospheric environment includes a series of factors whose relationship with organisms should be made at the beginning in separate, although, exist interrelations that multiply or reduce the observable effects. In the same way, it is verified in a lot of occasions the epidemics or burden diseases tend to appear periodically on human being. The disease, that appears or consonant disappears due the atmospheric changes is called meteorotropic disease. This study presents the development of a methodology that became possible the creation of a Risk Meteorotropic Index (RMI), considering the consequences of the environmental variables (pollutant and meteorological) about morbidity variability for all respiratory diseases (ICD 9 and ICD 10) in São Paulo city during for period 1993-2002. The elaboration of RMI will be made via Artificial Neural Networks (ANN's) in agreement with the age classification and sex. Besides ANN's, two statistical multivariate techniques were used: Principal Components Analysis (PCA) and Cluster Analysis (CA). The CA was used in exploratory character to verify the similarity of the involved variables while PCA was used in confirmative character. The multivariate techniques corroborated with the information obtained by neural network. The results obtained through the methodology developed in this research became possible the creation Risk Meteorotropic Index (RMI), as well as the forecasting respiratory diseases according to the age classification and sex via a multilayer perceptron with the backpropagation algorithm. The prognostics of the respiratory morbidity for all age groups and sex [RESP6 - F, M and MF] presented a more robust performance for the problem solution with an accuracy degree around 82% when compared with the observed values. The seasonality patterns and weekdays of the respiratory diseases were detected during the net training.

ÍNDICE

Resumo.....	I
Abstract.....	II
Lista de Figuras.....	III
Lista de Tabelas.....	VII
Lista de Abreviaturas e Símbolos.....	IX
CAPITULO I: Introdução.....	1
1.1. Motivação.....	3
1.2. Objectivo.....	4
1.3. Revisão Bibliográfica.....	4
CAPITULO II: Redes Neuronais Artificiais – Fundamentos.....	24
2.1. Histórico das Redes Neuronais Artificiais (RNA's).....	24
2.2. O Neurónio Biológico.....	27
2.3. Neurónios Artificiais: O Modelo MCP (McCulloch & Pitts).....	28
2.4. Funções de Activação.....	31
2.5. Aprendizado.....	32
2.5.1. Aprendizado Supervisionado.....	33
2.5.2. Correção de Erros.....	35
2.5.3. Aprendizado não Supervisionado.....	35
2.5.4. Aprendizado Hebbiano.....	36
2.5.5. Aprendizado por Competição.....	38
2.5.6. Aprendizado por Reforço.....	38
2.6. Arquitecturas de RNA's.....	40
2.6.1. Perceptrão de Camada Única e seu Algoritmo de Aprendizado.....	41
2.6.2. Perceptrão Multicamadas.....	42
2.6.2.1. Arquitectura PMC.....	43
2.6.2.2. Definição de uma Topologia Ideal para o PMC.....	44

CAPITULO VI: Conclusões.....	170
7. Referências Bibliográficas.....	175
Anexo A.....	188

Lista de Figuras

Figura 1.1 – Esquema da classificação da biometeorologia humana e ciências adjacentes do ponto de vista actual (Höppe, 1997).....	6
Figura 2.1 - Componentes de um Neurônio Biológico.....	27
Figura 2.2 – Exemplo de Neurônio Artificial.....	30
Figura 2.3 – Funções de activação: (a) função linear, (b) função rampa, (c) função degrau (<i>step</i>) e (d) função sigmoidal.....	31
Figura 2.4 – Aprendizado Supervisionado.....	34
Figura 2.5 – Aprendizado Não Supervisionado.....	35
Figura 2.6 – Aprendizado por Reforço.....	39
Figura 2.7 – Topologia de um perceptrão simples com uma única saída.....	41
Figura 2.8 – Rede PMC com camada intermediária.....	44
Figura 2.9 – Fluxo de Processamento do algoritmo backpropagation.....	50
Figura 2.10 – Superfície do Erro de uma Rede PMC.....	52
Figura 3.1 – Rede Telemétrica da CETESB na Cidade de São Paulo.....	63
Figura 3.2 – Identificação dos Sistemas Sinópticos segundo as Cartas do DHN...	67
Figura 4.1 – Esquematização da rede neural para o prognóstico da morbidade respiratória segundo a faixa etária e sexo.....	82
Figura 5.1 – Número de RESP segundo a faixa etária e sexo por dias da semana para todo o período de estudo.....	87
Figura 5.2 – Número de RESP segundo a faixa etária e sexo por sazonalidade para todo o período de estudo.....	88
Figura 5.3 – Frequência de sistemas sinópticos para todo o período de estudo.....	89
Figura 5.4 – Série Temporal dos Poluentes e Padrões Primários: (a) PM ₁₀ , (b) O ₃ , (c) CO e (d) SO ₂	91
Figura 5.5 – Dendograma de árvore através do método de Ward e distâncias euclidianas para todo o período de estudo: RESP1.....	94
Figura 5.6 – Comparação entre os valores observados de RESP1F e os valores	

modelados pela RNA através da Função de Distribuição Cumulativa Empírica (FDCE) e Série Temporal (ST).....	95
Figura 5.7 – Comparação entre os valores observados de RESP1M e os valores modelados pela RNA através da Função de Distribuição Cumulativa Empírica (FDCE) e Série Temporal (ST).....	97
Figura 5.8 – Comparação entre os valores observados de RESP1MF e os valores modelados pela RNA através da Função de Distribuição Cumulativa Empírica (FDCE) e Série Temporal (ST).....	99
Figura 5.9 – Análise de Sensibilidade: (a) RESP1F_LOG_BP – [62-130-1]; (b) RESP1M_TGH_BP – [62-144-1] e (c) RESP1MF_TGH_BP – [62-117-1].....	101
Figura 5.10 – Dendograma de árvore através do método de Ward e distâncias euclidianas para todo o período de estudo: RESP2.....	107
Figura 5.11 – Comparação entre os valores observados de RESP2F e os valores modelados pela RNA através da Função de Distribuição Cumulativa Empírica (FDCE) e Série Temporal (ST).....	108
Figura 5.12 – Comparação entre os valores observados de RESP2M e os valores modelados pela RNA através da Função de Distribuição Cumulativa Empírica (FDCE) e Série Temporal (ST).....	110
Figura 5.13 – Comparação entre os valores observados de RESP2MF e os valores modelados pela RNA através da Função de Distribuição Cumulativa Empírica (FDCE) e Série Temporal (ST).....	113
Figura 5.14 – Análise de Sensibilidade: (a) RESP2F_TGH_BP – [62-108-1]; (b) RESP2M_LOG_BP – [62-32-1] e (c) RESP2MF_TGH_BP – [62-86-1].....	115
Figura 5.15 – Dendograma de árvore através do método de Ward e distâncias euclidianas para todo o período de estudo: RESP3.....	120
Figura 5.16 – Comparação entre os valores observados de RESP3F e os valores modelados pela RNA através da Função de Distribuição Cumulativa Empírica (FDCE) e Série Temporal (ST).....	121
Figura 5.17 – Comparação entre os valores observados de RESP3M e os valores modelados pela RNA através da Função de Distribuição Cumulativa Empírica (FDCE) e Série Temporal (ST).....	123
Figura 5.18 – Comparação entre os valores observados de RESP3MF e os valores modelados pela RNA através da Função de Distribuição Cumulativa Empírica	

(FDCE) e Série Temporal (ST).....	125
Figura 5.19 – Análise de Sensibilidade: (a) RESP3F_TGH_BP – [62-141-1]; (b) RESP3M_LOG_BP – [62-135-1] e (c) RESP3MF_LOG_BP – [62-92-1]	127
Figura 5.20 – Dendograma de árvore através do método de Ward e distâncias euclidianas para todo o período de estudo: RESP4.....	133
Figura 5.21 – Comparação entre os valores observados de RESP4F e os valores modelados pela RNA através da Função de Distribuição Cumulativa Empírica (FDCE) e Série Temporal (ST).....	134
Figura 5.22 – Comparação entre os valores observados de RESP4M e os valores modelados pela RNA através da Função de Distribuição Cumulativa Empírica (FDCE) e Série Temporal (ST).....	136
Figura 5.23 – Comparação entre os valores observados de RESP4MF e os valores modelados pela RNA através da Função de Distribuição Cumulativa Empírica (FDCE) e Série Temporal (ST).....	138
Figura 5.24 – Análise de Sensibilidade: (a) RESP4F_TGH_BP – [62-95-1]; (b) RESP4M_TGH_BP – [62-118-1] e (c) RESP4MF_TGH_BP – [62-138-1].....	140
Figura 5.25 – Dendograma de árvore através do método de Ward e distâncias euclidianas para todo o período de estudo: RESP5.....	146
Figura 5.26 – Comparação entre os valores observados de RESP5F e os valores modelados pela RNA através da Função de Distribuição Cumulativa Empírica (FDCE) e Série Temporal (ST).....	147
Figura 5.27 – Comparação entre os valores observados de RESP5M e os valores modelados pela RNA através da Função de Distribuição Cumulativa Empírica (FDCE) e Série Temporal (ST).....	149
Figura 5.28 – Comparação entre os valores observados de RESP5MF e os valores modelados pela RNA através da Função de Distribuição Cumulativa Empírica (FDCE) e Série Temporal (ST).....	151
Figura 5.29 – Análise de Sensibilidade: (a) RESP5F_TGH_BP – [62-115-1]; (b) RESP5M_LOG_BP – [62-99-1] e (c) RESP5MF_LOG_BP – [62-71-1].....	153
Figura 5.30 – Dendograma de árvore através do método de Ward e distâncias euclidianas para todo o período de estudo: RESP6.....	159
Figura 5.31 – Comparação entre os valores observados de RESP6F e os valores modelados pela RNA através da Função de Distribuição Cumulativa Empírica	

(FDCE) e Série Temporal (ST).....	160
Figura 5.32 – Comparação entre os valores observados de RESP6M e os valores modelados pela RNA através da Função de Distribuição Cumulativa Empírica (FDCE) e Série Temporal (ST).....	162
Figura 5.33 – Comparação entre os valores observados de RESP6MF e os valores modelados pela RNA através da Função de Distribuição Cumulativa Empírica (FDCE) e Série Temporal (ST).....	164
Figura 5.34 – Análise de Sensibilidade: (a) RESP6F_LOG_BP – [62-115-1]; (b) RESP6M_TGH_BP – [62-100-1] e (c) RESP6MF_LOG_BP – [62-31-1].....	167

Lista de Tabelas

Tabela 1.1 – Efeitos dos poluentes (PM ₁₀ , SO ₂ , O ₃ , CO e NO ₂) na atmosfera (Fonte: CETESB, 2002).....	19
Tabela 1.2 – Padrões nacionais de qualidade do ar e critérios para episódios agudos de poluição do ar (resolução CONAMA nº 3 de 8/06/1990. Fonte: CETESB, 1994).....	22
Tabela 5.1 – Pesos das componentes principais após a rotação varimax para todo o período de estudo (01/01/1993 a 31/12/2002) – RESP1.....	93
Tabela 5.2 – Parâmetros da RNA e Estatística para RESP1F.....	96
Tabela 5.3 – Parâmetros da RNA e Estatística para RESP1M.....	98
Tabela 5.4 – Parâmetros da RNA e Estatística para RESP1MF.....	100
Tabela 5.5 – Matriz de Confusão para RESP1F_LOG_BP – [62-130-1].....	103
Tabela 5.6 – Matriz de Confusão para RESP1M_TGH_BP – [62-144-1].....	104
Tabela 5.7 – Matriz de Confusão Para RESP1MF_TGH_BP – [62-117-1].....	104
Tabela 5.8 – Pesos das componentes principais após a rotação varimax para todo o período de estudo (01/01/1993 a 31/12/2002) – RESP2.....	106
Tabela 5.9 – Parâmetros da RNA e Estatística para RESP2F.....	109
Tabela 5.10 – Parâmetros da RNA e Estatística para RESP2M.....	111
Tabela 5.11 – Parâmetros da RNA e Estatística para RESP2MF.....	114
Tabela 5.12 – Matriz de Confusão para RESP2F_TGH_BP – [62-108-1].....	116
Tabela 5.13 – Matriz de Confusão para RESP2M_LOG_BP – [62-32-1].....	117
Tabela 5.14 – Matriz de Confusão Para RESP2MF_TGH_BP – [62-86-1].....	117
Tabela 5.15 – Pesos das componentes principais após a rotação varimax para todo o período de estudo (01/01/1993 a 31/12/2002) – RESP3.....	119
Tabela 5.16 – Parâmetros da RNA e Estatística para RESP3F.....	122
Tabela 5.17 – Parâmetros da RNA e Estatística para RESP3M.....	124
Tabela 5.18 – Parâmetros da RNA e Estatística para RESP3MF.....	126
Tabela 5.19 – Matriz de Confusão para RESP3F_TGH_BP – [62-141-1].....	129
Tabela 5.20 – Matriz de Confusão para RESP3M_LOG_BP – [62-135-1].....	130
Tabela 5.21 – Matriz de Confusão Para RESP3MF_LOG_BP – [62-92-1].....	130
Tabela 5.22 – Pesos das componentes principais após a rotação varimax para todo	

o período de estudo (01/01/1993 a 31/12/2002) – RESP4.....	132
Tabela 5.23 – Parâmetros da RNA e Estatística para RESP4F.....	135
Tabela 5.24 – Parâmetros da RNA e Estatística para RESP4M.....	137
Tabela 5.25 – Parâmetros da RNA e Estatística para RESP4MF.....	139
Tabela 5.26 – Matriz de Confusão para RESP4F_TGH_BP – [62-95-1].....	142
Tabela 5.27 – Matriz de Confusão para RESP4M_TGH_BP – [62-118-1].....	143
Tabela 5.28 – Matriz de Confusão Para RESP4MF_TGH_BP – [62-138-1].....	143
Tabela 5.29 – Pesos das componentes principais após a rotação varimax para todo o período de estudo (01/01/1993 a 31/12/2002) – RESP5.....	145
Tabela 5.30 – Parâmetros da RNA e Estatística para RESP5F.....	148
Tabela 5.31 – Parâmetros da RNA e Estatística para RESP5M.....	150
Tabela 5.32 – Parâmetros da RNA e Estatística para RESP5MF.....	152
Tabela 5.33 – Matriz de Confusão para RESP5F_TGH_BP – [62-115-1].....	155
Tabela 5.34 – Matriz de Confusão para RESP5M_LOG_BP – [62-99-1].....	156
Tabela 5.35 – Matriz de Confusão Para RESP5MF_LOG_BP – [62-71-1].....	156
Tabela 5.36 – Pesos das componentes principais após a rotação varimax para todo o período de estudo (01/01/1993 a 31/12/2002) – RESP6.....	158
Tabela 5.37 – Parâmetros da RNA e Estatística para RESP6F.....	161
Tabela 5.38 – Parâmetros da RNA e Estatística para RESP6M.....	163
Tabela 5.39 – Parâmetros da RNA e Estatística para RESP6MF.....	165
Tabela 5.40 – Matriz de Confusão para RESP6F_LOG_BP – [62-115-1].....	168
Tabela 5.41 – Matriz de Confusão para RESP6M_TGH_BP – [62-100-1].....	169
Tabela 5.42 – Matriz de Confusão Para RESP6MF_LOG_BP – [62-31-1].....	169

Lista de Abreviaturas e Símbolos

f – Função de activação

η – Taxa de aprendizagem

μ – Taxa de momento

θ – limiar (*threshold*) do neurónio.

AC – Alta Continental

AC – Análise de Cluster

ACP – Análise em Componentes Principais

AIH's – Autorizações de Internamento Hospitalar

AP – Alta Polar

AS – Alta Subtropical do Atlântico Sul

b_k – Bias ou polarização

BP – Algoritmos Back-Propagation ou de retropropagação

CETESB – Companhia de Tecnologia de Saneamento Ambiental

CID – Classificação Internacional de Doenças

CO – Monóxido de Carbono

CONAMA – Conselho Nacional do Meio Ambiente

CP – Componente Principal

CSP – Cidade de São Paulo

dd – Direcção do vento

DGC – Algoritmo Descida do Gradiente Conjugada

DHN – Directoria de Hidrografia e Navegação do Ministério da Marinha

F – Sistema Frontal

FDCE– Função de Distribuição Cumulativa Empírica

IA – Inteligência Artificial

IAG – USP – Instituto de Astronomia, Geofísica e Ciências Atmosféricas –
Universidade de São Paulo

IBAMA – Instituto Brasileiro do Meio Ambiente dos Recursos Naturais Renováveis

IBGE – Instituto Brasileiro de Geografia e Estatística

IRM – Índice de Risco Meteorotrópico

IRM 1 – Morbilidade respiratória segundo a faixa etária e sexo com uma Baixa
Frequência, em relação ao valor esperado

IRM 2 – Morbilidade respiratória segundo a faixa etária e sexo com uma
Frequência Normal, em relação ao valor esperado

IRM 3 – Morbilidade respiratória segundo a faixa etária e sexo com uma
Frequência Acima do Normal, em relação ao valor esperado

IRM 4 – Morbilidade respiratória segundo a faixa etária e sexo com uma
Frequência de Alerta, em relação ao valor esperado

LOG – Função de activação logística

MCP – Modelo de McCulloch & Pitts

MP₁₀ – Material Particulado

MUNIC_MO – Município de localização do hospital

n – Número de entradas do neurónio

NCEP – National Center Environmental Predictor

Net – Entrada líquida do neurónio

NO – Óxido Nítrico

NO₂ – Dióxido de Azoto

NO_x – Óxidos de Azoto

O₃ – Ozono

PMC – Percepção de Multicamadas

PoF – Sistema Pós-Frontal

PrF – Sistema Pré-Frontal

Rad – Radiação solar

RESP1F – Morbilidade por doenças respiratórias em crianças de 0 a 13 anos de idade do sexo feminino

RESP1M – Morbilidade por doenças respiratórias em crianças de 0 a 13 anos de idade do sexo masculino

RESP1MF – Morbilidade por doenças respiratórias em crianças de 0 a 13 anos de idade do sexo masculino e feminino

RESP2F – Morbilidade por doenças respiratórias em adolescentes de 14 a 19 anos de idade do sexo feminino

RESP2M – Morbidade por doenças respiratórias em adolescentes de 14 a 19 anos de idade do sexo masculino

RESP2MF – Morbidade por doenças respiratórias em adolescentes de 14 a 19 anos de idade do sexo masculino e feminino

RESP3F – Morbidade por doenças respiratórias em adultos jovens de 20 a 44 anos de idade do sexo feminino

RESP3M – Morbidade por doenças respiratórias em adultos jovens de 20 a 44 anos de idade do sexo masculino

RESP3MF – Morbidade por doenças respiratórias em adultos jovens de 20 a 44 anos de idade do sexo masculino e feminino

RESP4F – Morbidade por doenças respiratórias em adultos maduros de 45 a 64 anos de idade do sexo feminino

RESP4M – Morbidade por doenças respiratórias em adultos maduros de 45 a 64 anos de idade do sexo masculino

RESP4MF – Morbidade por doenças respiratórias em adultos j maduros de 45 a 64 anos de idade do sexo masculino e feminino

RESP5F – Morbidade por doenças respiratórias em idosos com ou mais de 65 anos de idade do sexo feminino

RESP5M – Morbidade por doenças respiratórias em idosos com ou mais de 65 anos de idade do sexo masculino

RESP5MF – Morbidade por doenças respiratórias em idosos com ou mais de 65 anos de idade do sexo masculino e feminino

RESP6F – Morbidade por doenças respiratórias envolvendo todas as faixas etárias do sexo feminino

RESP6M – Morbidade por doenças respiratórias envolvendo todas as faixas etárias do sexo masculino

RESP6MF – Morbilidade por doenças respiratórias envolvendo todas as faixas etárias do sexo masculino e feminino

RMSP – Região Metropolitana de São Paulo

RN – Rede Neuronal

RNA – Rede Neuronal Artificial

RNA's – Redes Neurais Artificiais

RNRG – Redes Neurais com Regressão Generalizada

SC – Condição Sinóptica

SEADE – Sistema Estadual de Análise de Dados

SO₂ – Dióxido de Enxofre

SUS – Sistema Único de Saúde

TGH – Função de activação tangente hiperbólica

TMG – Time Mean Greenwich

Tmin – Temperatura mínima

Uemin – Humidade específica mínima

wv – Velocidade do vento

w_{ij} – Peso da conexão entre os neurónios i e j

CAPÍTULO I: Introdução

A poluição do ar tem sido um tema extensamente pesquisado nas últimas décadas e representa um factor de grande importância na busca da preservação do meio ambiente e na implementação de um desenvolvimento sustentável, pois os seus efeitos afectam de diversas formas a saúde humana, os ecossistemas e os materiais.

A Região Metropolitana de São Paulo (RMSP) é uma área prioritária, já que apresenta uma forte degradação da qualidade do ar, condição característica da maior parte dos grandes centros urbanos. Os poluentes presentes na atmosfera da RMSP estão principalmente relacionados a grande emissão proveniente dos veículos automotores leves e pesados e secundariamente através de emissões originadas em processos industriais.

Os seres vivos reagem e adaptam-se ao meio ambiente atmosférico através de complexos mecanismos bioquímicos, cuja compreensão é o objecto de um moderno e importante ramo da ciência do tempo: a biometeorologia. O meio ambiente atmosférico inclui uma série de factores cuja relação com organismos deve ser feita de início em separado, ainda que, existam inter-relações que multiplicam ou reduzem os efeitos observáveis. Do mesmo modo, verificou-se que, em muitas ocasiões, tendem a aparecer no homem epidemias ou doenças agudas com acessos quase periódicos. A uma doença generalizada, que aparece ou desaparece consoante as mudanças atmosféricas, denomina-se doença meteorotrópica (Deryapa, 1986).

A evidência estatística demonstra que os tipos de condições do tempo têm uma correlação directa com a saúde do homem. Assim, a passagem de frentes e a mudança de massas de ar induzem fenómenos meteorotrópicos facilmente corroboráveis. Por outro lado, as substâncias de origem industrial, existentes no ar contaminado das cidades, produzem doenças da mais variada ordem, se bem

que se possa fazer referência aos níveis de poluição, nesse caso, só são quantificáveis os acontecimentos patológicos extremos (Persinger, 1980).

A saúde dos seres humanos não é influenciada somente pelas condições do clima e do tempo, mas também pela maneira como o habitat afecta as variáveis climáticas tais como a temperatura, a velocidade do vento e a humidade do ar. As actividades adaptáveis da fisiologia humana, tais como suar, respirar, locomoção e o grau de conforto são determinados em sua grande parte por factores do clima e pelas condições do tempo, entretanto, em ambientes fechados, os factores podem ser modificados. O aquecimento e o condicionamento de ar podem influenciar consideravelmente as circunstâncias atmosféricas como os seres humanos estão expostos. Os estilos de vida individuais, os hábitos de vestimentas e as circunstâncias ocupacionais influenciam também os níveis de exposição das pessoas que compartilham o mesmo habitat (Jendritzky & Sievers, 1989).

O tempo age directamente na saúde humana e, também, indirectamente, na sobrevivência através de seus efeitos em ecossistemas, ciclos hidrológicos (fonte de água), espécie de alimentos, agentes e vectores de doenças (Kalkstein & Smoyer, 1993).

Dentro de determinadas escalas de tolerância, a biologia humana pode avaliar a maioria das variações no clima, se estas se relacionam com a taxa de mudança ou grau de mudança. Mas as flutuações a curto prazo marcadas no tempo podem causar efeitos agudos adversos, indicados frequentemente pelo aumento nas taxas de mortalidade, admissões hospitalares e por aumentos no número de reclamações de indivíduos com estresse mental tal como a depressão (Jones *et al.*, 1982).

Desde que os seres humanos respondam a todos impactes dos elementos meteorológicos a que são expostos simultaneamente, combinações de medidas podem ser empregadas para se avaliar os efeitos da totalidade do tempo na saúde, quando comparados com uma série de variáveis componentes do tempo tratadas separadamente (Kalkstein & Davis, 1989).

Os distúrbios respiratórios são provocados tipicamente por reacções alérgicas, infecções, inalações de poeiras ou produtos químicos, e podem ser influenciados pelo tempo e pelo clima, directamente, através de quedas súbitas na temperatura e, indirectamente, através do aumento em níveis de poluentes (Schwartz *et al.*, 1990).

A poluição atmosférica na Cidade de São Paulo responde por um número significativo de internamentos. Actualmente, ela é responsável pelo agravamento de doenças pulmonares e portadores de moléstias cardíacas. Os que sofrem de patologias crónicas, como a asma, apresentam grandes dificuldades no seu tratamento (Böhm *et al.*, 1989).

1.1. Motivação

Vários estudos (secção 1.3) comprovam que o impacte das mudanças climáticas sobre os seres humanos pode provocar várias reacções adversas à saúde, principalmente, em crianças e idosos, que habitam grandes centros urbanos. Além disso, as condições do tempo actuam de modo favorável ou desfavorável para a dispersão de poluentes.

Conde (2001) realizou um estudo biometeorológico na Cidade de São Paulo. Os resultados obtidos, utilizando métodos estatísticos multivariados de Análise em Componentes Principais (ACP) e a Taxonomia Numérica ou Análise de "Cluster" (AC), mostraram que o impacte das variáveis ambientais, no caso variáveis meteorológicas e poluentes, assim como o conhecimento das condições sinópticas na região, estão fortemente associadas com as admissões hospitalares por doenças respiratórias em crianças, na Cidade de São Paulo.

A partir deste estudo verificou-se a importância da construção de uma metodologia baseada na aplicação de um índice que auxiliasse os prognósticos dos internamentos hospitalares causados por doenças do aparelho respiratório na Cidade de São Paulo.

De entre as mais diversas técnicas adequadas a este tipo de aplicação, destacam-se os modelos de redes neuronais artificiais (ver Capítulo 2) como uma alternativa interessante na construção de tecnologias de suporte à decisão. Redes neuronais artificiais apresentam características específicas, intrínsecas à

sua formulação, que possibilitam a aproximação de qualquer função contínua, funções não lineares complexas, e também podem generalizar os resultados obtidos para dados previamente desconhecidos, ou seja, produzir respostas coerentes e apropriadas para padrões ou exemplos que não foram utilizados em seu treino (Fausset, 1994; Haykin, 1994; Bishop, 1995).

Desta forma, a metodologia desenvolvida nesta pesquisa visa a aplicação de um sistema (meteorológico) de monitoramento das condições atmosféricas e poluentes sobre as admissões por doença respiratória nos hospitais da Cidade de São Paulo (CSP), através da criação de um Índice de Risco Meteorotrópico (IRM). O IRM permitirá estabelecer um sistema de avisos a centros de decisão competentes nas áreas da saúde e da protecção civil, com a finalidade de determinar prognósticos que viabilizem informações em termos de prevenção para a população.

1.2. Objectivo

Conforme abordado na justificação, este trabalho tem como objectivo o desenvolvimento de uma metodologia que viabilize a criação de um Índice de Risco Meteorotrópico (IRM), levando em consideração as consequências das variáveis ambientais (poluentes e meteorológicas) sobre a variabilidade da morbidade por todas as doenças do aparelho respiratório (conforme, o Código de Doenças Internacionais, CID 9 e CID 10), na Cidade de São Paulo, durante o período de 1993 a 2002.

A elaboração do IRM será efectuada através do emprego das Rede Neurais Artificiais (RNA's), considerando a classificação etária e o sexo do paciente. Além das RNA's, foram utilizadas técnicas de análises multivariadas de Análise de Cluster (AC) e Análise em Componentes Principais (ACP). A AC foi utilizado em carácter exploratório para verificar a similaridade das variáveis envolvidas e a ACP em carácter confirmatório. Essas duas técnicas corroboraram as informações obtidas pela rede neuronal.

1.3. Revisão Bibliográfica

Os primeiros casos descritos de poluição excessiva que causou um aumento do número de mortes ocorreram em 1930, no vale do rio Meuse na

Bélgica (Fircket, 1931). Esta região fica localizada entre as cidades de Huy e Liège, apresentando uma grande concentração de indústrias, muitas utilizando carvão, além de muitas outras produtoras de material particulado. Nos cinco primeiros dias de Dezembro, ocorreu um bloqueio anticiclónico, apresentando ausência de chuva, vento e inversão térmica que impediu a dispersão dos poluentes na região. Resultado: 60 mortes até dois dias após o evento, além de centenas de casos de internamentos por problemas respiratórios.

O caso mais clássico, entretanto, ocorreu em Londres, no Inverno de 1952, (Logan, 1953; Munn, 1968). Durante três dias, um outro bloqueio anticiclónico gerou um “smog” (fumaça + névoa) bastante intenso, composto principalmente por material particulado e enxofre (com concentrações de até nove vezes superior a média daquele período). Neste caso, o número de mortes atingiu 4.000 acima da média mensal normalmente esperada. Munn (1968) cunha o termo meteorotrópico, proveniente deste evento. Um efeito meteorotrópico significa o efeito que um parâmetro ou conjunto afectariam directamente um sistema vivo (um espécime) ou um conjunto (uma espécie).

Höppe (1997) ressaltou que o estudo da biometeorologia é considerado uma ciência antiga: desde o tempo de Hipócrates, na Grécia antiga, existia a influência de mudanças de tempo em processos fisiológicos no corpo humano. Entretanto, com o progresso em estatísticas modernas, a física e a fisiologia no transcurso deste século forneceram métodos quantitativos que fizeram com que a biometeorologia humana se tornasse uma ciência natural reconhecida.

Os métodos modernos da biometeorologia humana são reconhecidos por pesquisadores nas disciplinas de aplicação potencial, tais como planeadores ou coordenadores urbanos ou regionais do condicionamento de ar. A biometeorologia humana é um ramo da ciência que envolve a meteorologia e medicina ambiental, avaliando todas as influências atmosféricas, incluindo o padrão da poluição do ar, ou seja, a biometeorologia humana é a ciência das influências do ambiente atmosférico sobre o homem, que podem ser térmicas, hídricas (em relação a humidade) e eléctricas, além da composição do ar ambiente.

A biometeorologia humana, frequentemente chamada meteorologia médica, é considerada uma subdisciplina da meteorologia. Ao contrário de outras subdisciplinas da meteorologia, como por exemplo, a micrometeorologia ou a sinóptica, uma colaboração interdisciplinar com cientistas de outros campos como biologia e medicina são essenciais para a biometeorologia humana, especialmente quando os efeitos no corpo humano são observados mais de perto. Na figura 1, apresenta-se um esquema da classificação da biometeorologia humana do ponto de vista actual.

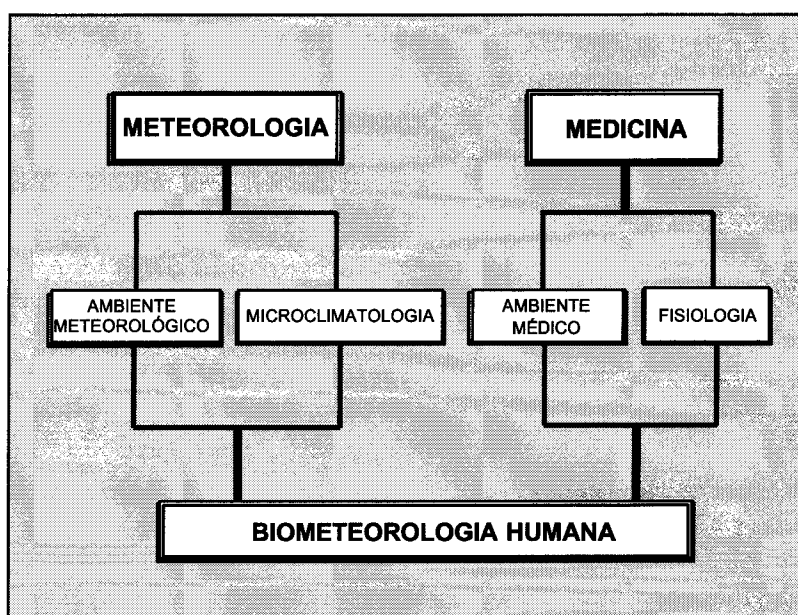


Figura 1.1 – Esquema da classificação da biometeorologia humana e ciências adjacentes do ponto de vista actual (Höppe, 1997).

A doença respiratória mais frequente é a inflamação. Durante os episódios de poluição, quando a CETESB (Companhia de Tecnologia de Saneamento Ambiental) avisa que o índice de qualidade do ar é ruim, muitas pessoas apresentam sintomas de ardência nos olhos, nariz, garganta, traqueia e, por vezes, tosse. A inflamação é uma das formas com que os tecidos reagem perante irritantes químicos, físicos ou microrganismos. Enquanto agudas e passageiras, as inflamações não são alterações preocupantes, entretanto, se crónicas, transformam-se em doenças que poderão agravar o quadro clínico do indivíduo. A mais grave das inflamações é a pneumonia, que é uma doença grave que

necessita de socorro médico; em crianças e idosos, podem acarretar em óbito (Böhm *et al.*, 1989).

O aprimoramento de técnicas de análise estatística de séries temporais, em estudos do tipo ecológico, predominante quando se analisa mortalidade e sua associação com poluentes, confere confiabilidade aos resultados (Pope, 1992; Schwartz, 1996).

Assim como a mortalidade pode ser um marcador dos efeitos da poluição do ar sobre a saúde, parâmetros de morbidade também o podem, visto que pessoas levadas à morte devem ter apresentado toda uma história de alterações clínicas anteriormente. Estas alterações têm sido documentadas na forma de exacerbações de sintomas respiratórios e cardiovasculares, aumento das crises de asma, dor precordial, limitação funcional, maior utilização de medicamentos, número de consultas em pronto-socorro e internamentos hospitalares (Braga, 1998).

Um estudo sobre atendimentos de crianças no pronto-socorro do Instituto da Criança da Faculdade de Medicina da USP, mostrou um aumento maior que 20%, nos dias mais poluídos, nas admissões hospitalares. Verificou-se uma forte associação com o PM₁₀ e, em menor intensidade, com o O₃, sugerindo que a poluição atmosférica representa um factor de risco importante na saúde da população infantil em São Paulo (Lin, 1997).

Setzer *et al.* (1980), Manfredini (1988), Abreu (1993) e Sanchez-Ccoyllo (1997) estudaram a relação entre sistemas sinópticos – meteorológicos e índices extremos de poluição para a RMSP. Mostraram que altas concentrações de poluentes ocorrem quando a cidade está sob domínio de um anticiclone (polar, subtropical do Atlântico Sul e continental), ou seja, esses sistemas favorecem a condição de estagnação atmosférica. Enquanto que baixas concentrações dos poluentes estão associadas a sistemas frontais.

Kalkstein (1991) avaliou o impacto do clima sobre a mortalidade humana em um período de 10 verões em St. Louis – USA. Utilizando dados de concentrações de poluentes, meteorológicos, condições sinópticas e mortalidade verificou através de análise em componentes principais e análise de Cluster que

em períodos consecutivos de dias quentes (8 a 10 dias) foi observado um aumento na mortalidade quando comparados com as concentrações dos poluentes que não ultrapassaram os padrões de qualidade do ar.

Andrade *et al.* (1994), estudaram a relação entre fontes de aerossóis e parâmetros meteorológicos durante o Inverno de 1989 no Campus da Universidade de São Paulo, onde foram realizadas a colecta de partículas atmosféricas em fracções finas e grossas. As amostras depois de analisadas foram submetidas ao método de Análises em Componentes Principais (ACP). Através da ACP foram identificadas cinco fontes para as partículas finas: emissões industriais, no qual explicava 13% da massa fina; emissões de óleo residual e diesel, explicando 41%; ressuspensão de poeira do solo com 28%; e emissões de cobre e manganês com 18%.

Para as partículas grossas identificaram-se quatro fontes: poeira do solo, explicando 59% da massa grossa; emissões industriais com 19%; queima de óleo com 8%; e aerossol marinho (Na e Cl) com 14%. Através da ACP observou-se que os parâmetros meteorológicos foram altamente correlacionados com as partículas finas e grossas (fontes de poeiras do solo), com excepção, da componente da brisa marítima que apresentou uma correlação negativa. As partículas finas e grossas (componentes industriais) tiveram também uma alta correlação positiva com os parâmetros meteorológicos, sendo mais correlacionada a direcção nordeste do vento e com pouca correlação a componente da brisa marítima.

Braga (1998) estudou a quantificação dos efeitos da poluição do ar na cidade de São Paulo e a morbidade por causas respiratórias em crianças menores de 13 anos, realizando análises de modelos de regressão linear múltipla e de regressão de Poisson. Observou que os poluentes PM₁₀, O₃ e CO apresentaram associações estaticamente significantes com a morbidade. Estas associações foram robustas para resistir à inclusão de variáveis de controlo. Os resultados encontrados no estudo avaliaram a proposta de uma estrutura de monitorização do comportamento da poluição do ar e seus efeitos sobre a saúde.

Leighton & Spark (1997) verificaram que em dias elevados e moderados de poluição na área de Sydney, eram maiores no início até a metade da década de

80 quando comparados com o fim da década de 80 e início da década de 90. A agência de protecção ambiental New South Wales (EPA) pensou que isto fora causado por uma diminuição dos poluentes emissores devido a controlos de emissões mais rígidos. Entretanto; uma avaliação do potencial previsto da poluição mostrou uma diminuição similar. A pergunta levantou-se consequentemente a respeito de que extensão a diminuição observada da poluição era uma função das situações sinópticas e se havia alguma mudança a longo prazo identificável nesses padrões sinópticos. Os resultados verificados foram:

- ✓ A identificação das principais situações sinópticas que cobriram a região australiana oriental foi associada com os eventos médios a elevados da poluição em Sydney. Os principais contribuintes que conduziram para eventos significativos de poluição foram as situações sinópticas em que uma corrente de ar de noroeste (leve a moderada) actuava sobre a região de Sydney.
- ✓ Relações entre o número de dias de poluição e a quantidade de vezes que os anticiclones ficavam situados em uma região (“anticlonicity”), mostraram anomalias de “anticlonicity” por 4 meses (uma para cada estação), onde um grande número de dias de poluição foi registrado.

Existiram tendências nos padrões sinópticos desde 1980 que contribuíram parcialmente para uma diminuição dos níveis dos poluentes observados sobre Sydney, mas essas tendências não puderam explicar toda a diminuição na poluição observada. Os padrões sinópticos para cada ano foram altamente variáveis, ocorrendo anos com situações sinópticas mais favoráveis à poluição do ar sobre Sydney para o período 1965-1993. Entretanto se os poluentes forem contidos em níveis actuais, a poluição actual observada deve permanecer abaixo das concentrações verificadas no início da década de 1980.

Panero *et al.* (1997) descreveram e analisaram as tendências temporais de observações horárias e médias diárias dos poluentes CO, NO₂, NO e SO₂ na área urbana da cidade de Salamanca para o período correspondido de Dezembro de 1993 a Novembro de 1994. Usando a série diária que correspondia a diferentes variáveis meteorológicas colectadas no mesmo local e período (direcção do vento, humidade relativa, pressão, radiação, temperatura e velocidade do vento),

aplicaram as técnicas de regressões lineares que foram usadas para estudar as relações dessas variáveis com cada um dos poluentes em uma base individual e multivariada. Os resultados obtidos explicaram 40% da variância da série para o NO₂ e 66% no caso do SO₂.

Nganga & Ngugi (1986) fizeram um estudo sobre a influência do tempo em duas das mais comuns doenças causadas em crianças na cidade de Nairobi para pacientes que sofriam de sarampo e pneumonia. Utilizaram uma análise de correlação com os registros diários do hospital nacional de Kenyatta e parâmetros meteorológicos tais como; temperatura do bulbo seco, insolação, humidade relativa, temperatura média e precipitação para um período de dois anos. Os autores mostraram que a maioria dos pacientes foram internados durante o período de Maio a Setembro, os meses frios, enquanto que, durante o período seco, Dezembro a Fevereiro, a frequência era mais baixa. Os coeficientes de correlação entre os elementos meteorológicos considerados e o número dos pacientes mostraram uma correlação negativa com a insolação, temperatura do bulbo seco e temperatura média e, uma correlação positiva com a humidade relativa do ar e a precipitação, que apresentaram registros elevados durante o período de estudo.

Um dos problemas principais de estudos biometeorológicos é a avaliação do efeito comparativo do clima em uma pessoa saudável e uma doente, revelando o mecanismo do impacto e, nesta base, o desenvolvimento de medidas para prevenir doenças e fortalecer a saúde humana. Deve-se levar em consideração que tanto o tempo como o clima actuam simultaneamente, afectando as pessoas de modo diferente, dependendo da capacidade de aclimação, sexo, idade, condição de saúde e outros factores (Rusanov, 1986).

Jauregui (1986) verificou a incidência de doenças respiratórias e níveis de poluição do ar na Cidade do México para o período 1975-1983. Observou que a frequência para os valores elevados da morbidade devido às doenças respiratórias agudas, no Inverno, coincide com os níveis elevados da poluição, enquanto que, a qualidade do ar melhora durante a estação húmida, a morbidade de doenças respiratórias declina.

As avaliações sinópticas são usadas frequentemente por biometeorologistas para caracterizar as condições meteorológicas e tipos de massas de ar em uma região. Essas avaliações combinam uma série complexa de variáveis meteorológicas em uma combinação característica (Perry, 1983; Yarnal, 1993).

As avaliações sinópticas foram usadas para identificar a periodicidade de sistemas atmosféricos em estudos feitos na Europa e na América do Norte, incluindo estudos de densidades do pólen na atmosfera, admissões hospitalares por asma, estresses biológicos, mortalidade humana, impacto do clima e da mudança do clima na poluição do ar (Wax, Borengasser & Muller, 1978; Balling, 1984; Muller & Jackson, 1985; Antó & Sunyer, 1986; Sanchez *et al.*, 1990; Kalkstein, 1991a; Schwartz, 1991; Bucher & Haase, 1993).

McGregor *et al.* (1999) utilizaram a climatologia sinóptica para investigar a relação entre os tipos de massa de ar actuaentes com média diária das concentrações do material particulado (PM_{10}) e admissões respiratórias hospitalares para a região de Birmingham, Inglaterra. O período correspondeu ao Inverno que possui em média, as maiores taxas de admissões respiratórias hospitalares. A metodologia correspondeu a dois passos, primeiro distinguiu os tipos de massas de ar, utilizando os padrões meteorológicos durante os Invernos (Novembro - Fevereiro) de 1988-1989 a 1993-1994 e o segundo, caracterizou os tipos de massa de ar em termos de todas as admissões respiratórias e as variáveis de qualidade do ar. O procedimento usado para a identificação dos tipos de massas foi através da Análise de Componente Principal (ACP) e Análise de Cluster (AC). Os resultados mostraram respostas distintas de taxas de admissão respiratória para os tipos de massas de ar identificadas para os Invernos. Dois das três massas de ar associadas com as taxas de admissão acima da média (anticiclone continental frio sem nebulosidade e o anticiclone continental encoberto) favorecem níveis elevados de PM_{10} e sugere uma possível ligação entre as condições de tempo, qualidade do ar e saúde, enquanto que, o tipo de massa de ar de característica marítima húmida e fria não favorece níveis elevados de PM_{10} . Assim, o mecanismo proposto é susceptível para responder os acoplamentos entre tipos de massa de ar, concentrações de PM_{10} e a resposta respiratória.

As RNA's são amplamente empregadas em diversos campos da ciência, entre eles, o reconhecimento de padrão e classificação, processamento de imagem e sinal, geografia (Spellman, 1999), previsão financeira (Tkacz, 2001), diagnósticos médico (Ellenius e Groth, 2000).

Na área da meteorologia pode-se citar alguns estudos de RNA's no prognóstico de precipitação (Hall, T., 1998; Silverman & Dracup, 1999; Cigizoglu, 2003; Michaelides *et al.*, 2001); no prognóstico de concentrações de poluentes (Benvenuto & Marani, 2000; Lu *et al.*, 2002; Cannon & Lord, 2000; Viotti *et al.*, 2002). Na área médica, no diagnóstico de doenças respiratórias e cardíacas (Bibi *et al.*, 2002; Hirsch *et al.*, 1997; Yana *et al.*, 2006), entre outros.

Nazir *et al.* (2002) desenvolveram um método rápido e seguro para prognosticar o padrão de deposição de partículas de aerossol no interior dos pulmões humanos, usando redes neuronais artificiais (RNA's). Foram usados dados experimentais da literatura para o treino da rede multicamadas perceptron (RMP's) para permitir o prognóstico regional e total do padrão de deposição das partículas de aerossol. Estes dados cobriram tamanhos de partícula que variam de 0.05-15 μm e três padrões diferentes de respiração (variando de respirar "calmo" a "sob de condições de trabalho físicas"). Foram treinadas três diferentes RMP's, para prover predições separadas da deposição de partícula de aerossol no laríngeo, regiões brônquicas e alveolares. As frações de deposição total para uma determinada condição de respiração são computadas com a soma das produções gerada da deposição regional correspondente a RMP's. Os resultados indicaram que as RNA's mostraram prognósticos altamente precisos para o padrão regional e total da deposição de partículas para todos os tamanhos de partículas e condições de respiração, com erros menores que 0.04%.

Abdul-Wahab & Al-Alawi (2002) desenvolveram o uso de uma rede neuronal para prognosticar as concentrações de ozono na baixa atmosfera em função das condições meteorológicas e vários parâmetros de qualidade de ar.. Os resultados indicaram a dependência da concentração de ozono com os demais poluente e com as condições meteorológicas. A contribuição da meteorologia nas variações da concentração de ozono foi dentro da escala de 33.1 – 40.6% e a variabilidade restante foi atribuída a poluente químicos como óxido de nitrogênio,

dióxido de enxofre, hidrocarboneto e dióxido de nitrogênio. A humidade relativa é o parâmetro meteorológico com a contribuição mais elevada para variações de ozono, além da influência da temperatura e da radiação solar. Os autores concluíram que a rede neuronal pode ser usada em modelamento e prognóstico das concentrações do ozono em superfície, indicando o potencial da aproximação de rede neural por capturar as interações não lineares entre ozono e outros factores e, para a identificação da importância relativa destes factores.

Miraglia *et al.* (2005) estimaram os custos do impacto da poluição do ar na saúde em São Paulo, utilizando o Método de Anos de Vida Ajustados (DALY). Os resultados mostraram que 28,2 anos de vida perdidos e vividos com inaptidão demonstram o perigo da poluição do ar. O risco para crianças, 12,3 anos, indicou que a poluição afecta vidas prematuras antes de alcançar a idade de produtividade. Os danos para os idosos, 15,9 anos, é devido a susceptibilidade dos idosos a doenças respiratórias e cardiovasculares que são agravadas através do poluente e pode ser agravado com situações de eventos de poluição severos, com o DALY aumentar substancialmente. Outro resultado importante recorre à redução de probabilidade de vida devido a poluição do ar: 4 anos (considerando doenças das vias respiratórias), 10 anos (considerando doenças cardiovasculares), e 19 anos (doenças das vias respiratórias de crianças) na probabilidade de vida anciã. A magnitude dos resultados reconhece o potencial desta ferramenta, aumentar a manutenção de qualidade de saúde pública e encorajar recurso, poupanças, uma vez que permite a predição de ganhos de saúde devido a melhorias ambientais.

Neste estudo será dado ênfase ao Material Particulado (MP₁₀), dióxido de enxofre (SO₂), ozono (O₃), monóxido de carbono (CO) e dióxido de nitrogênio (NO₂). Esse grupo de poluentes que servem como indicadores de qualidade do ar está ligado à sua maior frequência de ocorrência e aos efeitos adversos que causam sobre a saúde da população e ao meio ambiente em grandes centros urbanos como a Cidade de São Paulo.

➤ Material Particulado

O material particulado é uma mistura de partículas líquidas e sólidas em suspensão no ar, excepto a água pura existente, sob condições normais de temperatura e pressão atmosférica e que tenha dimensões microscópicas ou sub-microscópicas, com impacte relevante no tempo e no clima. Sua composição e tamanho dependem das fontes de emissão. O tamanho das partículas é expresso em relação ao seu tamanho aerodinâmico, definido como o diâmetro de uma esfera densa que tem a mesma velocidade de sedimentação que a partícula em questão (Dockery, 1992; Pope, 1991).

O material particulado pode ser tanto de origem natural como antropogénica, sendo o primeiro por actividades vulcânicas, natureza abiótica, “*spray*” marinho, além da entrada de meteoros na atmosfera terrestre e o segundo por queima de combustíveis fósseis, com transformações sofridas na própria atmosfera, de certos gases em partículas.

Em geral, as partículas podem ser divididas em dois grupos:

- ◆ Partículas grandes, com diâmetro ($d \leq 10 \mu\text{m}$, adoptado pela CETESB), também chamadas de moda grossa. Essas partículas ficam retidas na parte superior do sistema respiratório, sendo derivadas das combustões descontroladas dos veículos automotores, dispersão mecânica do solo, materiais da crosta terrestre e *spray* marinho;
- ◆ Partículas derivadas da combustão de fontes móveis e estacionárias, como automóveis, incineradores e termoeléctricas, em geral, são de menor tamanho, apresentando diâmetro menor que $2,5 \mu\text{m}$, também chamado de moda fina e têm maior acidez, podendo atingir as porções mais distais do tracto respiratório, prejudicando as trocas gasosas (Bascom *et al.*, 1996).

Os efeitos adversos do material particulado na atmosfera começam pelos aspectos estéticos, pois este interfere na visibilidade e está associado à produção de corrosão e sujeira em superfícies (edifícios, tecidos e outros materiais). Os efeitos sobre a saúde estão associados à:

- Capacidade do sistema respiratório remover as partículas no ar inalado, retendo-as nos pulmões;
- Presença nas partículas de substâncias minerais que possuem propriedades tóxicas;
- Presença nas partículas de compostos orgânicos, como os hidrocarbonetos policíclicos, que possuem propriedades carcinogênicas;
- Capacidade das partículas finas de aumentar os efeitos fisiológicos de gases irritantes também presentes no ar ou de catalisar e transformar quimicamente estes gases, criando espécies mais nocivas.

A capacidade do material particulado de aumentar os efeitos fisiológicos dos gases presentes no ar é um dos aspectos mais importantes a ser considerado. Os efeitos de uma mistura de material particulado e dióxido de enxofre, por exemplo, são mais acentuados que os provocados na presença isolada de cada um deles.

Estudos observacionais, comparando as variações dos níveis ambientais dos poluentes e das consultas de pronto-socorro e internamentos hospitalares, são plenos em exemplos que mostram serem os particulados directamente responsáveis por danos à saúde em crianças (Pope, 1991).

➤ **Dióxido de Enxofre (SO₂)**

Os gases sulfurosos presentes no ar podem causar uma série de distúrbios fisiológicos no homem, animais e plantas. A inalação de SO₂, que é um dos mais frequentes poluentes atmosféricos, mesmo em concentrações muito baixas, provoca espasmos passageiros dos músculos lisos dos brônquios pulmonares. O SO₂ é um gás incolor, cheiro desagradável e pode ser percebido pelo olfacto de algumas pessoas em concentrações que variam de 0,3 a 1 ppm (CETESB, 1993).

Os efeitos dos gases na saúde humana estão intimamente associados à sua solubilidade nas paredes do aparelho respiratório, facto estes que governa a

quantidade do poluente capaz de atingir as porções mais profundas do aparelho respiratório.

Existem evidências de que o dióxido de enxofre agrava as doenças respiratórias pré-existentes e também contribui para seu desenvolvimento, produzindo irritação no sistema respiratório, e absorvido em partículas pode ser conduzido mais profundamente, podendo produzir danos aos tecidos do pulmão.

Estudos epidemiológicos e clínicos mostram que certas pessoas são mais sensíveis ao dióxido de enxofre que outras. Exposições prolongadas a baixas concentrações de dióxido de enxofre têm sido associadas ao aumento de morbidade cardiovascular em pessoas idosas.

O SO₂ está frequentemente associado aos internamentos por doenças respiratórias em geral (Bates, 1983), especificamente por Doenças Pulmonares Obstrutivas Crônicas – DOPC (Sunnyer *et al.*, 1991), mesmo dentro dos padrões de segurança e, também, por asma e outras doenças.

Níveis baixos de SO₂ observados na Cidade do Porto, Portugal (Queirós *et al.*, 1990), não estão associados directamente com aumento das admissões em pronto-socorro ou de internamentos por asma, porém, os autores sugerem que, a longo prazo, estes níveis podem levar à diminuição do limiar a outros estímulos espásticos em crianças asmáticas.

➤ Ozono (O₃)

Oxidantes fotoquímicos são a denominação que se dá à mistura de poluentes secundários formados pela reacção dos hidrocarbonetos e óxidos de azoto na presença de luz solar.

O principal ingrediente desta mistura é o gás ozono (O₃), por isso mesmo utilizado como parâmetro indicador da presença dos oxidantes fotoquímicos, que têm em sua composição também quantidades pequenas de compostos oxigenados derivados dos hidrocarbonetos, derivados de fontes de combustão móveis, como os veículos automotores, fontes estacionárias, como usinas termoeléctricas, e até mesmo fontes naturais como as árvores, que contribuem na

produção de compostos orgânicos voláteis (National Research Council, 1991; Apud Bascom *et al.*, 1996).

O efeito mais relatado dos oxidantes fotoquímicos é a irritação dos olhos. Os principais componentes da mistura associados a este efeito são os peróxi-acitilnitratos (PAN - nitratos de peróxi-acitila e outros), o formaldeído e a acroleína (CETESB, 1993).

A presença dos oxidantes fotoquímicos na atmosfera tem sido associada à redução de capacidade pulmonar e ao agravamento de doenças respiratórias, como a asma. Estudos realizados em animais mostraram que o ozono causa o envelhecimento precoce, provoca danos na estrutura pulmonar e diminui a capacidade de resistir às infecções respiratórias (Koren *et al.*, 1989).

Mesmo pessoas saudáveis, como os atletas, têm se mostrados sensíveis aos efeitos do ozono pela diminuição da capacidade de executar exercícios físicos (Hazucha & Bates, 1989).

Os níveis de ozono aumentam consideravelmente entre o fim da primavera e o começo do Outono, em regiões periféricas de grandes centros urbanos, localizadas nas direcções em que sopram os ventos. Caracteristicamente, os seus picos ocorrem entre 12 e 15 horas da tarde e declinam à noite (CETESB, 1993).

O ozono é tóxico para o sistema respiratório, pois devido a sua baixa solubilidade em água, consegue atingir os alvéolos, assim como, para as plantas que prejudica o seu tecido, inibindo seu processo de fotossíntese. O poluente ozono, também, é gerado secundariamente a partir das emissões dos carros (Zarkzewski, 1997). De acordo com Boelter & Davidson (1997), uma exposição curta ao ozono pode causar ataque de tosse. Uma exposição longa pode causar mal-estar afectando muito o sistema respiratório. Uma exposição contínua ao O₃ pode também resultar em cancro no pulmão.

➤ **Monóxido de Carbono (CO)**

Os efeitos da exposição de seres humanos ao monóxido de carbono estão associados à capacidade de transporte de oxigénio pelo sangue. O monóxido de

carbono compete com o oxigénio na combinação com a hemoglobina do sangue, uma vez que a afinidade de hemoglobina pelo monóxido de carbono é cerca de 210 vezes maior que pelo oxigénio. Quando uma molécula de hemoglobina recebe uma molécula de monóxido de carbono forma-se a carboxi-hemoglobina, que diminui a capacidade do sangue de transportar oxigénio.

Os sintomas da exposição ao monóxido de carbono dependem da quantidade de hemoglobina combinada com monóxido de carbono. Tem sido demonstrado experimentalmente que baixos níveis de carboxi-hemoglobina já podem causar diminuição na capacidade de estimar intervalos de tempo e, também, diminuir os reflexos e a acuidade visual da pessoa exposta.

Com excepção dos fumantes, que possuem suas próprias fontes emissoras de CO, os demais habitantes dos grandes centros urbanos têm no trânsito intenso a sua maior fonte deste poluente. Pessoas que passam várias horas do dia dentro de um automóvel, ou que tenham que andar a pé ou de bicicleta são as mais afectadas (Bascom *et al.*, 1996b).

➤ **Dióxido de Azoto**

No caso dos óxidos de azoto (NO e NO₂), somente o NO₂ é motivo de preocupação, pois devido à sua baixa solubilidade, quando inalado, é capaz de penetrar profundamente no sistema respiratório, podendo dar origem as nitrosaminas, algumas das quais podem ser carcinogênicas. O dióxido de azoto (NO₂) é também um poderoso irritante, podendo conduzir o sintoma de enfisema (CETESB, 1993). Seu efeito tóxico está relacionado ao facto de ser um agente oxidante (Pryor & Lightsey, 1981).

As principais fontes de óxidos de azoto na Cidade de São Paulo são os motores dos automóveis, usinas termoeléctricas, incineradores e indústrias que utilizam combustíveis fósseis também contribuem, porém, em menor escala. Durante a combustão sob elevadas temperaturas, o oxigénio reage com o azoto formando o óxido nítrico (NO), dióxido de azoto (NO₂) e outros óxidos de azoto (NO_x). Esses compostos são extremamente reactivos e na presença de oxigénio (O₂), ozono e hidrocarbonetos, o NO se transforma em NO₂. Por sua vez, NO₂ na presença de luz solar, hidrocarbonetos e oxigénio se transforma em ozono (O₃),

sendo um dos principais precursores deste poluente na troposfera (CETESB, 2002). Os efeitos adversos desses poluentes tanto na saúde como ao meio ambiente são mostrados na tabela 1.

Tabela 1 - Efeitos dos poluentes (PM₁₀, SO₂, O₃, CO e NO₂) na atmosfera (Fonte: CETESB, 2002).

Poluente	Efeitos Gerais sobre a Saúde	Efeitos Gerais ao Meio Ambiente
MP ₁₀	Aumento de atendimentos hospitalares e mortes prematuras	Danos à vegetação, deterioração da visibilidade e contaminação do solo.
SO ₂	Desconforto na respiração, doenças respiratórias, agravamento de doenças respiratórias e cardiovasculares já existentes. Pessoas com asma, doenças crônicas de coração e pulmão são mais sensíveis ao SO ₂ .	Pode levar à formação de chuva ácida, causar corrosão aos materiais e danos à vegetação: folhas e colheitas.
O ₃	Irritação nos olhos e vias respiratórias, diminuição da capacidade pulmonar. Exposição a altas concentrações pode resultar em sensações de aperto no peito, tosse e chiado na respiração. O O ₃ tem sido associado ao aumento de admissões hospitalares	Danos às colheitas, à vegetação natural, plantações agrícolas; plantas ornamentais.
CO	Altos níveis de CO estão associados a prejuízo dos reflexos, da capacidade de estimar intervalos de tempo, no aprendizado, de trabalho e visual	
NO ₂	Aumento da sensibilidade à asma e à bronquite, abaixar a resistência às infecções respiratórias.	Pode levar à formação de chuva ácida, danos à vegetação e à colheita.

➤ Padrões de Qualidade do Ar

Um padrão de qualidade do ar define legalmente um limite máximo para a concentração de um componente atmosférico que garanta a protecção da saúde e do bem-estar das pessoas. Os padrões de qualidade do ar são baseados em estudos científicos dos efeitos produzidos por poluentes específicos e são fixados em níveis que possam propiciar uma margem de segurança adequada.

Os principais objectivos do monitoramento da qualidade do ar são:

- Fornecer dados para activar acções de emergência durante períodos de estagnação atmosférica quando os níveis de poluentes na atmosfera possam representar risco à saúde pública;

- Avaliar a qualidade do ar à luz de limites estabelecidos para proteger a saúde e o bem-estar das pessoas;
- Acompanhar as tendências e mudanças na qualidade do ar devidas a alterações nas emissões dos poluentes.

Para atingir estes objectivos, torna-se necessária a fixação de padrões de qualidade do ar.

Através da Portaria Normativa nº 348 de 14/03/90 o IBAMA (Instituto Brasileiro do Meio Ambiente dos Recursos Naturais Renováveis) estabeleceu os padrões nacionais de qualidade do ar, ampliando o número de parâmetros anteriormente regulamentados através da Portaria GM 0231 de 27/04/76. Os padrões estabelecidos através dessa portaria foram submetidos ao CONAMA em 28/06/1990 e transformados na Resolução CONAMA (Conselho Nacional do Meio Ambiente) nº 03/90. São estabelecidos dois tipos de padrões de qualidade do ar: os primários e secundários.

São padrões primários de qualidade do ar as concentrações de poluentes que, ultrapassadas, poderão afectar a saúde da população. Podem ser entendidos como níveis máximos toleráveis de concentração de poluentes atmosféricos, constituindo-se em metas de curto e médio prazo.

São padrões secundários de qualidade do ar as concentrações de poluentes atmosféricos abaixo das quais se prevê o mínimo efeito adverso sobre o bem-estar da população, assim como o mínimo dano à fauna e à flora, aos materiais e ao meio ambiente em geral. Podem ser entendidos como níveis desejados de concentração de poluentes, constituindo-se em meta de longo prazo.

O objectivo do estabelecimento de padrões secundários é criar uma base para uma política de prevenção da degradação da qualidade do ar. Deve ser aplicado a áreas de preservação como parques nacionais, áreas de protecção ambiental, estâncias turísticas, etc. Não se aplicam, pelo menos em curto prazo, a áreas de desenvolvimento, onde devem ser aplicados os padrões primários. Como prevê a própria Resolução CONAMA nº 03/90, a aplicação diferenciada de padrões primários e secundários requer que o território nacional seja dividido em

classes I, II e III conforme o uso pretendido. A mesma resolução prevê ainda que enquanto não for estabelecida a classificação das áreas os padrões aplicáveis serão os primários.

Levando em consideração os poluentes verificados neste trabalho, os padrões nacionais de qualidade do ar e critérios para episódios agudos de poluição do ar, segundo a resolução do CONAMA, são apresentados na tabela 2.

Os monitoramentos dos padrões de qualidade do ar são feitos pela CETESB, que possui uma rede automática de vinte e cinco estações fixas de amostragem, localizadas na RMSP (22 estações, sendo que destas, 13 estão localizadas na CSP), e em Cubatão (3 estações), e mais dois laboratórios móveis. Estas fornecem medidas de dióxido de enxofre, partículas inaláveis, monóxido de carbono, óxidos de azoto, hidrocarbonetos, direção e velocidade do vento, temperatura e humidade (CETESB, 1993).

Com o surgimento da indústria automobilística, no final dos anos cinquenta, a frota circulante tem aumentado de forma progressiva, apresentando hoje algo em torno de cinco milhões de veículos circulando na região metropolitana de São Paulo (CETESB, 2000). Frota esta composta de carros, caminhões e autocarros, muitos com mais de dez anos de fabricação e, portanto, desprovidos de mecanismos de controlo de emissão de poluentes.

Tabela 2 – Padrões nacionais de qualidade do ar e critérios para episódios agudos de poluição do ar (resolução CONAMA n° 3 de 8/06/1990. Fonte: CETESB, 2002).

Poluente	Tempo de Amostragem	Padrão Primário $\mu\text{g}\cdot\text{m}^{-3}$	Padrão Secundário $\mu\text{g}\cdot\text{m}^{-3}$	Método de Medição	Níveis		
					Atenção	Alerta	Emergência
Dióxido de Enxofre	24 horas	365	100	Pararosanilina	800	1.600	2.100
	MAA (2)	80	40				
Monóxido de Carbono	1 hora (1)	40.000 (35 ppm)	40.000 (35 ppm)	Infravermelho não dispersivo	15	30	40
	8 horas	10.000 (9 ppm)	10.000 (9 ppm)				
Ozono	1 hora (1)	160	160	Quimiluminescência	400*	800	1.000
Partículas Inaláveis	24 horas (1)	150	150	Separação Inercial/Filtração	250	420	500
	MAA (2)	50	50				
Dióxido de Nitrogênio	1 hora (1)	320	190	Quimiluminescência	1.130	2.260	3.000
	MAA (2)	100	100				

(1) Não deve ser excedido mais que uma vez ao ano.

(2) Média aritmética anual.

* O nível de atenção é declarado pela CETESB com base na legislação Estadual que é mais restritiva ($200 \mu\text{g}\cdot\text{m}^{-3}$).

➤ Características Geomorfológicas e Climáticas da Cidade de São Paulo

As características topográficas da RMSP onde está localizada a cidade têm influência marcante na qualidade do ar. Localizada em uma bacia sedimentar, compartimento rebaixado do Planalto Atlântico, cortado pelo Trópico de Capricórnio, apresenta colinas que variam de 650 a 1.200 m. Ao norte, está situada entre a Serra da Cantareira, e entre os rios Tietê, Pinheiros e Tamanduateí temos o Espigão Central (Espigão da Paulista), com altitudes de mais de 800 m, toda essa complexidade topográfica associada à proximidade do oceano e à intensa urbanização da área, influenciam o padrão da circulação atmosférica.

Além disso, a Cidade de São Paulo está situada entre os maiores conglomerados humanos do mundo, com uma população de aproximadamente

10,5 milhões de pessoas e um grande parque industrial, com uma enorme frota de veículos (IBGE, 2000).

O clima da cidade de São Paulo pode ser classificado como tropical húmido de altitude (Martyn, 1992), apresentando Invernos frios e secos, enquanto o verão é quente e húmido. As temperaturas variam de 12° C (média das temperaturas mínimas), até 28° C (média das temperaturas máximas), com média anual de 19° C (Estação Meteorológica IAG - SP). Sugahara (1991) mostrou que a pluviosidade média chega em torno de 1500 mm. O regime anual de precipitação no Estado de São Paulo é altamente sazonal, basicamente com verão chuvoso e Inverno seco. A estação chuvosa começa em Setembro ou Outubro e se estende até Abril.

CAPÍTULO II: Redes Neurais Artificiais – Fundamentos

Neste capítulo, apresentaremos os principais fundamentos relacionados às RNA's, e que serão importantes no desenvolvimento deste trabalho. Em especial, descreveremos com grande detalhe os algoritmos de retro propagação de erro e descida do gradiente conjugada utilizados no treino das redes neurais artificiais multicamadas.

2.1. Histórico das Rede Neurais Artificiais (RNA's)

As primeiras informações mencionadas sobre a neurocomputação datam de 1943, em artigos de McCulloch & Pitts, em que sugeriam a construção de uma máquina baseada ou inspirada no cérebro humano. Muitos outros artigos e livros surgiram desde então, porém, por um longo período de tempo, pouco resultado foi obtido. Até que em 1949 Hebb escreveu um livro intitulado "The Organization of Behavior" (A Organização do Comportamento) que perseguia a ideia de que o condicionamento psicológico clássico está presente em qualquer parte dos animais pelo fato de que esta é uma propriedade de neurónios individuais. Suas ideias não eram completamente novas, mas Hebb foi o primeiro a propor uma lei de aprendizagem específica para as sinapses dos neurónios. Este primeiro passo serviu de inspiração para que muitos outros pesquisadores perseguissem a mesma ideia. E embora muito tenha sido estudado e publicado nos anos que seguiram (1940-1950), estes serviram mais como base para desenvolvimento posterior que para o próprio desenvolvimento.

Também proveniente deste período de tempo foi a construção do primeiro neuro computador, denominado Snark, por Minsky em 1961. O Snark operava com sucesso a partir de um ponto de partida técnico, ajustando seus pesos automaticamente, entretanto, ele nunca executou qualquer função de

processamento de informação interessante, mas serviu de inspiração para as ideias de estruturas que o sucederam.

Em 1956 no "Darthmouth College" nasceram os dois paradigmas da Inteligência Artificial, a simbólica e o conexionista. A Inteligência Artificial Simbólica tenta simular o comportamento inteligente humano desconsiderando os mecanismos responsáveis por tal. Já a Inteligência Artificial Conexionista acredita que construindo-se um sistema que simule a estrutura do cérebro, este sistema apresentará inteligência, ou seja, será capaz de aprender, assimilar, errar e aprender com seus erros.

O primeiro neuro computador a obter sucesso (Mark I Perceptron) surgiu em 1958, criado por Rosenblatt *et al.*. Devido a profundidade de seus estudos, suas contribuições técnicas e de sua maneira moderna de pensar, muitos o vêem como o fundador da neurocomputação na forma actual. Seu interesse inicial para a criação do Perceptron era o reconhecimento de padrões.

Após Rosenblatt, Widrow (1962) desenvolveu um novo tipo de elemento de processamento de redes neurais chamado de Adaline, equipado com uma poderosa lei de aprendizado, que diferente do Perceptron ainda permanece em uso. Widrow também fundou a primeira companhia de hardware de neurocomputadores e componentes.

Nos anos seguintes foram marcados por um entusiasmo exagerado de muitos pesquisadores, que passaram a publicar artigos e livros que faziam uma previsão pouco confiáveis para a época, sobre máquinas tão poderosas quanto o cérebro humano que surgiriam em um curto espaço de tempo, proporcionando a perda da credibilidade nos estudos desta área.

Um período de pesquisa silenciosa seguiu-se durante 1967 a 1982, quando poucas pesquisas foram publicadas devido aos factos ocorridos anteriormente. Entretanto, aqueles que pesquisavam nesta época, e todos os que se seguiram no decorrer de treze anos conseguiram novamente estabelecer um campo concreto para o renascimento da área.

Nos anos 80, muitos dos pesquisadores passaram a publicar diversas propostas para a exploração de desenvolvimento de redes neuronais bem como suas aplicações. Porém, o facto mais importante deste período tenha ocorrido



quando Skurnick, um administrador de programas da DARPA (Defence Advanced Research Projects Agency) decidiu ouvir os argumentos da neurocomputação e seus projectistas, e divergindo dos caminhos tradicionais dos conhecimentos convencionais, fundou em 1983 pesquisas em neurocomputação. Este acto não só abriu as portas para a neurocomputação, como também deu à DARPA o status de uma das líderes mundiais em se tratando de "moda" tecnológica.

Outra "potência" que emergiu neste período foi Hopfield (1982), renomado físico de reputação mundial, se interessou pela neurocomputação, e escreveu artigos que percorreram o mundo todo persuadindo centenas de cientistas, matemáticos, e tecnólogos altamente qualificados a se unirem esta nova área emergente.

Apesar de um terço dos pesquisadores da área terem aderido à mesma pela influência de Hopfield, foi em 1986 que este campo de pesquisa "explodiu" com a publicação do livro "Parallel Distributed Processing" (Processamento Distribuído Paralelo) editado por Rumelhart & McClelland.

Em 1987 muitas universidades anunciaram a formação de institutos de pesquisa e programas de educação em neurocomputação, ocorrendo neste ano em São Francisco, a primeira conferência de redes neurais em tempos modernos, a IEEE International Conference on Neural Networks, em que foi formada a International Neural Networks Society (INNS). A partir destes acontecimentos decorreram a fundação do INNS journal em 1989, seguido do Neural Computation e do IEEE Transactions on Neural Networks em 1990.

Em síntese, segue abaixo os factos históricos em ordem cronológica:

- 1943: McCulloch & Pitts estabeleceram as bases da neurocomputação, com modelos matemáticos.
- 1949: Hebb traduziu matematicamente a sinapse dos neurônios biológicos.
- 1951: Minski constrói o Snark, primeiro neurocomputador com capacidade de aprendizado, ou seja, ajustava automaticamente os pesos entre as sinapses. Não executou nenhuma função útil.
- 1957: Rosenblatt concebeu o "perceptron", que era uma rede neural de duas camadas, usado no reconhecimento de caracteres.

- 1962: Widrow desenvolveu um processador para redes neurais e fundou a primeira empresa de circuitos neurais digitais, a Memistor Corporation.
- 1974: Werbos lançou bases para o algoritmo de retropropagação (backpropagation).

2.2. O Neurónio Biológico

Como as RNA's são baseadas na estrutura do cérebro (neurónio), serão mostradas as principais características de um neurónio biológico, para uma posterior compreensão, a respeito, do funcionamento das RNA's.

O neurónio biológico é uma célula que pode ser dividida em três secções: o *corpo da célula*, os *dendrites* e o *axónio*, cada uma com funções específicas, porém complementares.

O corpo do neurónio mede apenas alguns milésimos de milímetros e os dendrites apresentam poucos milímetros de comprimento. O axónio pode ser mais longo. Os dendrites têm como função, receber informações, ou *impulsos nervosos*, oriundas de outros neurónios, e conduzi-las até o corpo celular. Ali, a informação é processada e novos impulsos são gerados. Estes impulsos são transmitidos a outros neurónios (nodos), passando pelo axónio e atingindo os dendrites dos neurónios seguintes.

O ponto de contacto entre a terminação axónica de um neurónio e o dendrito de outro é chamado *sinapse*. É pelas sinapses que os nodos se unem funcionalmente, formando as redes neuronais. As sinapses funcionam como válvulas, sendo capazes de controlar a transmissão de impulsos, isto é, o fluxo da informação entre os nodos na rede neural. O efeito das sinapses é variável, é esta variação que dá ao neurónio capacidade de adaptação. Na figura 2.1 está ilustrada, de forma simplificada, os componentes do neurónio.

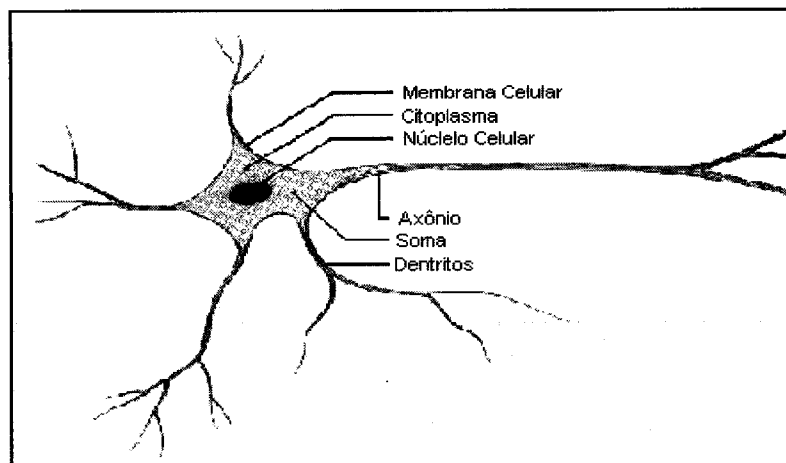


Figura 2.1 - Componentes de um Neurônio Biológico

Os sinais oriundos dos nodos *pré-sinápticos* são passados para o corpo do neurônio, onde são comparados com outros sinais recebidos pelo neurônio. Se o percentual em um intervalo curto de tempo é suficientemente alto, a célula “dispara”, produzindo um impulso que é transmitido para as células seguintes (nodos *pós-sinápticos*). Este sistema simples é responsável pela maioria das funções realizadas pelo nosso cérebro. A capacidade de solucionar funções complexas surge com a operação em paralelo de todos os 10^{11} nodos do nosso cérebro.

2.3. Neurônios Artificiais: O Modelo MCP (McCulloch & Pitts)

O modelo de neurônio proposto por McCulloch & Pitts (MCP) em 1943 é uma simplificação do que se sabia sobre o neurônio biológico naquela época. A sua descrição matemática resultou em um modelo com n terminais de entrada x_1, x_2, \dots, x_n (que representam os dendrites), e apenas um terminal de saída y (que representa o axônio). Para emular o comportamento das sinapses, os terminais de entrada do neurônio têm pesos acoplados w_1, w_2, \dots, w_n , cujo valores podem ser positivos ou negativos, dependendo das sinapses correspondentes serem inibitórias ou excitatórias. O efeito de uma sinapse particular i no neurônio pós-sináptico é dado por: $x_i w_i$. Os pesos determinam “em que grau” o neurônio deve considerar sinais de disparo que ocorrem naquela conexão. Uma descrição de um neurônio artificial é ilustrada na figura 2.2.

Um neurônio dispara quando a soma dos impulsos que ele recebe ultrapassa o seu limiar de excitação chamado de *threshold*. O corpo do neurônio,

por sua vez, é emulado por um mecanismo simples que faz a soma dos valores $x_i w_i$ recebidos pelo neurónio (soma ponderada). E decide se o neurónio deve ou não disparar (saída igual a 1 ou a 0) comparando a soma obtida ao limiar ou *threshold* do neurónio. No modelo MCP, a activação do neurónio é obtida através da aplicação de uma “função de activação”, que activa a saída ou não, dependendo do valor da soma ponderada das suas entradas. Na descrição original do modelo MCP, a função de activação é dada pela função limiar descrita a seguir. O nodo MCP terá então sua saída activa quando:

$$\sum_{i=1}^n x_i w_i + b_k \geq \theta \quad (2.1)$$

Onde:

n : é o número de entradas do neurónio

w_i : é o peso associado à entrada x_i ,

θ : é o limiar (*threshold*) do neurónio.

O combinador linear calcula a entrada líquida (*Net*) do neurónio:

$$Net = \sum_{i=1}^n x_i w_i + b_k \quad (2.2)$$

Em seguida, a função de activação f é aplicada sobre *Net* gerando o valor de saída y do neurónio:

$$y = f(Net) \quad (2.3)$$

O bias (b_k) ou polarização é uma espécie de excitador ou inibidor com o efeito de aumentar ou diminuir a entrada líquida da unidade, dependendo se o seu valor for positivo ou negativo, respectivamente.

A função de activação f é também conhecida como função restritiva, já que limita o intervalo possível da saída (y) a um valor finito.

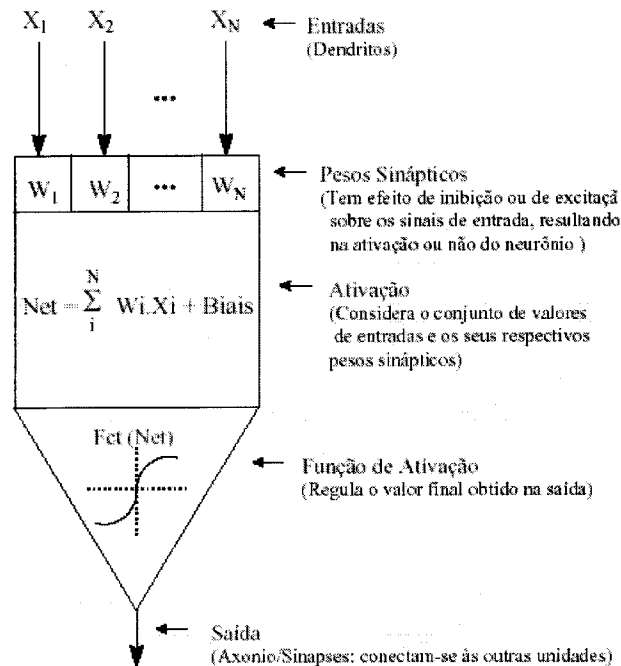


Figura 2.2 – Exemplo de Neurônio Artificial.

McCulloch & Pitts simplificaram seu modelo assumindo que os nodos em cada camada da rede disparam sincronamente, isto é, que todos os nodos são avaliados ao mesmo tempo e também que as entradas em um instante de tempo t produzem a sua saída no tempo $t + 1$. Em sistemas biológicos, sabe-se que não existe um mecanismo para sincronizar as acções de nodos, nem há restrição para que as suas saídas sejam activadas em tempos discretos como no modelo MCP. Sabe-se também que o valor da próxima saída dos nodos biológicos depende enormemente das activações dos neurónios anteriores, já que até mesmo os neuro-transmissores liberados anteriormente levam algum tempo para se recombinarem, influenciando, assim, as activações seguintes.

Finalmente, pode-se levantar algumas limitações na descrição do modelo MCP original, como:

- Redes MCP com apenas uma camada conseguem implementar funções linearmente separáveis;
- Pesos negativos são mais adequados para representar disparos inibidores;
- O modelo foi proposto com pesos fixos, não ajustáveis. Não há a presença de um algoritmo de aprendizado.

2.4. Funções de Activação

A partir do modelo proposto por McCulloch & Pitts foram derivados vários outros modelos que permitem a produção de uma saída qualquer, não necessariamente zero ou um, e com diferentes funções de activação. A figura 2.3 ilustra, graficamente, quatro funções de activação diferentes: a função linear, a função rampa, a função degrau (*step*) e a função sigmoidal.

A função de activação linear (figura 2.3a) é definida pela equação abaixo:

$$f = \alpha \cdot x \quad (2.4)$$

Onde α é um número real que define a saída linear para os valores de entrada, y é a saída e x é a entrada.

A função linear pode ser restringida para produzir valores constantes em uma faixa $[-\gamma, +\gamma]$, e neste caso, passa a ser a função rampa mostrada, graficamente, na figura 2.3.b definida pela notação abaixo:

$$f = \begin{cases} +\gamma & \text{se } x \geq \gamma \\ x & \text{se } x = \gamma \\ -\gamma & \text{se } x \leq -\gamma \end{cases} \quad (2.5)$$

Os valores máximo e mínimo da saída são $-\gamma$, $+\gamma$. A função rampa é usada geralmente, como uma função linear simplificada.

A função *step* (2.3c) é similar a uma função rampa, pois a função produz a saída $+\gamma$ para valores de x maiores que zero, caso contrário a função produz o valor de $-\gamma$. A função *step* é definida por:

$$f = \begin{cases} +\gamma & \text{se } x > 0 \\ -\gamma & \text{se } x \leq 0 \end{cases} \quad (2.6)$$

A função sigmoidal, conhecida também como *S-shape* (figura 2.3d) é uma função semi-linear. É possível definir várias funções sigmoidais. Uma das funções sigmoidais mais importantes é a função logística definida por:

$$f = \frac{1}{1 + e^{-x/T}} \quad (2.7)$$

Onde parâmetro T determina a suavidade de curva.

Em muitas situações é desejável termos uma função sinusoidal que varie entre -1 e +1. Nestas situações, uma função usualmente empregada é a *tangente hiperbólica*, dada por:

$$f = \frac{1 - e^{-x/T}}{1 + e^{-x/T}} \quad (2.8)$$

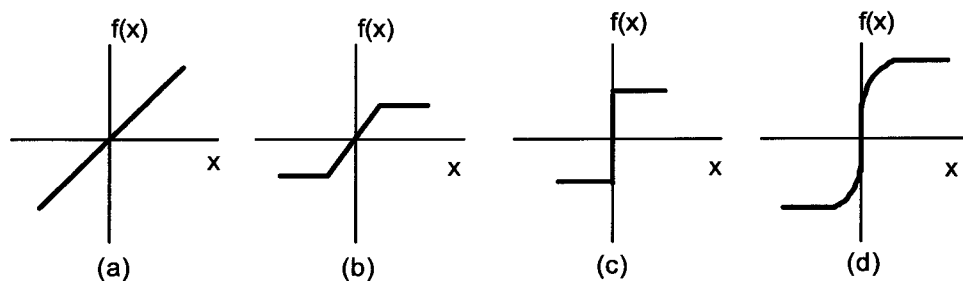


Figura 2.3 – Funções de ativação: (a) função linear, (b) função rampa, (c) função degrau (*step*) e (d) função sigmoidal.

2.5. Aprendizado

As RNA's possuem a capacidade de aprender através exemplos e fazer interpolações e extrapolações do que aprenderam. No aprendizado conexionista não se procura obter regras como na abordagem simbólica da Inteligência Artificial (IA), mas determinar a intensidade das conexões entre os neurónios. Um conjunto de procedimentos bem definidos para adaptar os parâmetros de uma RNA, afim de que a mesma possa *aprender* uma determinada função é chamado de *algoritmo de aprendizado*. Como era de se esperar, não há um único algoritmo de aprendizado. O que temos é um conjunto de ferramentas representadas por diversos algoritmos, cada qual com suas vantagens e desvantagens. Estes algoritmos, basicamente, diferem pela maneira através da qual o ajuste dos pesos é feito.

A utilização de uma RNA na solução de uma tarefa passa, inicialmente, por uma fase de aprendizagem, onde a rede extrai características relevantes de padrões de informações apresentados para a mesma, criando assim uma representação própria para o problema. A etapa de aprendizagem consiste em um processo iterativo de ajuste de parâmetros da rede, os pesos das conexões entre as unidades de processamento, que guardam o conhecimento que a rede adquiriu do ambiente em que está operando. Uma definição geral do que vem a ser aprendizagem pode ser expressa da seguinte forma (Mendel & McLaren, 1970):

“Aprendizagem é o processo pelo qual os parâmetros de uma rede neural são ajustados através de uma forma continuada de estímulos fornecidos pelo ambiente no qual a rede está operando, sendo que o tipo específico de aprendizagem realizado, é definido pela maneira particular como ocorrem os ajustes realizados nos parâmetros”.

Diversos métodos para treino de redes foram desenvolvidos, podendo estes serem agrupados em dois paradigmas principais: Aprendizado Supervisionado e Aprendizado Não Supervisionado. Outros dois paradigmas bastante conhecidos são os de Aprendizado por Reforço (que é um caso particular de Aprendizado Supervisionado) e Aprendizado por Competição (que é um caso particular de Aprendizado Não Supervisionado). Nas próximas secções estes paradigmas serão estudados, juntamente, com dois mecanismos de aprendizados encontrados na literatura, Correção de Erros e o Aprendizado Hebbiano.

2.5.1. Aprendizado Supervisionado

Este método de aprendizado é o mais comum no treinamento das RNA's, tanto de neurónios com pesos, quanto de neurónios sem pesos, sendo chamado *Aprendizado Supervisionado* porque a saída desejada que a rede deve resultar é fornecida por um supervisor (professor) externo. O objectivo é ajustar os parâmetros da rede, de forma a encontrar uma ligação entre os pares de entrada e saída fornecidos. A figura 2.4 ilustra o mecanismo de Aprendizado

Supervisionado. O “professor” indica, explicitamente, um comportamento bom ou mau para a rede, visando direccionar o processo do treino.

A rede tem sua saída corrente (calculada) comparada com a saída desejada, recebendo informações do supervisor sobre o erro da resposta actual. A cada padrão de entrada submetido à rede compara-se a resposta desejada (que representa uma acção óptima a ser realizada pela rede) com a resposta calculada, e os pesos das conexões são ajustados para minimizar o erro. A minimização da diferença é incremental, já que pequenos ajustes são feitos nos pesos à cada etapa do treino, de tal forma que estes caminhem, se possível, para uma solução.

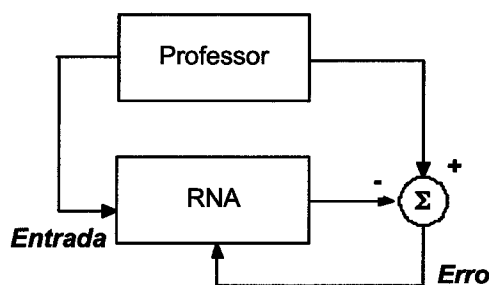


Figura 2.4 - Aprendizado Supervisionado

A desvantagem do Aprendizado Supervisionado é que, na ausência do “professor”, a rede não conseguirá aprender novas estratégias para situações não cobertas pelos exemplos do treino da rede. Os exemplos mais conhecidos de algoritmos de Aprendizado Supervisionado são a regra delta ou correcção de erros (Widrow & Hoff, 1960) e a sua generalização para redes de múltiplas camadas, o algoritmo *backpropagation* (Rumelhart *et al.*, 1986).

O Aprendizado Supervisionado pode ser implementado, basicamente, de duas formas: *off-line* e *on-line*. Para o treino *off-line*, os dados do conjunto do treino não mudam, sendo que uma vez obtida uma solução para a rede, esta deve permanecer fixa. Por sua vez, no aprendizado *on-line*, o conjunto de dados muda, continuamente, visto que a rede deve estar em contínuo processo de adaptação.

2.5.2. Correção de Erros

O processo de aprendizado por correção de erros consiste em ajustar os pesos e a polarização de suas unidades, para que a saída desejada seja obtida. Para adaptação da polarização juntamente com os pesos, pode-se considerá-los como sendo o peso associado a uma conexão cuja entrada é sempre igual a -1 e adaptar o peso relativo a essa entrada. Quando um padrão é inicialmente apresentado à rede, ela produz uma saída. Após medir a distância entre a resposta actual e a desejada, são realizados os ajustes apropriados nos pesos das conexões de modo a reduzir esta distância. Esse procedimento é conhecido como Regra de Delta.

A Regra de Delta é utilizada para minimizar o erro gerado na saída do perceptrão através do gradiente descendente (Anexo A). A Regra de Delta pode ser expressa matematicamente por:

$$\Delta w_{ij} = \eta(d_i - x_i \theta_j) \quad (2.9)$$

Onde:

w_{ij} - peso da conexão entre os neurónios i e j

Δw_{ij} - ajustes dos pesos das conexões

η - taxa de aprendizagem

d_i - estado de activação desejado do neurónio i

x_i - estado de activação real do neurónio i

θ_j - saída do neurónio j

Esta expressão aparece tanto no algoritmo de treino do perceptrão (Rosenblatt, 1958) quanto no algoritmo do treino do ADALINE (Widrow & Hoff, 1960) e posterior generalização para o algoritmo *backpropagation* (Rumelhart et al., 1986). Este método caracteriza o conceito de aprendizado supervisionado descrito na secção 2.5.1.

2.5.3. Aprendizado não Supervisionado

No Aprendizado Não Supervisionado, como o próprio nome sugere, não há um professor ou supervisor para acompanhar o processo de aprendizado.

Este método é ilustrado na figura 2.5. Para este tipo de aprendizado, somente os padrões de entrada estão disponíveis para rede, ao contrário do Aprendizado Supervisionado, cujo conjunto do treino possui pares de entrada e saída. A partir do momento em que a rede estabelece uma harmonia com as regularidades estatísticas da entrada de dados, desenvolve-se nela uma habilidade de formar representações internas para codificar características da entrada e criar novas classes ou grupos automaticamente. Este tipo de aprendizado, só se torna possível, quando existe redundância nos dados de entrada. Sem redundância seria impossível encontrar quaisquer padrões ou características dos dados de entrada.



Figura 2.5 – Aprendizado Não Supervisionado

Uma das vantagens do Aprendizado Não Supervisionado pode ser visualizada na ausência do supervisor ou professor (responsável pela indicação da saída desejada). A desvantagem está, intimamente, ligada ao conjunto de dados para o treino, onde este deve ser redundante para que a rede consiga abstrair características durante o seu treino.

2.5.4. Aprendizado Hebbiano

Os resultados obtidos por Hebb (1949) motivaram os primeiros métodos de aprendizado em RNA's. Como originalmente postulado, a regra de Hebb propõe que o peso de uma conexão sináptica deva ser ajustado, caso exista sincronismo entre os "níveis de actividade" das entradas e saídas. Caso dois neurónios, em lados distintos da sinapse, são activados sincronamente, teremos um fortalecimento desta sinapse. Entretanto, se os neurónios forem activados assincronamente, a sinapse será enfraquecida ou mesmo eliminada. Em outras palavras, se o neurónio pré-sináptico tiver grande influência na activação do neurónio pós-sináptico, a conexão entre eles dever ser reforçada.

A sinapse Hebbiana possui quatro características principais:

- *Mecanismo interactivo*: dentro do aprendizado Hebbiano, não se pode analisar actividades pré-sinápticas de forma isolada; qualquer modificação na sinapse Hebbiana depende (estatisticamente ou deterministicamente) da interacção entre os dois tipos de actividades (pré e pós-sinápticos).
- *Mecanismo local*: uma sinapse Hebbiana é uma transmissão com sinais contínuos que produz modificações sinápticas locais que são entradas específicas. É a partir do mecanismo local que sinapses Hebbianas efectuem Aprendizado Não Supervisionado.
- *Mecanismo dependente do tempo*: as modificações em uma sinapse Hebbiana dependem do momento exacto de ocorrência das actividades pré e pós-sinápticas.
- *Mecanismo correlacional ou conjuncional*: a sinapse Hebbiana pode ser chamada de sinapse conjuncional pelo fato de que a ocorrência conjunta de actividades pré e pós-sinápticas é suficiente para que haja uma modificação. Além disso, pode também ser chamada de sinapse correlacional, porque uma correlação entre estas mesmas actividades é suficiente para gerar mudanças.

Pode-se expressar o postulado de Hebb em termos matemáticos a partir da seguinte equação, que consiste em uma regra para a mudança do peso sináptico w_{ij} (regra do produto de actividade):

$$\Delta w_{ij}(t) = \eta \cdot y_i(t) x_j(t) \quad (2.10)$$

Onde η é uma constante positiva que determina a taxa de aprendizagem, y_i é a saída desejada no instante de tempo t e x_j é a entrada do neurónio j no instante de tempo t .

Apesar dos vectores de entrada e saída serem fornecidos (como no Aprendizado Supervisionado), a regra de Hebb é classificada como Aprendizado Não Supervisionado, já que não existe supervisor externo para verificar a qualidade da resposta da rede para prover um ajustes de pesos. Neste caso, o treino da rede é feito, independentemente, da resposta actual, através de um mecanismo local à sinapse.

2.5.5. Aprendizado por Competição

Aprendizado por Competição é um caso particular de Aprendizado Não Supervisionado (Grossberg, 1976). A ideia neste caso é, dado um padrão de entrada, fazer com que as unidades de saída disputem entre si para serem activadas. Existe, portanto, uma competição entre as unidades de saída para decidir qual delas será a vencedora e, conseqüentemente, terá a sua saída activada e seus pesos actualizados no treino.

As unidades de entrada são directamente conectadas às unidades de saída, sendo que estas últimas também podem estar ligadas entre si via conexões laterais inibitórias, ou negativas. A unidade de saída com maior activação inicial terá uma maior probabilidade de vencer a disputa com as outras unidades, que perderão o poder de inibição ao longo do tempo sobre as unidades de maior activação. A unidade mais forte fica ainda mais forte e seu efeito inibidor sobre as outras unidades de saída torna-se dominante. Com o tempo, todas as outras unidades de saída ficarão completamente inactivas, excepto a vencedora. Este tipo de inibição mútua também é conhecido como "*winner tasks all*".

Um algoritmo simples de aprendizado competitivo é:

- Apresentar um vector de entrada;
- Calcular a activação inicial de cada unidade de saída;
- Deixar as unidades de saída competirem até que apenas uma permaneça activa;
- Aumentar os pesos sobre as conexões entre as unidades de saída activa e as unidades de entrada activas. Com isso, a unidade de saída terá maior probabilidade de ficar activa na próxima repetição do padrão.

Um problema neste algoritmo é que uma unidade de saída pode se tornar dominante e ficar activa todo tempo, podendo captar para si todo o espaço de entrada. A solução para tal problema está em racionar os pesos, de tal forma que a soma dos pesos sobre as linhas de entrada de uma unidade seja limitada a 1. Para aumentar o peso de uma conexão, é preciso diminuir o peso de outra, tal que:

$$\Delta w_j = \frac{\eta X}{M} - \eta w_j \quad (2.11)$$

Para todo $j = 1, 2, \dots, n$, onde w_j é o peso da conexão da unidade de entrada j com a unidade activa, x_j é o valor do j -ésimo bit de entrada, M é o número de unidades de entrada activas no vector de entrada e η é a taxa de aprendizagem.

Este algoritmo funciona em muitos casos, porém, às vezes, uma unidade de saída ganha sempre. Uma solução para este problema é a *leaky learning* que foi desenvolvido por Grossberg em 1987. O Aprendizado por Competição é a base dos modelos ART de Grossberg e dos mapas de Kohonen.

2.5.6. Aprendizado por Reforço

Aprendizado por Reforço pode ser visto como um caso particular de Aprendizado Supervisionado. A principal diferença entre o Aprendizado Supervisionado clássico e o Aprendizado por Reforço é a medida de desempenho usada em cada um dos sistemas. No Aprendizado Supervisionado, a medida de desempenho é baseada no conjunto de respostas desejadas usando um critério de erro conhecido, enquanto que, no Aprendizado por Reforço o desempenho é baseado em qualquer medida que possa ser fornecida ao sistema.

No Aprendizado por Reforço, a única informação de realimentação fornecida à rede é se uma determinada saída está correcta ou não, isto é, não é fornecida a rede a resposta correcta para o padrão de entrada. O Aprendizado por Reforço é ilustrado na **Figura 2** figura 2.6.

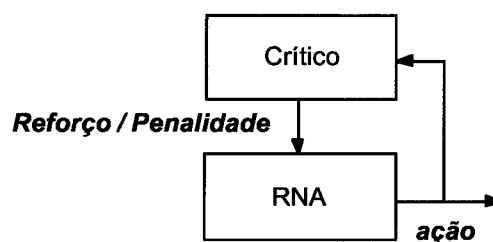


Figura 2.6 – Aprendizado por Reforço

Aprendizado por Reforço é uma forma de aprendizado *on-line* obtido por um mapeamento de entrada-saída, através de um processo de triagem de erro desenvolvido para maximizar o índice de desempenho escalar chamada de *signal de reforço*. O termo Aprendizagem por Reforço foi, inicialmente, citado por Minsky (1961) em seus estudos iniciais de IA. A ideia básica que está por trás do termo “reforço” tem sua origem em estudos experimentais sobre aprendizado dos animais. Neste contexto, é interessante lembrar a Lei do Efeito (Thorndike, 1911) que diz quanto maior a satisfação obtida com uma certa experiência em um animal, maiores as chances dele aprender. Sutton (1991) conceitualiza o Aprendizado por Reforço como:

“Se uma acção tomada pelo sistema de aprendizagem é seguida de estados satisfatórios, então a tendência do sistema de produzir esta acção particular é reforçada. Se não for seguida de estados satisfatórios, a tendência do sistema de produzir esta acção é enfraquecida”.

2.6. Arquitecturas de RNA's

Nesta secção será abordado o modelo perceptrão de única e múltiplas camadas. O perceptrão foi proposto por Rosenblatt (1958) e originou-se do trabalho de McCulloch & Pitts (1943).

2.6.1. Perceptrão de Camada Única e seu Algoritmo de Aprendizado

O trabalho original que enfocou o modelamento de um neurónio e sua capacidade computacional de solucionar problemas, teve início com as pesquisas de McCulloch & Pitts (1943). Naquela época surgiram os primeiros computadores digitais e entende-se que a associação de um modelo artificial de um neurónio com a capacidade de resolver funções Booleanas está, provavelmente, ligada ao pensamento vigente na época de que era possível construir uma máquina inteligente através de operadores lógicos básicos. Foi somente com o trabalho de Rosenblatt (1958) que o conceito *aprendizado* em RNA's foi introduzido.

O modelo proposto por Rosenblatt, conhecido como perceptrão, era composto por uma estrutura de rede tendo como unidades básicas nodos MCP e

uma regra de aprendizado. Alguns anos mais tarde, Rosenblatt demonstrou o teorema de convergência do perceptrão, o qual mostra que um nodo MCP treinado com o algoritmo de aprendizado do perceptrão sempre converge, caso o problema em questão seja linearmente separável (Sycara, 1988).

A topologia original descrita por Rosenblatt era composta por unidades de entrada (retina), por um nível intermediário formado pelas unidades de associação e por um nível de saída formado pelas unidades de resposta. Embora, esta topologia original possua três níveis, ela é conhecida como perceptrão de uma única camada, já que somente o nível de saída (unidade de resposta) possui propriedades adaptativas. A figura 2.7 mostra um esboço da topologia do perceptrão.

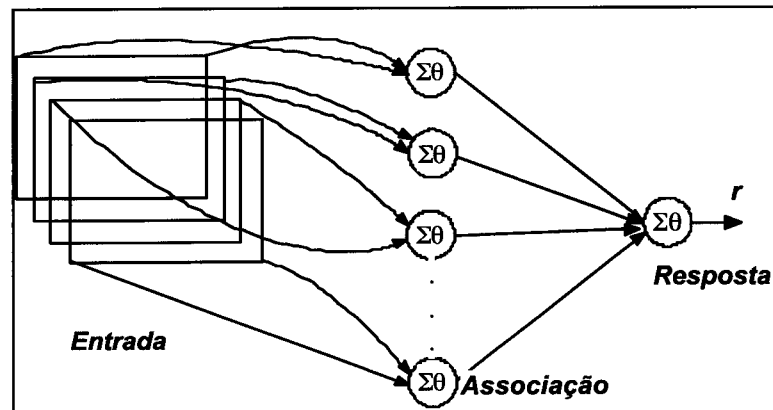


Figura 2.7 – Topologia de um perceptrão simples com uma única saída

Apesar de ter causado grande euforia na comunidade científica da época, o perceptrão não teve vida muito longa, já que duras críticas de Minsky & Papert (1969) foram feitas em relação à sua capacidade computacional, causando, assim, um grande impacto sobre as pesquisas em RNA's, o que levou a um grande desinteresse pela área durante os anos 70 e início dos anos 80.

Esta visão pessimista sobre a capacidade do perceptrão e das RNA's de uma maneira geral mudou, devidas as descrições da rede de Hopfield em 1982 e do algoritmo de "backpropagation" (que será visto nas próximas seções). Foi em consequência destes trabalhos que a área de RNA's ganhou novo impulso, ocorrendo a partir do final dos anos 80, uma forma de expansão no número de trabalhos de aplicação e teóricos envolvendo RNA's e técnicas correlatadas.

2.6.2. Perceptrão de Multicamadas – PMC

Conforme citado na secção anterior, as redes de uma só camada resolvem apenas problemas linearmente separáveis. A solução de problemas não linearmente separáveis passa pelo uso de redes com uma ou mais camadas intermediárias, ou escondidas. Segundo Cybenko (1989), uma rede com uma camada intermediária pode implementar qualquer função contínua. A utilização de duas camadas intermediárias permite a aproximação de qualquer função (Cybenko, 1988).

Com a utilização desta arquitectura é necessário, então, encontrar uma forma de treinar tal rede. Uma possibilidade seria a divisão da rede em um conjunto de sub-rede para cada camada, com treinamento independente. Uma rede com uma camada intermediária pode ser formada por um conjunto de redes perceptrão, uma rede para cada grupo de entradas linearmente separáveis. Na camada de saída, uma outra rede combina as saídas produzidas pelas redes da primeira camada, gerando, assim, a classificação final. Esta solução esbarra na dificuldade de dividir um problema em sub problemas. Geralmente, esta divisão ou não é possível ou muito complicada.

Uma outra alternativa seria treinar a rede completa de uma só vez. O problema agora está em como treinar os nodos da camada intermediária desta rede e assim, quais seriam as respostas desejadas para estes nodos, isto é, como seria definido o erro destes nodos. O problema passa agora a ser a definição do erro dos nodos das camadas intermediárias. Se uma função de activação do tipo limiar (função *threshold*) for utilizada, a avaliação do erro será bastante complexa, uma vez, que os nodos das camadas intermediárias e de saída não saberão quão distantes as respostas de seus nodos estarão dos valores desejados, este é um dos problemas da rede perceptrão.

Uma alternativa seria a utilização de funções de activação lineares. No entanto, uma rede com mais de uma camada cujos nodos utilizarão funções de activação lineares é equivalente a uma rede de uma só camada. Esta afirmação pode ser, facilmente, provada:

Seja uma Rede Neural com uma camada intermediária e funções de activação lineares. Seja x o vector de entrada e w_A e

w_B as matrizes de pesos para nodos da primeira e segunda camadas, respectivamente. O vector de saída da rede é gerado de acordo com:

$$y = (xw_A)w_B \quad (2.12)$$

Utilizando propriedades da Álgebra Linear (Lay, 1997) tem-se que:

$$\forall (w_A \text{ e } w_B) \exists w_C \mid w_C = w_A w_B \quad (2.13)$$

Logo:

$$y = xw_C \quad (2.14)$$

Onde x é o vector de entrada, y o vector de saída e w_A , w_B , e w_C , são vectores de pesos.

Para treinar redes com mais de uma camada foi proposto um método que se baseia em gradiente descendente (Rumelhart *et al.*, 1986). Afim de que, este método possa ser utilizado, a função de activação precisa ser contínua, diferenciável e, de preferência não decrescente. A função de activação deve informar os erros cometidos pela rede para as camadas anteriores com a maior precisão possível. A alternativa utilizada para superar os problemas acima citados foi a utilização de funções de activação do tipo sinusoidal.

2.6.2.1. Arquitectura PMC

Conforme visto nas secções anteriores, a inexistência ou desconhecimento de algoritmos para treinar redes com uma ou mais camadas intermediárias foi uma das causas para redução das pesquisas em RNA's na década de 70. Tal problema foi primeiro apontado por Minsky & Papert (1969), no livro "*Perceptrons*" e, por representar uma redução no universo de aplicações de RNA's, causou uma diminuição drástica tanto do número de pesquisadores que trabalhavam na área quanto no financiamento de projectos envolvendo RNA's.

As redes PMC apresentam um poder computacional maior do que aquele apresentado pelas redes sem camadas intermediárias. As redes PMC's podem

tratar com dados que não são linearmente separáveis. A precisão obtida e a implementação de uma rede PMC dependem do número de nodos utilizados nas camadas intermediárias, e um dos problemas é a definição de uma função de activação. Na figura 2.8 é apresentada uma rede PMC típica.

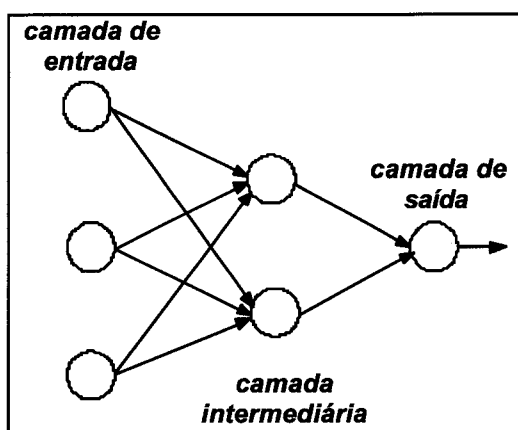


Figura 2.8 - Rede PMC com camada intermediária

Conforme observado, um dos principais aspectos relacionados ao projecto de redes PMC diz respeito à função de activação utilizada. Diversas funções de activação têm sido propostas para redes multicamadas. Estas funções são não lineares e diferenciáveis. As funções precisam ser diferenciáveis para que o gradiente possa ser calculado, direccionando o ajuste dos pesos. A função de activação mais utilizada é *sigmoidal logística*. A orientação da sinusoidal é determinada pela direcção do vector de pesos w . O valor do termo de polarização, corresponde ao peso w_0 , determina a localização da função sinusoidal, ou seja, define a posição da função com relação ao eixo da ordenada.

2.6.2.2. Definição de uma Topologia Ideal para o PMC

Na elaboração da arquitectura da RNA ideal para solucionar um determinado problema, uma das tarefas mais difíceis é determinar o número de elementos de processamento da camada oculta, bem como o número de camadas ocultas.

A definição de uma topologia de uma rede do tipo PMC está intimamente ligada com a definição dos nodos na camada intermediária; este número é, geralmente, definido empiricamente e depende, fortemente, da distribuição dos padrões de treino e validação da rede. Alguns métodos, no entanto, têm sido propostos. Os mais utilizados são:

- Definir o número de unidades em função do número de entradas e saídas;
- Utilizar um número de conexões dez vezes menor que o número de exemplos. Este método apenas reduz a incidência de “*overfitting*” (rede memoriza padrões de treinamento, ao invés de extrair as características). Se o número de exemplos for muito maior que o número de conexões, “*overfitting*” é improvável, mas pode ocorrer “*underfitting*” (a rede não converge durante o seu treinamento).

O número adequado de nodos na camada intermediária depende de vários factores, como:

- Número de exemplos de treino;
- Quantidade de ruído presente nos exemplos;
- Complexidade da função a ser aprendida;
- Distribuição estatística dos dados do treino.

Existem problemas que necessitam apenas de uma unidade de entrada e uma unidade de saída, e outros que podem precisar de milhares de unidades intermediárias. O número de unidades intermediárias pode também, em alguns casos, crescer exponencialmente, com o número de entradas. A solução neural mais eficiente é aquela onde o número de unidades cresce apenas polinomialmente com o aumento dos números de unidades de entrada.

Para solução de problemas práticos de reconhecimento de padrões, aloca-se para a rede um número de unidades intermediárias suficiente para a solução do problema. Deve-se ter cuidado para não utilizar unidades demais, o que pode levar a rede memorizar padrões de treinamento, ao invés de extrair as características gerais que permitirão a generalização ou reconhecimento de

padrões não vistos durante o treino (este problema é chamado de *overfitting*), não utilizar um número muito pequeno, que possa forçar a rede a gastar tempo em excesso tentando encontrar uma representação ótima (as unidades utilizadas podem ficar sobrecarregadas, tendo que lidar com um elevado número de restrições).

2.6.2.3. Otimização da Topologia utilizando *Pruning*

Conforme observado nos parágrafos anteriores, uma das maiores dificuldades em se definir a estrutura de uma RNA é o fiel dimensionamento de sua topologia. Normalmente, o número de camadas e o número de nodos em cada camada é definido em função de uma inspeção prévia nos dados e da complexidade do problema. Uma vez definida a topologia inicial, a estrutura final mais adequada para o modelamento é, normalmente, obtida através de refinamentos sucessivos, que pode levar a um tempo de dimensionamento alto, já que este tem um grande componente empírico.

O objectivo da etapa do ajuste é a obtenção de uma topologia de rede que modele com precisão os dados do conjunto do treino, mas também resulte em uma aproximação com boa capacidade de generalização. Como, na maioria dos casos, os conjunto do treino de uma RNA é composto de dados experimentais, estes contêm, implicitamente, erros inerentes aos processos de amostragem. Desta forma, a aproximação através de RNA's deve ser feita visando a obtenção de uma estrutura que seja capaz de modelar os dados sem modelar o ruído contido neles.

Este é um problema no projecto de RNA's conhecida na literatura como "*bias and variance dilemma*" (Geman *et al.*, 1992), que envolve a obtenção de um modelo que não seja muito rígido a ponto de não modelar fielmente os dados, mas que também não seja, excessivamente, flexível a ponto de modelar também o ruído.

O equilíbrio entre a rigidez e flexibilidade da rede é obtido através de seu dimensionamento. Quanto maior a sua estrutura, maior o número de parâmetros livres ajustáveis e, conseqüentemente, maior a sua flexibilidade. Porém, quando

os dados são apresentados à rede, não se tem real conhecimento de sua complexidade, daí a dificuldade do problema de dimensionamento.

Uma forma de evitar o “*overfitting*” é estimar o erro de generalização durante a fase de treino. Para isto, o conjunto de dados é dividido em *conjunto de treinamento* e *conjunto de validação*. O conjunto de treinamento é utilizado na modificação dos pesos e o conjunto de validação é utilizado para estimar a capacidade de generalização da rede durante o processo de aprendizagem. O treinamento deve ser interrompido quando o erro do conjunto de validação começar a subir, ou seja, quando a rede começar a incorporar o ruído presente nos dados o que causa degradação na sua capacidade de generalização. Embora, esta alternativa possa se mostrar eficiente em algumas situações, a sua utilização é limitada para os casos em que um conjunto de treinamento muito grande está disponível, já que os dados do conjunto de validação não podem ser utilizados para treinamento. Uma outra solução conhecida é adoção de técnicas de pruning que, por sua vez, envolvem a eliminação de pesos e nodos irrelevantes para a função executada pela rede. Existem, basicamente, dois tipos de métodos de pruning: métodos baseados na avaliação da sensibilidade da saída e os métodos que envolvem modificações na função de custo (Reed, 1993).

A filosofia dos métodos do primeiro grupo é a retirada de elementos da rede e subsequente verificação da variação do erro de saída (Karnin, 1990). Caso a retirada do elemento não cause grande variação no erro de saída, conclui-se que a rede é pouco sensível a este e que o mesmo pode ser retirado sem perda da capacidade da rede modelar dados.

A desvantagem deste método é que a retirada dos elementos não considera a correlação entre eles, Um exemplo de situação em que este método poderia falhar seria o caso de dois nodos que anulam as suas saídas em uma camada intermediária da rede. Como um par eles não causam nenhum efeito na saída de rede, mas individualmente, podem ter um grande efeito quando a remoção de um deles é analisada (Reed, 1993).

Por sua vez, os métodos que envolvem modificação na função de custo, utilizam, em uma primeira etapa, termos de regularização (Tikhonov & Arsenin, 1977) adicionados ao termo do erro padrão. A ideia deste algoritmo é a obtenção

de soluções com pesos de norma mínima, para isto, as soluções com valores altos de pesos são penalizadas durante o treinamento. Começa-se com uma rede com topologia super-dimensionada e desta são retirados os pesos irrelevantes.

Como exemplo, forma geral da função custo para o treinamento do algoritmo “*weight decay*” é apresentada na equação seguinte. Pode-se então observar que se minimiza não somente a soma dos erros quadráticos, mas também a norma de vector de pesos.

$$J = \frac{1}{2} \sum_{i=1}^k (d_i - y_i)^2 + \frac{1}{2} \lambda \|w\|^2 \quad (2.15)$$

Onde k é o número de nodos de saída, d_i é saída desejada do nodo i , y_i a saída gerada pelo nodo i e w o vector de pesos.

Uma vez obtida uma solução com pesos de norma mínima, aqueles com valores muito pequenos são então considerados irrelevantes e eliminados, espera-se então obter uma solução com um fiel equilíbrio entre a rigidez e a flexibilidade da rede. A contrapartida para esta abordagem, que parece mais elegante, é a busca de uma solução com norma mínima também depende do ajuste de forma empírica do parâmetro de regularização λ . Quando este é nulo somente a soma dos erros quadráticos é minimizada quando possui valor muito grande a solução para o vector de pesos tende para $w = 0$. Portanto, um valor satisfatório intermediário deve ser encontrado.

Uma solução para contornar este problema é incluir o parâmetro λ como uma das variáveis a serem ajustadas, porém, isto incorpora mais complexidade ao processo do treino. Outros procedimentos para a construção de redes com estrutura otimizada são a utilização de *redes construtivas*.

2.6.2.4. Treino de Redes PMC

Existem, actualmente, vários algoritmos para treinar rede PMC (Rumelhart & McClelland, 1986; Riedmiller, 1994; Hagan & Menhay, 1994). Estes algoritmos

são do tipo supervisionado. De acordo com os parâmetros que eles actualizam, os algoritmos para o treino de redes do tipo PMC podem ser classificados em:

- Estáticos;
- Dinâmicos.

Como os algoritmos estáticos não alteram a estrutura da rede, variando apenas os valores de seus pesos, os algoritmos dinâmicos podem tanto reduzir quanto aumentar o tamanho da rede (número de camadas, número de nodos nas camadas intermediárias e número de conexões). Quando se utiliza o aprendizado estático, a mesma regra de aprendizado é empregada para redes do tipo PMC com diferentes tamanhos e formatos. Interessante observar que topologias diferentes podem resolver o mesmo problema.

O algoritmo de aprendizado mais conhecido para o treino destas redes é o “*backpropagation*”. A maioria dos métodos de aprendizado para RNA’s do tipo PMC utiliza variações deste algoritmo. O “*backpropagation*” foi um dos principais responsáveis pelo ressurgimento do interesse em RNA’s, por ocasião da publicação do livro “*Parallel Distributed Processing*”, mais conhecido como PDP (Rumelhart & McClelland, 1986). Embora a popularização deste algoritmo tenha surgido a partir de 1986, ele foi proposto muito antes, com diferentes propósitos, por diferentes pesquisadores como: Bryson & Ho (1969), Werbos (1974), Parker (1985) e Le Cun (1985).

O “*backpropagation*” é um algoritmo supervisionado, que utiliza pares (entrada, saída desejada), através de um mecanismo de correcção de erros, para ajustar os pesos da rede. O treino ocorre em duas fases, onde cada fase percorre a rede em um sentido. Estas duas fases são chamadas de fase “*forward*” e fase “*backward*” (figura 2.9).

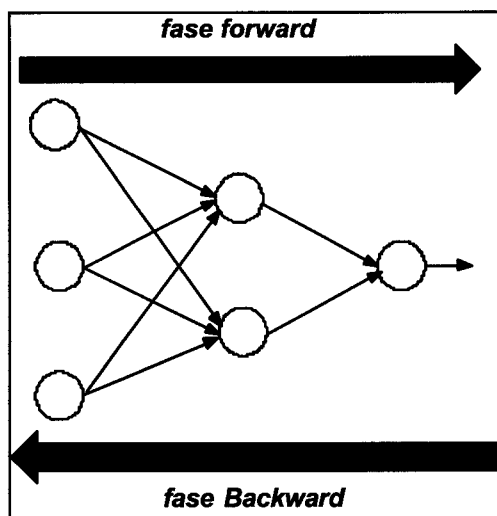


Figura 2.9 – Fluxo de Processamento do algoritmo backpropagation.

A fase “*forward*” envolve os seguintes passos:

- 1 – A entrada é apresentada à primeira camada da rede, a camada C^0 ;
- 2 – Para cada camada C^i a partir da camada de entrada;
 - 2.1 – Após os nodos da camada C^i ($i > 0$);

Calcular os sinais de saída, estes servem como entrada para a definição das saídas produzidas pelos nodos da camada C^{i+1} .

- 3 – As saídas produzidas pelos nodos da última camada são comparadas às saídas desejadas.

A fase “*backward*” envolve as etapas listadas a seguir:

- 1 – A partir da última camada, até chegar na camada de entrada:
 - 1.1 – Os nodos da camada actual ajustam seus pesos de forma a reduzir seus erros
 - 1.2 – O erro de um nodo da camada intermediária é calculado, utilizando os erros dos nodos da camada seguinte conectados a ele, ponderado pelos pesos das conexões entre eles

O “*backpropagation*” é baseado na regra proposta por Widrow & Hoff (1960), sendo também chamada de regra delta generalizada. Este algoritmo propõe uma forma de definir o erro dos nodos das camadas intermediárias, possibilitando o ajuste de seus pesos. Os ajustes dos pesos são realizados

utilizando o método do gradiente. Na literatura da análise numérica, a regra delta generalizada é conhecida como “método da bola pesada” (Bertsekas, 1995).

A derivação da regra delta generalizada é simples e semelhante à derivação da regra delta. Também neste caso, a função a ser minimizada é uma função erro, definida pela soma dos erros quadráticos e representada por:

$$E = \frac{1}{2} \sum_p \sum_{i=1}^k (d_i^p - y_i^p)^2 \quad (2.16)$$

Onde E é a medida de erro total, p é o número de padrões, k é o número de unidades de saída, d_i é a i -ésima saída desejada e y_i é a i -ésima saída gerada pela rede. Esta equação define o erro total cometido pela rede, ou a quantidade que, para todos os padrões p de um dado conjunto, as saídas geradas pela rede diferem das saídas desejadas.

A regra delta generalizada requer que as funções de activação utilizadas pelos nodos sejam contínuas, diferenciáveis, geralmente, não decrescente da entrada total recebida pelo nodo. Estas funções são chamadas de funções semi-lineares. A Equação abaixo ilustra o cálculo do valor de activação.

$$y_j^p = f_j(\text{net}_j^p) \quad (2.17)$$

Onde:

$$\text{net}_j^p = \sum_{i=1}^n x_i^p w_{ji} \quad (2.18)$$

A constante n representa o número de conexões de entrada do nodo j e w_{ji} o peso da conexão entre a entrada x_i^p e o nodo j .

Há garantia de convergência se a superfície do erro for simples. Uma forma de visualizar a influência do treino da rede no valor do erro está ilustrada pela superfície de erro na figura 2.10. Esta figura demonstra um caso simples de uma

rede com apenas um nodo e duas conexões. Os valores dos pesos da rede definem a coordenada de um ponto da superfície do erro. O erro produzido pela rede, para cada combinação de valores de pesos, é dado pela altura da superfície naquele ponto.

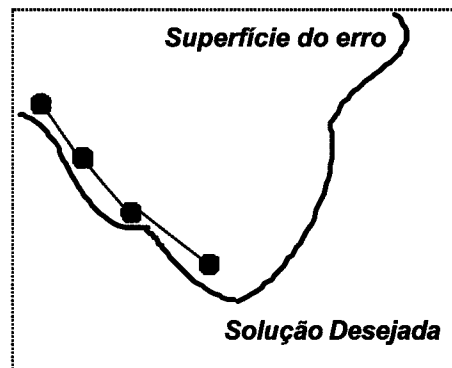


Figura 2.10 – Superfície do Erro de uma Rede PMC

As três equações anteriores mostram como encontrar o erro global da rede e como encontrar o erro local de um determinado padrão de entrada, processadas as duas equações anteriores deve-se aplicar a equação seguinte que ajustará os pesos das diversas camadas:

$$\Delta w_{ji} = \eta \delta_j x_i \quad (2.18)$$

ou

$$w_{ij}(t+1) = w_{ij}(t) + \eta \delta_j(t) x_i(t) \quad (2.19)$$

Onde $\delta_j = (d_j - y_j)$ para camada de saída e $\delta_j = \sum w_{ij} \delta_i$ para camadas intermediárias, i unidade em questão, j camada em questão.

2.6.2.5. Problemas Enfrentados no Treino de Redes PMC

Um dos problemas enfrentados no treino de rede PMC diz respeito à definição de seus parâmetros. A selecção dos parâmetros de treino do algoritmo “*backpropagation*” é um processo pouco conhecido, muitas vezes chamado de

“magia negra”. Pequenas diferenças nestes parâmetros podem levar a grandes divergências tanto no tempo de treino como na generalização obtida.

Esta é uma dúvida que surge, naturalmente, em que momento parar de treinar a rede. Existem vários métodos para a determinação do momento que o treino deva ser encerrado. Estes métodos são chamados de *critérios de parada*. Os critérios de paradas mais utilizados são:

- Encerrar o treino após n ciclos;
- Encerrar o treino após o erro quadrático médio ficar abaixo de uma constante c .
- Encerrar o treino quando a percentagem de classificações correctas estiverem acima de uma constante c (mais indicados para saídas binárias);
- Combinação dos métodos acima.

Outro aspecto que precisa ser observado é frequência das actualizações dos pesos. A frequência de ajustes dos pesos influencia o desempenho obtido durante o treinamento. Duas abordagens diferentes têm sido utilizadas quanto à frequência (periodicidade) para o ajuste de pesos pelo algoritmo “*backpropagation*”:

- Por padrão (on-line);
- Por ciclo (batch).

Existem vantagens e desvantagens em cada uma destas abordagens. Na abordagem por padrão, os pesos são actualizados após a apresentação de cada padrão de treino. Esta abordagem é estável, se a taxa de aprendizado for pequena (é aconselhável reduzir progressivamente esta taxa). Quando as taxas de aprendizagem são elevadas, a rede torna-se instável. A abordagem por padrão é mais rápida, principalmente, se o conjunto do treino for grande e redundante. Outra vantagem desta técnica é que ela requer menos memória.

Na abordagem por ciclo, por sua vez, os pesos são actualizados após todos os padrões terem sido apresentados. Esta técnica é mais estável, mas pode ser lenta, se o conjunto de treino for grande e redundante. Esta abordagem apresenta uma estimativa mais precisa do vector gradiente, ao custo da

necessidade de mais memória. A escolha da abordagem a ser utilizada depende da aplicação e da distribuição estatística dos dados.

O principal problema diz respeito a lentidão do algoritmo para superfícies complexas. Uma forma de minimizar este problema é considerar os efeitos de segunda ordem (taxa de momento) para o gradiente descendente. Não é raro o algoritmo convergir para mínimos locais.

Mínimos locais são pontos na superfície de erro que apresentam uma solução estável, embora não seja correcta. Algumas técnicas são utilizadas tanto para acelerar o algoritmo “*backpropagation*” como para reduzir a incidência de mínimos locais:

- Utilizar taxa de aprendizado decrescente;
- Adicionar nós intermediários;
- Utilizar um termo momento;

Entre as várias técnicas utilizadas para acelerar o processo de treinamento e evitar mínimos locais, a adição de um termo *momento* é uma das mais frequentes. O termo *momento* é representado por:

$$w_{ij}(t+1) = w_{ij}(t) + \eta \delta_j(t) x_i(t) + \mu(w_{ij}(t) - w_{ij}(t-1)) \quad (2.20)$$

Onde μ é taxa de *momento* a ser utilizada no processamento do aprendizado da rede.

A inclusão da taxa de *momento* na fórmula de ajuste dos pesos aumenta a velocidade de aprendizado (aceleração), reduzindo o perigo de instabilidade. A taxa *momento* pode acelerar o treino em regiões muito planas para a superfície de erro. Além disso, ele suprime a oscilação de pesos em valores de ravinas.

2.6.2.6. Descida do Gradiente Conjugado

Descida de Gradiente Conjugada (DGC) é um método avançado de treinar perceptrões de multicamadas, normalmente executa significativamente melhor

que o algoritmo BP. Assim como o BP, este algoritmo pode ser usado para qualquer rede com um grande número de pesos.

DGC é um algoritmo de actualização de grupo (batch): considerando que o algoritmo BP ajusta os pesos da rede depois de cada caso, o algoritmo DGC trabalha com o gradiente médio da superfície de erro para todos os casos, antes de actualizar os pesos uma vez ao término da iteração, por isto, não há nenhuma necessidade de seleccionar a taxa de aprendizagem e momento.

DGC trabalha construindo uma série de linhas de procura através do erro de superfície. Primeiro actua na direcção de descida mais íngreme, da mesma maneira que o BP, porém, em vez de dar um passo proporcional para uma taxa de aprendizagem, a descida de gradiente conjugada projecta uma linha directa naquela direcção, e então localiza um mínimo ao longo desta linha – um processo que é bastante rápido pois só envolve a procura em uma dimensão. Subsequentemente, são administradas procuras de linhas adicionais (uma por iteração). A direcção das linhas de procura (método das direcções conjugadas) são escolhidas para tentar assegurar que as direcções que já foram minimizadas permaneçam minimizadas.

As direcções conjugadas são calculadas supondo que a superfície do erro é quadrática, que geralmente não é o caso. Porém, é uma boa suposição, caso o algoritmo descubra que a direcção de procura da linha actual não está realmente em declive, então, simplesmente calcula a linha de descida mais íngreme e reinicia a procura naquela direcção. Uma vez que o ponto próximo de um mínimo é achado, a suposição quadrática rectifica que a solução encontrada é verdadeira e o mínimo pode ser localizado muito rapidamente.

A DGC possui uma desvantagem em relação ao tempo de cálculo por iteração (tipicamente de 3 – 10 vezes maior) quando comparado com o algoritmo BP. Maiores detalhes do algoritmo DGC na forma não linear são vistos no Anexo A.

CAPÍTULO III: Dados Experimentais

Este trabalho mostra um estudo do tipo ecológico ao longo de uma série de dez anos sobre a variabilidade dos internamentos associados ao grupo de doenças respiratórias na população da Cidade de São Paulo, levando em consideração as condições atmosféricas e poluição do ar. Neste Capítulo será abordado a utilização dos dados de morbidade respiratória, meteorológicas e poluição do ar.

3.1. Dados

Para este estudo foram utilizados três grupos de dados: hospitalares, meteorológicos e poluentes, durante o período correspondido entre 1 de Janeiro de 1993 a 31 de Dezembro de 2002. Os parâmetros meteorológicos, tais como temperatura do ar (máxima, mínima e média), humidade relativa e específica do ar (máxima, mínima e média), precipitação diária total, irradiação solar global diária, foram fornecidos pela Estação Meteorológica do Instituto de Astronomia, Geofísica e Ciências Atmosféricas (IAG), que está localizada na latitude de 23°39'S e longitude de 46°37'W em uma altitude de 799,2 m.

Foram utilizados ainda, os dados de poluentes atmosféricos como material particulado inalável (MP₁₀), ozono (O₃), dióxido de enxofre (SO₂), monóxido de carbono (CO) fornecidos pela Companhia Tecnológica de Saneamento Básico (CETESB), cartas sinópticas de superfície e reanálises do National Center Environmental Predictor (NCEP).

Os dados hospitalares utilizados correspondem ao número de entradas (morbidade) nos hospitais com doenças respiratórias, fornecidos pelo Sistema de Informações Hospitalares do Sistema Único de Saúde – SIH/SUS.

3.1.1. Dados de Morbilidade

Os registos de morbilidade foram referentes a todos os hospitais associados ao SUS já disponíveis junto ao Sistema de Informações Hospitalares do SUS, que constavam dos registos das Autorizações de Internamento Hospitalar (AIH's).

Na época, o Sistema Único de Saúde (SUS), contava com 112 hospitais credenciados na cidade de São Paulo, entre públicos e privados, que prestavam atendimento médico à parcela da população que não dispunha de planos de saúde privados, de carácter particular ou custeados por empresas (Braga, 1998). Segundo relatório publicado pela Fundação Sistema Estadual de Análise de Dados – SEADE (Fundação SEADE, 2002) sobre condições de vida na região metropolitana de São Paulo, esta população correspondia a aproximadamente 38 % dos habitantes.

Os internamentos hospitalares, bem como todos os procedimentos realizados durante o período de internamento, devem ser notificados ao SUS através dos preenchimentos das AIH's. Estas informações são então registadas e arquivadas. O banco de dados utilizado para este estudo é composto pelo registo dos pagamentos efectuados pelo SUS aos prestadores de serviço (Braga, 1998).

❖ Critérios de Inclusão das Admissões Hospitalares

Os arquivos reduzidos de AIH's mensais para o Estado de São Paulo, elaborados pelo SUS, possuem uma grande quantidade de informações que variam aproximadamente de 120.000 a 300.000 informações. Dentro dos critérios de inclusão optou-se, primeiramente, incluir somente AIH's que pertenciam ao município de São Paulo que possuíam códigos "MUNIC_MO" (município de localização do hospital) " igual a 355030, em seguida, inseriu-se AIH's com o código identificador igual a 1, ou seja, movimento de AIH's que correspondiam aos pagamentos normais efectuados pelo SUS referente a data de internamento dos pacientes, assim como as datas dos internamentos.

Todas as doenças do aparelho respiratório foram seleccionadas de acordo com a Classificação Internacional de Doenças – CID. Os anos de 1993 a 1997 o SUS adoptava a CID9 (9ª Revisão) que era codificado pelos diagnósticos

principais (CID9, Cap. VIII -- 460 a 519). Nos anos de 1998 a 2002 o SUS adoptava a CID10 (10ª Revisão) que era codificado pelos diagnósticos principais (CID10, Cap. VIII -- J000 a J999)

Nos estudos epidemiológicos é quase rotina especificar três características de qualquer pessoa – a idade, o sexo e o grupo étnico ou raça (Mausner & Bahn, 1999). Neste trabalho será especificado apenas a idade e sexo do paciente proveniente das AIH's.

➤ **Classificação Etária e Sexo**

Acima de tudo, a idade é a determinante mais importante entre as variáveis da epidemiologia descritiva. As taxas de morbidade e mortalidade de quase todas as afecções apresentam uma relação qualquer com esta variável, assim como o aspecto mais evidente da análise das taxas de doenças segundo o sexo é o contraste entre as taxas de mortalidade e morbidade. As taxas de mortalidade são mais altas nos homens que nas mulheres, mas as taxas de morbidade são geralmente mais elevadas nas mulheres (Mausner & Bahn, 1999).

Abaixo seguem a conotação das variáveis de morbidade adoptadas neste trabalho segundo o critério da classificação etária e sexo:

- ✓ RESP1M, RESP1F e RESP1MF – Morbidade por doenças respiratórias em crianças de 0 a 13 anos de idade (RESP1) segundo o sexo masculino (M), feminino (F) e ambos os sexos (MF), respectivamente;
- ✓ RESP2M, RESP2F e RESP2MF – Morbidade por doenças respiratórias em adolescentes de 14 a 19 anos de idade (RESP2) segundo o sexo masculino (M), feminino (F) e ambos os sexos (MF), respectivamente;
- ✓ RESP3M, RESP3F e RESP3MF – Morbidade por doenças respiratórias em adultos jovens de 20 a 44 anos de idade (RESP3) segundo o sexo masculino (M), feminino (F) e ambos os sexos (MF), respectivamente;
- ✓ RESP4M, RESP4F e RESP4MF – Morbidade por doenças respiratórias em adultos maduros de 45 a 64 anos de idade (RESP4) segundo o sexo masculino (M), feminino (F) e ambos os sexos (MF), respectivamente;

- ✓ RESP5M, RESP5F e RESP5MF – Morbidade por doenças respiratórias em idosos com ou mais de 65 anos de idade (RESP5) segundo o sexo masculino (M), feminino (F) e ambos os sexos (MF), respectivamente;
- ✓ RESP6M, RESP6F e RESP6MF – Morbidade por doenças respiratórias envolvendo todas as faixas etárias (RESP6) segundo o sexo masculino (M), feminino (F) e ambos os sexos (MF), respectivamente;

3.1.2. Dados Meteorológicos

A inclusão das variáveis meteorológicas é de fundamental importância para verificar a acção do tempo sobre a população que deram entrada nos hospitais com doenças respiratórias. As variáveis utilizadas nas análises serão sucintamente descritas a seguir.

❖ Temperatura do Ar

A temperatura do ar é o resultado do balanço energético entre a superfície da terra e a atmosfera, variando com a altitude e as condições atmosféricas, ou seja, a temperatura do ar pode variar em decorrência de actuações de sistemas sinópticos, sazonalidade e até mesmo dentro de perímetros urbanos (bairros) que possuam características distintas tais como vegetação, ocupação do solo, arquitectura e outros factores peculiares do local.

Os registos de temperatura são efectuados entre 07h e 24h com a leitura do capilar de mercúrio que constitui o bulbo seco de um psicrômetro de aspiração tipo Assmann com motor eléctrico (Fuess, modelo 32e, N.º D-3727). Um termógrafo de anel bimetálico (Fuess, modelo 79, N.º F-2243), com rotação diária, permitem as estimativas dos registos de temperatura do ar dos valores horários entre 01h e 06h. Valores extremos de temperatura são obtidos a partir de observações efectuadas com um termómetro de máxima com capilar de mercúrio e de um termómetro de mínima com capilar de álcool (Fuess), essas observações são registradas às 07h, 14h, 21h e 24h; no caso do termómetro de máxima, também, às 15h. Todos esses instrumentos encontram-se instalados no abrigo principal do cercado meteorológico.

❖ Humidade Relativa do Ar

Definida como a relação entre a quantidade real de vapor de água em um dado volume de ar, para a quantidade máxima que estaria presente se o ar estivesse saturado à mesma temperatura e pressão. É dada pela relação entre a pressão parcial de vapor de água e a pressão parcial de saturação do vapor de água, sendo dada em percentagem. Matematicamente é expressa por:

$$U = 100 \times \frac{e'}{ew(P,t)} \quad [\%] \quad (3.1)$$

e' É a pressão parcial de vapor de água, que é dado por:

$$e' = ew(P,t) - 0,000653 \times (1 + 0,000944 \times tw) \times P \times (t - tw) \quad [\text{hPa}] \quad (3.2)$$

Onde:

P → pressão atmosférica do ambiente

t → temperatura do bulbo seco

tw → temperatura do bulbo húmido

Enquanto que, $ew(P,t)$ é a pressão de saturação do vapor de água à temperatura do bulbo seco, expresso por:

$$ew(P,t) = f(P) \times ew(t) \quad [\text{hPa}] \quad (3.3)$$

Onde:

$$ew(t) = 6,112 \times \exp \left[17,62 \times \frac{t}{(243,12 + t)} \right] \quad [\text{hPa}] \quad (3.4)$$

Os registos de humidade são efectuados entre 07h e 24h com a leitura dos capilares de mercúrio que constituem os bulbos seco e húmido de um psicrômetro de aspiração tipo Assmann com motor eléctrico (Fuess, modelo 32e, N.º D-3727). Um higrógrafo dotado de harpa de fio de cabelo (Fuess, modelo 77g, N.º 2243), com rotação diária, permitem as estimativas dos registos de humidade do ar dos valores horários entre 01h e 06h. Todos esses instrumentos encontram-se instalados no abrigo principal do cercado meteorológico.

❖ Humidade Específica do Ar

É definida como sendo a razão entre a massa de vapor de água pela massa total de ar húmido num mesmo volume, expressa em gramas de vapor de água por quilograma de ar húmido, é obtida a partir da pressão parcial do vapor de água e da pressão atmosférica, matematicamente pode ser expressa por:

$$q = \frac{\left(\frac{M_v}{M_d}\right) \times \left(\frac{e}{P}\right)}{1 - \left(1 - \frac{M_v}{M_d}\right) \times \left(\frac{e}{P}\right)} \quad (3.5)$$

Sendo:

M_v e M_d são os valores assumidos pelo peso molecular da água (18,015g) e o ar seco (28,97g)

❖ Precipitação Diária Total

As observações foram obtidas através de um pluviógrafo de Hellmann (Lambrecht, modelo L-1507, N.º 54085) instalado no cercado meteorológico onde é empregada, posteriormente: (a) a avaliação de totais de horários de precipitação entre cada hora inteira, e (b) na estimativa de valores de intensidade de precipitação. Três pluviómetros encontram-se instalados à sua proximidade, sendo um do tipo “Ville de Paris” e dois do tipo “Paulista”; a diferença entre as leituras efectuadas às 07h, 14h, 21h e 24h com estes pluviómetros permite a dedução do total de precipitação em cada período.

❖ Irradiação Solar Global Diária

A quantidade de energia que atinge uma superfície plana, por unidades de área e de tempo é denominada de irradiância. A integração de irradiâncias sobre um determinado período de tempo é denominada irradiação. A irradiância solar global compreende quantidades de energia radiante provenientes do disco solar aparente bem como de todo o céu situado acima da superfície horizontal de

interesse. A irradiação solar global diária envolve a integração, ao longo de todo o período diurno, de valores instantâneos da irradiância solar global.

A irradiância solar global diária foi obtida a partir dos registos de um actinógrafo bimetalico (Fuess, modelo 58d, N.º C-9869) instalado no alto da torre da Estação. Os registos são analisados em termos do valor assumido pela área compreendida entre a linha de registo e a linha obtida por interpolação linear entre os valores mínimos de registo durante as noites anterior e posterior ao período diurno de interesse, onde o valor desta área é obtido mediante planimetria mecânica, efectuada diversas vezes e por técnicos diferentes.

❖ **Velocidade e Direcção do Vento**

Os registos de velocidade e direcção do vento foram obtidos através de um anemógrafo tipo universal (Fuess, modelo 82b, N.º A-9506) instalado na torre da Estação, onde é empregado para: (a) observação da direcção predominante e da velocidade instantânea, a cada hora entre 07h e 24h; e (b) a posteriori, obtenção da direcção predominante e da velocidade média entre cada hora inteira. A velocidade média entre cada hora inteira é deduzida a partir do registo da linha "caminho percorrido" associado ao deslocamento de uma parcela de ar imaginária.

Maiores detalhes sobre os procedimentos de medidas e instrumentações meteorológicas podem ser encontrados no Boletim Anual "Observações de Superfície Efectuadas na Estação Meteorológica" (ISSN 1415-4374).

3.1.3. Poluentes Atmosféricos

Durante o período de estudo, a CETESB dispunha de 13 estações fixas (figura 3.1), da rede automática, na Cidade de São Paulo. Porém, nem todas as estações registavam medidas de todos os poluentes. Apenas a estação do Parque D. Pedro, da Moóca e Congonhas (aeroporto), mediram todos os poluentes que constam nos padrões nacionais de qualidade do ar. As demais estações registavam em média dois ou três poluentes. Para todos os poluentes foram utilizadas as médias diárias dos registos das concentrações que compreendiam um período de 24 horas, que se iniciava às 16 horas do dia anterior até às 15 horas do próximo dia.

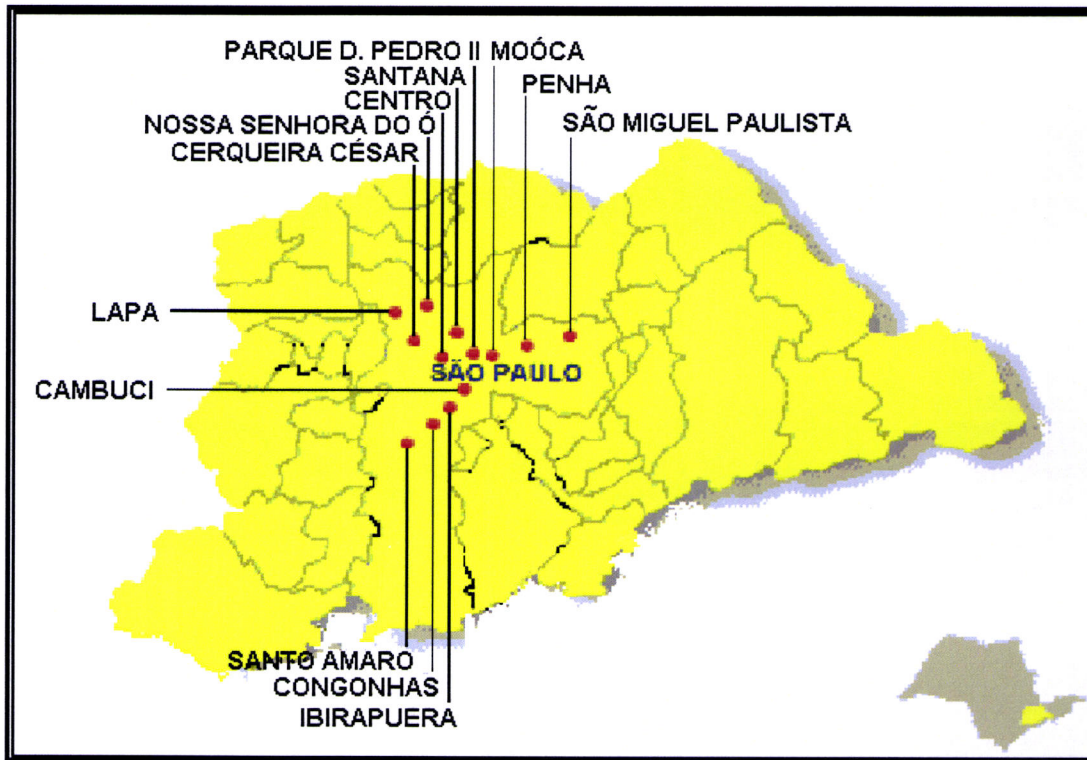


Figura 3.1 – Rede Telemétrica da CETESB na Cidade de São Paulo.

❖ Material Particulado Inalável (MP_{10})

Para o MP_{10} foi utilizada a média de 24 horas, com registos realizados em 12 estações. O método utilizado para medir os níveis deste poluente foi o monitor de radiação β , onde é avaliada a atenuação da intensidade de radiação β que passa através de material sólido retido em um filtro, através do qual a amostra de ar foi filtrada. Com isto é possível aferir a concentração de partículas inaláveis na amostra de ar analisada. São realizadas medições seriadas, calculando-se a média do período.

❖ Ozono (O_3)

O ozono foi medido em 4 estações, registando-se a maior média horária. O método utilizado para medir o ozono foi a quimiluminescência, que é a medida da energia liberada em forma de luz em uma reacção química exotérmica. No caso, a reacção do óxido de azoto com ozono libera dióxido de azoto (NO_2), produzindo liberação de energia na forma de luz e oxigénio e, também, através de colisão com outras moléculas presentes. Fixando-se um dos reagentes, a energia luminosa liberada é proporcional à quantidade do outro reagente.

❖ Dióxido de Enxofre (SO₂)

No caso do SO₂, também foi utilizada a média de 24 horas das medições realizadas em 13 estações. O método para quantificar as concentrações deste poluente foi a coulometria, que é a medida da corrente eléctrica produzida quando um forte oxidante, ou um poluente gasoso redutor, reage com iodeto de potássio ou solução de brometo em uma célula electroquímica. Este valor aferido é proporcional à quantidade de SO₂ que entrou no sistema.

❖ Monóxido de Carbono (CO)

Os registos dos níveis de CO foram realizados em cinco estações, registando-se a maior média móvel de 8 horas. O método empregado para medir os níveis de CO é o infravermelho não dispersivo. É um monitor composto de duas câmaras, uma contendo uma amostra de gás na qual está presente o CO (ar ambiente) e outra com um gás transparente, sem CO, que são atingidas por feixes de luz infravermelha. A presença de CO no ar ambiente irá absorver radiação, diminuindo a produção de calor na célula do monitor. Esta alteração é comparada com a produção de calor registada na outra célula pela luz infravermelho que passou antes pela câmara sem CO. Esta diferença indica a concentração de CO na amostra de ar.

Maiores detalhes sobre os métodos utilizados para medir os níveis dos poluentes podem ser encontrados no "Handbook of Air Pollution Analysis" (Perry; Young, 1977).

Neste estudo foi empregada a média (espacial) diária dos poluentes, para cada poluente, medidos em várias estações distribuídas pela Cidade de São Paulo, como indicativo dos níveis diários de poluição (Böhm; 1989; Lin, 1997; Braga; 1998 e Gonçalves *et al.*, 2005). A utilização das médias diárias dos poluentes, pode ser justificada pelo comportamento observado em relação aos níveis de PM₁₀ em diferentes regiões da cidade. De todos os cinco poluentes presentes neste estudo, o PM₁₀ é o que apresenta a menor percentagem de emissão por fontes móveis, com 40% de sua concentração sendo de origem automotiva directa.

3.1.4. Cartas Sinópticas de Superfície

Cartas de pressões (isobáricas) elaboradas pela Directoria de Hidrografia e Navegação do Ministério da Marinha (DHN, 1993 a 2002) no horário das 12:00 TMG.

3.1.5. Reanálises do NCEP

As reanálises serviram de suporte para avaliação das posições dos sistemas sinópticos que actuavam ou que influenciavam sob a Cidade de São Paulo durante o período de estudo.

3.2. Análises e Identificações dos Sistemas Sinópticos

A utilização das cartas de pressões em superfície, e reanálises do NCEP, serviram para a identificação e posição das frentes e anticiclones em relação a Cidade de São Paulo, onde foram utilizados os mesmos critérios adoptados por Abreu, M. L., 1993; Manfredini, L. M., 1988; Sánchez-Ccoyllo, 1998, Conde (2001) e Gonçalves et al. (2005).

- 1) Posição das frentes e dos anticiclones: determinados pelas cartas de superfície e dados da reanálise foi possível a determinação da área de domínio dos sistemas e do tipo de circulação que predominava na região de São Paulo.
- 2) Após a análise sinóptica, classificaram-se seis sistemas sinópticos que actuavam na região:
 - ✓ AS → Alta Subtropical do Atlântico Sul (figura 3.2 a). Este sistema de alta pressão é caracterizado por um regime semi-estacionário. Geralmente, este sistema provoca a inibição de nuvens e, conseqüentemente, a ausência de precipitações.
 - ✓ AP → Alta Polar (figura 3.2 b) que foi caracterizada pelo deslocamento marítimo do anticiclone extratropical. Este sistema de alta pressão geralmente provoca um declínio de temperatura e um pequeno aumento da humidade no período da manhã no Inverno;

- ✓ AC → Alta Continental (figura 3.2 c) caracterizada pelo deslocamento do anticiclone extratropical (seco e frio) totalmente sobre o continente;
- ✓ F → Sistema Frontal (figura 3.2 d) que se encontrava sobre a Cidade de São Paulo. Este sistema geralmente é acompanhado por precipitações. No Inverno, durante a actuação deste sistema, é comum apenas provocar um aumento da nebulosidade e humidade mas sem a ocorrência de precipitação.
- ✓ PrF → Sistema Pré-Frontal (figura 3.2 e) que se encontrava nos Estados do Rio Grande do Sul e/ou Paraná (sistema frontal antes de atingir São Paulo). Às vezes esse tipo de sistema não atinge o Estado São Paulo, principalmente no Inverno. Esse sistema muitas vezes ajuda a organizar a convecção sobre o Estado de São Paulo ou intensificar cavados em níveis superiores, porém, as vezes indicam situações de bloqueios atmosféricos.
- ✓ PoF → Sistema Pós-Frontal (figura 3.2 f) que se encontrava nos Estados do Rio de Janeiro, Minas Gerais e Espírito Santo (sistemas frontais depois de atingir São Paulo). Esta configuração sinóptica ocorre com uma maior frequência no Verão.

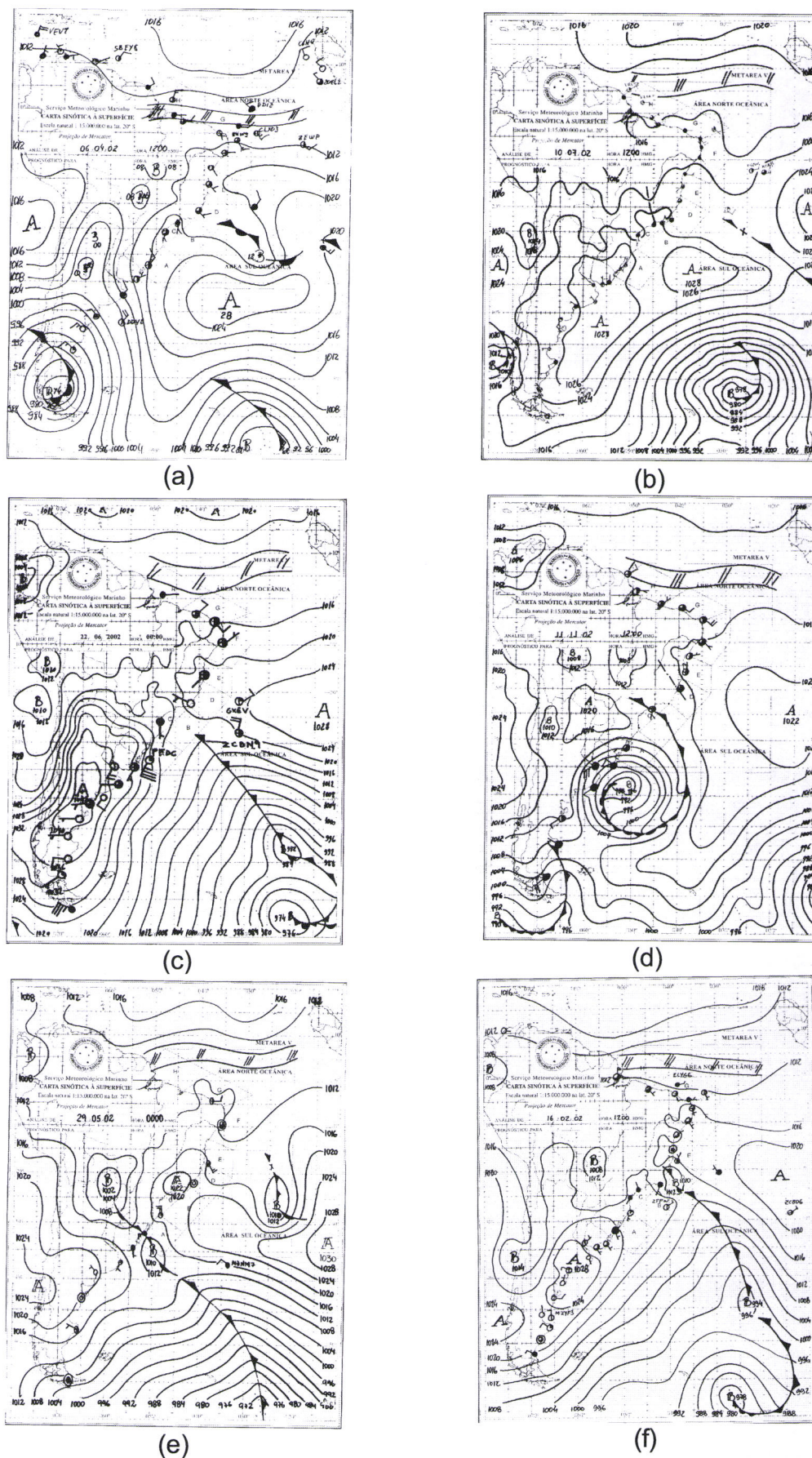


Figura 3.2 – Identificação dos Sistemas Sinóticos segundo as Cartas do DHN.

CAPÍTULO IV: Metodologia Desenvolvida

Conforme visto no *Capítulo 1*, vários estudos comprovam que a actuação do tempo/clima em seres humanos provocam várias reacções adversas sobre a saúde, principalmente, em crianças e idosos, que habitam grandes centros urbanos. Além disso, as condições de tempo actuam de modo favorável ou desfavorável para a dispersão de poluentes. As condições meteorológicas de grande escala se fazem necessárias para a avaliação do tempo e a disposição destas sobre uma determinada área.

Neste Capítulo será abordado a análise dos dados através de técnicas estatísticas multivariadas e a utilização de redes neuronais artificiais para o desenvolvimento de uma metodologia, que satisfaça a elaboração de um Índice de Risco Meteorotrópico (IRM), que quantifique e categorize as admissões hospitalares devido a doenças do aparelho respiratório para a área de estudo.

4.1. Métodos Estatísticos Multivariados

Os métodos estatísticos multivariados possuem várias técnicas distintas. Neste trabalho serão utilizadas duas técnicas: análise em componentes principais e análise de taxonomia numérica ou análise de “Cluster” (agrupamento).

O emprego da Análise de Cluster será utilizado em carácter exploratório para verificar a similaridade das variáveis envolvidas, enquanto que a Análise em Componentes Principais será empregada em carácter confirmatório. Essas duas técnicas deverão corroborar as informações obtidas pela rede neuronal. As análises de ACP e AC foram realizadas através do software “STATISTICA 6.0”.

4.1.1. Análise em Componentes Principais

Os métodos multivariados são bastante úteis para pesquisas que contabilizam um grande conjunto de dados que possuem medidas de variáveis

em várias unidades experimentais. O objectivo primário das análises multivariadas é resumir grandes quantidades de dados através de alguns parâmetros que serão obtidos, posteriormente, após o emprego da análise. O interesse de análise multivariada é encontrar relações entre: 1) a resposta da variável, 2) as unidades experimentais, e 3) a resposta da variável com as unidades experimentais. Essas relações existem quando algumas das variáveis têm uma resposta comum entre si.

Análise em Componentes Principais (ACP) transforma uma série de variáveis originais dentro de uma série menor de combinações lineares que explicam a maior parte da variância explicada da série original. O propósito da ACP é determinar os factores (Componentes Principais – CP's) de maneira a explicar o quanto a variância total da série pode ser explicada com poucos desses factores (Wilks, 1995).

As CP's são extraídas para que a primeira componente principal ($CP_{(1)}$), explique a maior parte da variação nos dados. $CP_{(1)}$ é a combinação linear das variáveis observadas. A segunda Componente Principal ($CP_{(2)}$) é a combinação linear dos pesos das variáveis observadas que não é correlacionada com a primeira combinação linear no qual explica o máximo da variação total restante que nem sempre é explicada por $CP_{(1)}$. Em geral, a m-ésima componente principal ($CP_{(m)}$) é a combinação linear dos pesos das variáveis observadas X_j ($j = 1, 2, \dots, p$), representado na equação abaixo.

$$CP_1 = w_{11}X_1 + w_{12}X_2 + \dots + w_{1p}X_p \quad (4.1)$$

Onde w é o peso da variável observada.

4.1.2. Entrada de Dados

O conjunto de dados utilizados neste trabalho possuem unidades físicas (escalas) diferentes, então, utilizou-se a normalização ou anomalia padronizada, com a finalidade de remover a influência da localização e extensão do conjunto de dados. Neste processo, as unidades físicas do conjunto de dados são canceladas, então, as variáveis normalizadas são sempre quantidades não dimensionais, onde foi usada a seguinte equação:

$$Z = \frac{X_j - \bar{X}}{\sigma_x} = \frac{X'}{\sigma_x} \quad (4.2)$$

Onde:

$X_j \rightarrow$ é a j – ésima variável da amostragem

$\bar{X} \rightarrow$ é a média da amostragem

$Z \rightarrow$ é a variável normalizada

$\sigma_x \rightarrow$ é o desvio padrão da amostra

Este processo converte cada observação dos dados originais em um valor padronizado; onde a média é igual a zero e desvio padrão igual a 1

4.1.3. Extracção das Componentes Principais

Suponha que os vectores das observações $X'=(X_1, X_2, \dots, X_p)$ tenham uma matriz de covariância – variância (Σ); por conveniência matemática, assume-se que a média de Z_i é zero para todo $i=1, 2, \dots, p$. Para encontrar a primeira componente principal $Y_{(1)}$, procura-se os coeficientes de vectores $\gamma'=(\gamma_1, \gamma_2, \dots, \gamma_p)$ tal que a variância de $\gamma'X=1$. A forçante que o produto escalar do coeficiente do vector γ com ele mesmo igual a um é utilizada para evitar um aumento na variância de $\gamma'X$ (Wilks, 1995).

Dadas p variáveis originais X_1, X_2, \dots, X_p com n elementos cada uma, a matriz de dados é escrita como:

$$X = \begin{bmatrix} X_{11} & X_{12} & \dots & X_{1p} \\ X_{21} & X_{22} & \dots & X_{2p} \\ \dots & \dots & \dots & \dots \\ X_{n1} & X_{n2} & \dots & X_{np} \end{bmatrix} \quad (4.3)$$

A partir dessa matriz, obtém-se a matriz de correlações Σ ($p \times p$):

$$\Sigma = \begin{bmatrix} 1 & S_{12} & \dots & S_{1p} \\ S_{21} & 1 & \dots & S_{2p} \\ \dots & \dots & \dots & \dots \\ S_{p1} & S_{p2} & \dots & 1 \end{bmatrix} \quad (4.4)$$

Cada elemento S_{jk} é obtido através da expressão:

$$S_{jk} = \frac{\frac{1}{n} \sum_{i=1}^n (x_{ij} - \bar{x}_j)(x_{ik} - \bar{x}_k)}{s_{xj} s_{xk}} \quad (4.5)$$

Onde S é o coeficiente de correlação e o σ é o desvio padrão dado por:

$$s_{xj} = \sqrt{\frac{\sum_{i=1}^n (x_{ij} - \bar{x}_j)^2}{n}} \quad (4.6)$$

$$s_{xk} = \sqrt{\frac{\sum_{i=1}^n (x_{ik} - \bar{x}_k)^2}{n}} \quad (4.7)$$

Esta matriz fornece p autovalores (λ_j), onde $\lambda_1 \geq \lambda_2 \geq \dots \geq \lambda_p$. A cada um dos p autovalores corresponde um autovector A_j , sendo:

$$A_j = (a_{1j}, a_{2j}, \dots, a_{pj}), \quad j = 1, 2, \dots, p \quad (4.8)$$

Que são as componentes da base ortogonal utilizada.

Cada um dos p componentes da base fornece uma explicação e_j da variância total proporcional ao seu autovalor, de modo que:

$$e_j = \frac{\lambda_j}{\lambda_1 + \lambda_2 + \dots + \lambda_p} \times 100\% \quad (4.9)$$

Para a determinação do número de autovalores significativos, foi utilizado o critério proposto por Kaiser (1960), ou seja, retêm-se somente os factores com autovalores maiores que 1.

A correlação ente os autovalores e as respectivas variáveis é chamada de “factor loadings” (pesos dos factores). Os “factor loadings” representam a informação mais importante para a interpretação dos autovalores.

4.1.4. Rotação dos Factores Retidos

Para os p factores retidos, é realizado o critério da rotação VARIMAX que é o método mais comumente usado. Consideremos a notação:

$A = a_{jp}$ - matriz dos pesos (“loadings”) dos factores iniciais

$B = b_{jp}$ - matriz dos pesos (“loadings”) dos factores finais

$T = t_{qp}$ - matriz da transformação ortogonal tal que,

$$B = AT \quad (4.10)$$

A comunalidade de qualquer variável permanece invariante.

$$\sum_{p=1}^m b_{jp}^2 = \sum_{p=1}^m a_{jp}^2 = h_j^2 \quad j=1, 2, \dots, n \quad (4.11)$$

O critério VARIMAX de rotação procura simplificar os factores, maximizando a variância dos pesos (pesos dos novos factores divididos pelas raízes quadradas das respectivas comunalidades), isto é, maximizando a função:

$$V = n \sum_{p=1}^m \sum_{j=1}^n \left(\frac{b_{jp}}{h_j} \right)^{4m} - \sum_{p=1}^m \left(\sum_{j=1}^n \frac{b_{jp}^2}{h_j^2} \right)^2 \quad (4.12)$$

Esse processo de rotação dos factores iniciais tem como objectivo tornar mais clara a separação entre os factores (Keiding *et al.*, 1986).

4.1.5. Análise de “Cluster”

A análise de “Cluster” é a denominação de um grupo de técnicas multivariadas, cujo propósito principal é conter variáveis que estão baseadas nas características que possuem. A análise de agrupamento classifica as variáveis de

modo que cada variável possua características similares em um conglomerado que considera alguns critérios de seleções predeterminados.

O objectivo principal da análise de Cluster é definir a estrutura dos dados que ligam as observações mais similares em grupos.

A análise de “Cluster” (AC) é inicialmente uma ferramenta exploratória de análises de dados. Geralmente, os processos de AC são hierárquicos, ou seja, a construção hierárquica das séries de grupos se faz pela fusão de pares de grupos definidos anteriormente.

O processo tem início considerando-se n observações de X que não possuem estruturas de grupos, ou que as séries de dados constituem de n grupos contendo uma observação cada. O primeiro passo é encontrar os dois grupos que são próximos em seus espaços dimensionais – k , e combiná-los dentro de um novo grupo. Os n dados agrupados nos seus espaços dimensionais – k dentro de algum número G ($1 < G < n$), isto é, grupos que reflectem dados similares gerados pelo processo (Wilks, 1995).

❖ Medida de Distância e Método de Agrupamento

As medidas de distância na análise de Cluster, representam as similaridades das observações em relação a proximidade das variáveis, onde os valores altos indicam uma menor similaridade, enquanto que valores baixos indicam uma maior similaridade.

Existem várias medidas de distância, neste estudo optou-se em empregar a distância euclidiana.

Os procedimentos hierárquicos consistem na construção de uma estrutura em forma de árvore. Existem basicamente dois tipos de procedimentos para se obter os conglomerados hierárquicos: agrupamentos e divisivos. Neste estudo utilizou-se o método de agrupamentos, ou seja, cada observação começa dentro de seu próprio conglomerado. Em fases anteriores, os dois conglomerados mais próximos se combinam em um novo conglomerado, reduzindo assim o número de conglomerado passo a passo.

Uma característica importante dos procedimentos hierárquicos é que os resultados obtidos em um passo prévio sempre precisam ser acoplados dentro dos resultados do passo seguinte, criando algo semelhante a uma árvore.

➤ Distância Euclidiana

Os agrupamentos devem consistir de pontos separados pela menor distância, relativa as distâncias entre os agrupamentos. O mais intuitivo e comumente usado é a distância euclidiana no espaço dimensional – k dos dados. Assim, a distância entre dois pontos x_i e x_j é:

$$d_{ij} = \|x_i - x_j\| = \left[\sum_{k=1}^k (x_{i,k} - x_{j,k})^2 \right]^{\frac{1}{2}} \quad (4.13)$$

➤ Método de Ward (ou Método da Variância Mínima)

Este método une os pares de agrupamentos que resultarão na soma mínima das distâncias quadráticas entre os pontos e os centróides dos respectivos grupos, somados sobre os grupos resultantes, isto é, para todas as possíveis maneiras de união dos grupos $G+1$ tornarem-se grupos G (Wilks, 1995), onde esta união é feita para minimizar, que é obtida pela seguinte equação:

$$W = \sum_{g=1}^G \sum_{i=1}^{n_g} \|x_i - \bar{x}_g\|^2 = \sum_{g=1}^G \sum_{i=1}^{n_g} \sum_{k=1}^k (x_{i,k} - \bar{x}_{g,k})^2 \quad (4.14)$$

Onde o centróide (ou grupo médio) para o grupo unido recentemente é recalculado usando os dados para ambos os grupos separados anteriormente, sendo que antes as distâncias quadráticas são calculadas.

4.2. Rede Neuronal Artificial Multicamadas – Metodologia Desenvolvida

Esta secção abordará os fundamentos vistos no Capítulo 2 para a construção e preparação da rede neuronal para a viabilização dos prognósticos das admissões hospitalares e elaboração do Índice de Risco Meteorotrópico (IRM). Todos os processos para o desenvolvimento das RNA's foram realizados através do software "ALYUDA NEURO INTELLIGENCE 2.2".

4.2.1. Tratamento da Série de Dados

A preparação dos dados é o passo mais importante quando se trabalha com redes neuronais. Os dados utilizados neste estudo seguiram os seguintes critérios em relação ao problema e sua complexidade a ser analisada:

4.2.1.1. “Missing Values e outliers”

Os “missing values” ou dados faltantes foram todos removidos, enquanto que os outliers (espúrios) foram considerados na solução do problema devido o estudo abordar um grupo de variáveis ambientais (meteorológicas e poluentes), sabendo-se a priori, que do ponto de vista meteorológico este conjunto de variáveis são grandezas que possuem uma variabilidade bem acentuada dentro de um ciclo diário ou sazonal.

4.2.1.2. Variáveis de Entrada e Saída

As variáveis de entrada corresponderam a 62 camadas de entradas que reuniam os dados meteorológicos, poluentes e variáveis de periodicidade, enquanto que a variável de saída (valor alvo) foram os dados de morbidade respiratória segundo a classificação etária e o sexo.

As variáveis meteorológicas entraram como variáveis numéricas e categóricas.

- ✓ Variáveis Meteorológicas Numéricas (9 entradas): Pressão (max e min), temperatura (max e min), humidade relativa e específica (max e min) e irradiação solar global diária.
- ✓ Variáveis Meteorológicas Categóricas: Velocidade e direcção do vento, precipitação diária total e classificação sinóptica.
 - Velocidade do Vento (4 entradas) – Foram classificadas segundo a escala de Beaufort:
 - Calmo: Velocidade do vento entre 0 a 1 km/h;
 - Brisa1: Velocidade do vento entre 2 a 6 km/h;

- Aragem: Velocidade do vento entre 7 a 12 km/h;
- Brisa 2: Velocidade do vento entre 13 a 18 km/h.
- Direcção do Vento (17 entradas): classificado segundo a rosa-dos-ventos (4 pontos cardeais, 4 pontos colaterais e 8 pontos subcolaterais) incluindo a condição de vento calmo (sem direcção).
- Precipitação Diária Total (8 entradas): Foram classificadas segundo o critério abaixo.
- NPRP – Ausência de precipitação;
- TRPRP – Traços de precipitação [0,1 mm a 0,9 mm];
- PRPL – Precipitação leve [1 mm a 5 mm];
- PRPFR – Precipitação fraca [6 mm a 10 mm];
- PRPMD – Precipitação moderada [11 mm a 20 mm];
- PRPFT – Precipitação forte [21 mm a 30 mm];
- PRPMFT – Precipitação muito forte [31 mm a 50mm];
- PRPSFT – Precipitação super forte [acima dos 50 mm].

Classificação Sinóptica (6 entradas): classificado segundo a secção 3.2.

Os poluentes CO, MP₁₀, SO₂ e O₃ entraram como variáveis numéricas (4 entradas).

As manifestações biológicas dos efeitos da poluição e das condições atmosféricas dominantes em uma determinada região, apresentam um comportamento importante nas afecções do aparelho respiratório sobre a saúde de uma população. As admissões hospitalares por doenças respiratórias, normalmente, apresentam uma resposta directa em relação ao ciclo sazonal (factores atmosféricos e poluentes) e ciclo semanal (agentes socio-económicos) ou conjugação de ambos.

Como essas manifestações demonstram uma dependência em relação ao ciclo semanal e sazonal adotou-se a inclusão das variáveis de periodicidade. A inclusão da periodicidade, nesse caso, é uma informação importante na redução do erro durante o prognóstico gerado pela rede.

As variáveis de periodicidades são citadas abaixo:

- Variável Sazonalidade (4 entradas): Essa variável foi categorizada como: Verão, Outono, Inverno e Primavera;
- Variável Dias da Semana (7 entradas): Essa variável foi categorizada como: segunda-feira, terça-feira, quarta-feira, quinta-feira, sexta-feira, sábado, domingo;
- Variável Data (2 entradas): Esta variável identifica o padrão da periodicidade na camada de entrada da rede neuronal sendo codificada em 2 entradas conforme a fórmula abaixo.

$$\text{Entrada}_1 = \text{sen}\left(\left(\frac{X}{P}\right) \times 2\pi\right)$$

$$\text{Entrada}_2 = \text{cos}\left(\left(\frac{X}{P}\right) \times 2\pi\right)$$

Onde:

X – Data

P – periodicidade – 7

4.2.2. Partição dos Dados

Os dados foram divididos dentro de três séries: série de treinamento, série de validação e série de teste.

- a) Série de treinamento: é uma amostra da série de dados usada para o treinamento da rede neural, i.e., para o ajustamento dos pesos da rede. A amostra utilizada compreendeu 68 %, no período entre 01/01/1993 a 17/02/2000.
- b) Série de Validação: é uma amostra da série dos dados usado para aperfeiçoar a topologia da rede ou parâmetros da rede para diferentes pesos, por exemplo, é usada para definir o número de unidades escondidas, para detectar o momento

quando o desempenho da rede neural começou a deteriorar. A série de validação calcula a perda de generalização, retendo a melhor rede (a rede com o menor erro na série de validação). A amostra utilizada compreendeu 16 % dos dados no período entre 18/02/2000 a 25/07/2001.

c) Série de teste: é uma porção da série de dados utilizada somente para testar a performance que a RN executará sobre novos dados. A série de teste é usada depois que a rede estiver pronta (treinada), para testar que erros acontecerão durante aplicação da rede no futuro. Esta série não é usada durante a fase do treino, sendo assim considerada como a composição de novos dados inseridos pelo usuário para a aplicação da RN. A amostra utilizada compreendeu 16 % dos dados no período entre 26/07/2001 a 31/12/2002.

4.2.3. Pré-Processamento dos Dados

O Pré-processamento dos dados significa a modificação dos dados de uma forma satisfatória e compreensível antes da RN ser alimentada. Este pré-processamento é realizado através do escalonamento e/ou codificação de valores numéricos através da seguinte fórmula:

$$SF = (SR_{\max} - SR_{\min}) / (X_{\max} - X_{\min}) \quad (4.15)$$

$$X_p = SR_{\min} + (X - X_{\min})^{SF} \quad (4.16)$$

Onde:

X – valor actual de uma coluna numérica

Xmin – valor actual mínimo da coluna

Xmax – valor actual máximo da coluna

SRmin – limite inferior de amplitude da escala

SRmax – limite superior de amplitude da escala

SF – factor de escala

Xp – valor pré processado

Para as colunas de entradas e saídas utilizou-se dois tipos de funções de activação: logística [0, 1] e tangente hiperbólica [-1, 1].

4.2.4. Método de Procura da Arquitectura da RN

Para a complexidade do problema proposto neste estudo para se obter uma boa performance no treino da RN, optou-se empregar o método da procura exaustiva. Este método faz uma pesquisa exaustiva entre todas as topologias da série especificada pelo pesquisador. Neste trabalho foi utilizado o seguinte critério para a procura da melhor arquitectura da RN:

- A arquitectura da rede apresentava 62 unidades na camada de entrada, 150 unidades na camada escondida e 1 unidade na camada de saída [62-150-1];
- Funções de activações logísticas e hiperbólica;
- Para o melhor critério de ajustamento foi aplicado o teste da correlação de Pearson. O maior valor da correlação de Pearson apresentava a melhor topologia da RN;

Para cada procura foram utilizados 500 interacções com 3 retreinamentos conforme a morbilidade respiratória segundo a classificação etária e sexo.

4.2.5. Treino das RN's

O principal objectivo para o treinamento de uma RN é a minimização da função de erro, onde o valor da função de erro é usado para determinar a qualidade da RN.

A soma dos quadrados é a função de erro mais utilizada e que apresenta um melhor ajuste para problemas de regressão. A função de erro soma dos quadrados é a soma das diferenças quadráticas entre o valor actual (valor alvo) e o valor de saída produzido pela RN.

O método aplicado neste estudo usa Redes Neurais com Regressão Generalizada (RNRG) devido ser mais rápido e sensível para contribuições irrelevantes durante o treino da rede.

Para o treino das redes foram utilizados as variáveis meteorológicas, poluentes e periodicidade na camada de entrada e morbidade respiratória segundo a faixa etária e o sexo na camada de saída.

Através da topologia da rede verificou-se que a arquitectura apresentou um elevado número de pesos (≈ 2000), no qual viabilizou que a rede a ser treinada seria através dos algoritmos backpropagation e descida do gradiente conjugada.

➤ **BackPropagation – BP**

O BackPropagation (BP) é um dos melhores e mais usado algoritmo de treinamento conhecido para redes neuronais, possui baixa exigência de memória que a maioria dos algoritmos, e normalmente alcança um nível de erro aceitável bastante depressa, embora seja muito lento para convergir correctamente a um erro mínimo. Pode ser usado na maioria dos tipos de rede, embora seja muito apropriado para treinar perceptrões de multicamadas, principalmente, com redes que apresentam elevado número de pesos.

Para o treinamento usando o algoritmo BP foram utilizadas as seguintes condições:

- ✓ Taxa de aprendizagem de 0,1 a 10;
- ✓ Taxa de momento de 0,1 a 5;
- ✓ Condição de parada do treino até 500 iterações;
- ✓ Funções de activação logística e tangente hiperbólica;
- ✓ Retenção e restauro da melhor rede;
- ✓ Variação do peso de 0,1 a 10;
- ✓ 3 Retreinamentos;
- ✓ Actualizações dos ajustes dos pesos foram online.

➤ Descida de Gradiente Conjugada

Para o treinamento usando o algoritmo DGC foram utilizadas as mesmas condições do algoritmo BP, excepto a taxa de aprendizagem e momento.

4.2.5.1. Prognósticos da Morbilidade Respiratória

A figura 4.1, mostra a esquematização da rede neuronal realizado para o prognóstico após os passos descritos na secção 4.2.5. Matematicamente o prognóstico de RESP pode ser descrito como:

$$Net = \sum_{i=1}^n x_i w_{ij} + b_k \quad (4.17)$$

Onde: Net = Sinal de n entrada (x_i) x peso (w_{ij}).

Enquanto que, a saída de um neurónio (output) será obtida em função do sinal de entrada (input), isto é, a saída será função do Net :

$$Y = Saída = f(Net) \quad (4.18)$$

Onde f é a função de activação logística (LOG) ou tangente hiperbólica (TGH).

Substituindo (4.17) na equação (4.18), temos:

$$Y = f\left(\sum_{i=1}^n x_i w_{ij} + b_k\right) \quad (4.19)$$

Conforme descrito no *Capítulo 2* o ajuste dos pesos é calculado pela equação (2.20):

$$w_{ij}(t+1) = w_{ij}(t) + \eta \delta_j(t) x_i(t) + \mu (w_{ij}(t) - w_{ij}(t-1))$$

Variando η (0,1 a 10), μ (0,1 a 5) e w (0,1 a 10) após 500 iterações até obter a melhor resposta da rede treinada, ou seja, encontra-se um nível de erro aceitável.

Reescrevendo a equação (4.19), temos a equação do prognóstico gerado pela rede neuronal após o aprendizado:

$$RESP = f\left(\sum_{i=1}^n (met + pol + per) \cdot w_{ij} + b_k\right) \quad (4.20)$$

Onde *met*, *pol* e *per* são as variáveis meteorológicas, poluentes e periodicidade, respectivamente.

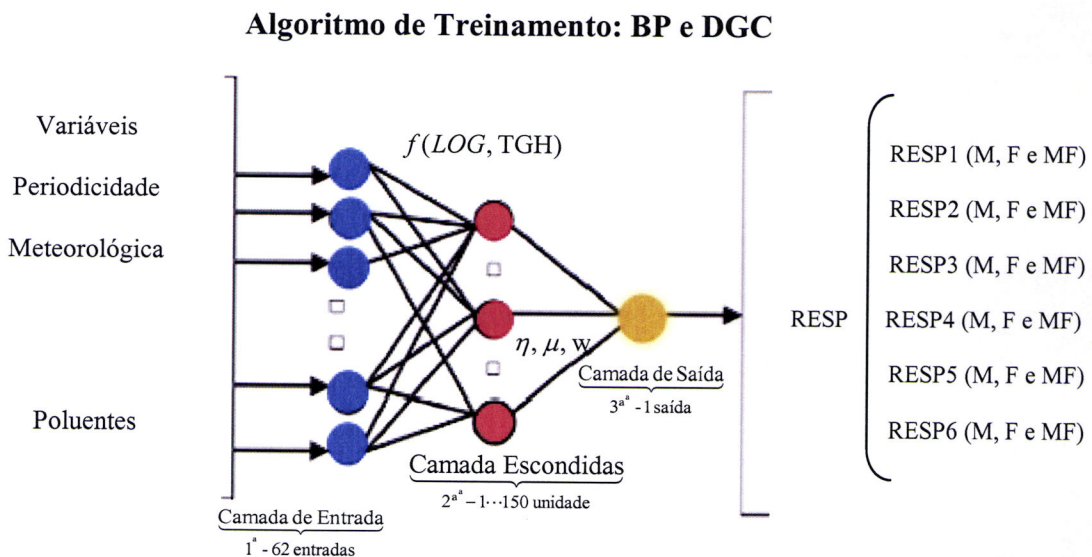


Figura 4.1: Esquematização da rede neural para o prognóstico da morbidade respiratória segundo a faixa etária e sexo.

4.2.5.2. Índice de Risco Meteorotrópico

Muitas medidas que resumem as propriedades do conjunto de dados usam os chamados “QUANTIS AMOSTRAIS”, (“*quantiles*” em inglês, ou também chamados de ‘fractiles’). Estes termos são essencialmente equivalentes ao termo também comum “PERCENTIL”. Um quantil amostral q_p é um número tendo a mesma unidade que o dado, o qual excede a proporção do dado dada pelo subscrito p , com $0 \leq p \leq 1$. O quantil amostral q_p pode ser interpretado aproximadamente como aquele valor do dado que excede um membro escolhido aleatoriamente do conjunto de dado, com probabilidade p . Analogamente, o

quantil amostral q_p poderia ser interpretado como o $[p \times 100]^{é-simo}$ percentil do conjunto de dados. A determinação dos quantis requer primeiro que os dados sejam ordenados. A notação utilizada comumente para designar os dados ordenados é a seguinte $\{x_{(1)}, x_{(2)}, x_{(3)}, x_{(4)}, \dots, x_{(n)}\}$, onde $x_{(1)}$ é o valor mais baixo e $x_{(n)}$ o mais alto.

Alguns quantis são utilizados mais comumente como a mediana (ou $q_{0.5}$) ou o 50º percentil. A mediana (M) é uma medida de localização do centro da distribuição dos dados, definida do seguinte modo: Ordenados os elementos da amostra, a mediana é o valor (pertencente ou não à amostra) que a divide ao meio, isto é, 50% dos elementos da amostra são menores ou iguais à mediana e os outros 50% são maiores ou iguais à mediana. Para a sua determinação utiliza-se a seguinte regra, depois de ordenada a amostra de n elementos ($X_{1:n}, X_{2:n}, \dots, X_{n:n}$). A expressão para o cálculo da mediana será:

$$M = \begin{cases} X_{\frac{n+1}{2}:n} & \text{se } n \text{ é ímpar} \\ \frac{1}{2} \left(X_{\frac{n}{2}:n} + X_{\frac{n+1}{2}:n} \right) & \text{se } n \text{ é par} \end{cases} \quad (4.21)$$

Se n é ímpar, a mediana é o elemento médio. Se n é par, a mediana é a semi-soma dos dois elementos médios

Outro uso tão comum quanto as medianas são os quantis $q_{0.25}$ e $q_{0.75}$. Usualmente, são chamados de quartil inferior e quartil superior. Estão localizados a meio caminho entre a mediana e os extremos $x_{(1)}$ e $x_{(n)}$.

Neste trabalho, a partir do valor prognosticado (valor desejado) gerado pela rede neuronal após o treinamento, optou-se adotar uma classificação das admissões hospitalares em quartis.

Portanto, a elaboração do Índice de Risco Meteorotrópico (IRM) associado às morbidades respiratórias segundo a classificação etária e o sexo, baseou-se na divisão dos dados observados (valor alvo) em 4 quartis, levando em consideração a seguinte classificação:

IRM 1= $RESP \leq q_{0.25} \rightarrow$ Morbidade respiratória segundo a faixa etária e sexo com uma **Baixa Frequência**, em relação ao valor esperado;

IRM 2= $q_{0.25} < RESP \leq q_{0.50} \rightarrow$ Morbidade respiratória segundo a faixa etária e sexo com uma **Frequência Normal**, em relação ao valor esperado;

IRM 3= $q_{0.50} < RESP \leq q_{0.75} \rightarrow$ Morbidade respiratória segundo a faixa etária e sexo com uma **Frequência Acima do Normal**, em relação ao valor esperado;

IRM 4= $RESP \geq q_{0.75} \rightarrow$ Morbidade respiratória segundo a faixa etária e sexo com uma **Frequência de Alerta**, em relação ao valor esperado.

4.2.5.3. Validação do Modelo

4.2.5.3.1. Função de Distribuição Cumulativa Empírica – FDCE

Em teoria da probabilidade, a função distribuição acumulada (fda) descreve completamente a distribuição da probabilidade de uma variável aleatória de valor real X . Para cada número real x , a fda é dada por:

$$F(x) = P(X \leq x) \quad (4.22)$$

Onde o lado direito representa a probabilidade de que a variável X tome um valor inferior ou igual a x . A probabilidade de que X se situe num intervalo (a, b) é deste modo $F(b) - F(a)$ se $a < b$.

Qualquer fda F é monótona crescente e contínua a partir da direita. Para além disso temos $\lim_{x \rightarrow -\infty} F(x) = 0$ e $\lim_{x \rightarrow +\infty} F(x) = 1$.

Se X é uma variável aleatória discreta, então ela obtém os valores x_1, x_2, \dots, x_n com probabilidade p_1, p_2, \dots, p_n , e a fda de X será descontínua nos pontos x_i e constante entre eles.

Se a fda F de X é contínua, então X é uma variável aleatória contínua; se para além disso F absolutamente contínua, então existe uma função Integral Lebesgue $f(x)$ tal que:

$$F(b) - F(a) = P(a \leq X \leq b) = \int_a^b f(x) dx \quad (4.23)$$

Para todos os números reais a e b , a função f é igual à derivada de F (quase em toda a parte), e é chamada de função densidade de probabilidade da distribuição de X .

Em estatística, uma função de distribuição empírica é uma função de distribuição de probabilidade que concentra as probabilidades $1/n$ para cada n números na amostra.

Onde X_1, \dots, X_n é uma variável aleatória com $x_i \in \mathfrak{R}, i = 1, \dots, n \in \mathbb{N}$

A Função de Distribuição Cumulativa Empírica (FDCE) $F_n(x)$ é uma função escada baseada na amostra x_1, \dots, x_n definida como:

$$F_n(x) = \frac{1}{n} \sum_{i=1}^n I(x_i \leq x) \quad (4.24)$$

4.2.5.3.2. Matriz de Confusão

A matriz de confusão é uma ferramenta de visualização usada no aprendizado supervisionado para analisar a performance da classificação da rede neural. A matriz de confusão é uma matriz quadrada onde cada coluna da matriz representa os exemplos em uma classe predita (modelado) gerada pela RN, enquanto cada linha representa os exemplos em uma classe actual (observado ou valor alvo). O benefício de uma matriz de confusão é a facilidade de verificar se o sistema estiver confundindo duas classes (i.e. geralmente estimando erroneamente uma classe com outra).

Uma rede neural que executa uma classificação perfeita teria zeros em todos lugares (linhas e colunas) com excepção das entradas diagonais.

CAPÍTULO V: Resultados e Discussões

Os resultados apresentados neste trabalho serão apresentados segundo a metodologia utilizada no *Capítulo 4*. Em uma primeira análise será feita uma breve discussão das admissões hospitalares por morbidade respiratória, sistemas sinópticos e poluentes durante o período estudado. Em seguida serão apresentados os resultados das morbidades respiratórias segundo a faixa etária e sexo levando em consideração as análises multivariadas, os prognósticos gerados pela rede neuronal comparado com os valores observados através das Funções de Distribuição Cumulativa Empírica (FDCE) normal, assim como a análise de sensibilidade da rede e o Índice de Risco Meteorológico verificado através da matriz de confusão.

5.1. Panorama Geral

Na figura 5.1 (a-d), tem-se a média de doenças respiratórias conforme a faixa etária e o sexo por dias da semana. O maior número de internamentos é observado durante os dias úteis, principalmente na segunda-feira, quando comparados com os finais de semana, assim como evidenciou-se que o sexo masculino apresentava uma maior frequência de afecções do aparelho respiratório. Segundo Braga (1998), o menor número de internamentos nos finais de semana é devido a falta de atendimento médico nos hospitais e, também, a diminuição da frota de transportes públicos e outros problemas socio-econômicos.

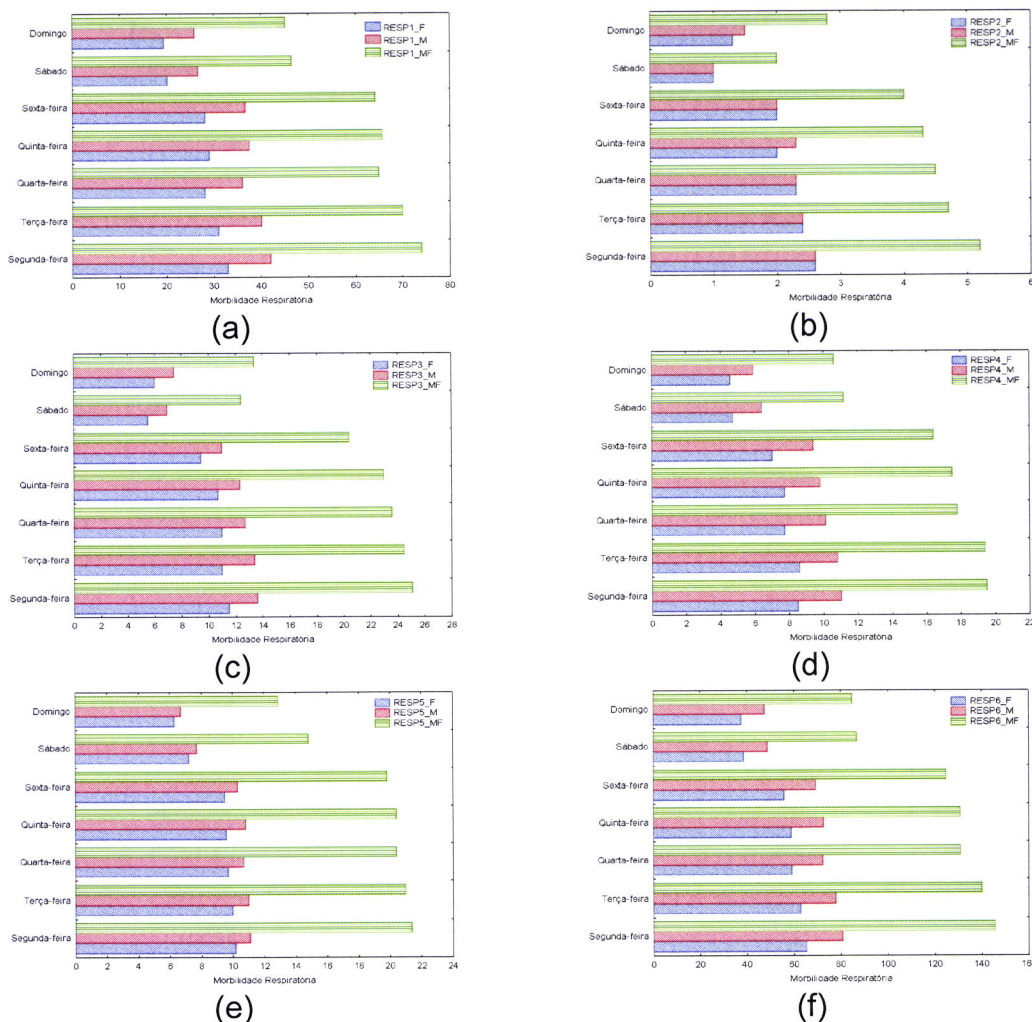


Figura 5.1 – Número de RESP (média) segundo a faixa etária e o sexo por dias da semana para todo o período de estudo.

Na figura 5.2 (a-d), verificou-se que há uma diminuição do número de RESP nos meses de verão e um aumento nas estações de Outono (pico) e Inverno, enquanto que na Primavera os internamentos voltam a diminuir. O pico identificado no Outono está supostamente relacionado ao mês de Maio, que marca a transição entre as duas estações, este acréscimo prolonga-se até o mês de Setembro, que são os meses considerados mais poluídos, segundo a CETESB (2000).

As associações entre doenças respiratórias e a sazonalidade, do ponto de vista meteorológico, provavelmente está relacionada com as incursões de massas de ar frias e secas, no mês de Maio, que favorecem um maior número de internamentos, possivelmente, provocada por uma mudança significativa no

uma frequência de 13% de actuações de sistemas frontais, 13% de sistemas pré frontais e 4% de sistemas pós frontais. Os sistemas frontais são acompanhados por precipitação (F), geralmente estratiformes (leve a moderado); enquanto que os sistemas frontais (pré e pós) podem organizar precipitações convectivas (moderado a forte). Em relação aos poluentes, estes sistemas favorecem a dispersão dos poluentes.

Em relação aos anticiclones foi observada uma frequência de 69% de actuações durante o período, sendo que 38% de actuação foi verificada pela presença da alta subtropical que possui características de uma massa de ar quente e húmida, 30% de actuação verificada pela presença da alta polar que possui características de uma massa de ar fria e pouco húmida e 1% verificada pela presença da alta continental que possui características de uma massa de ar fria e seca.

Deve-se ressaltar que os sistemas sinóticos de alta pressão podem variar significativamente durante o ano, conforme, a sua intensidade de actuação. Estes sistemas geralmente favorecem o aumento dos poluentes.

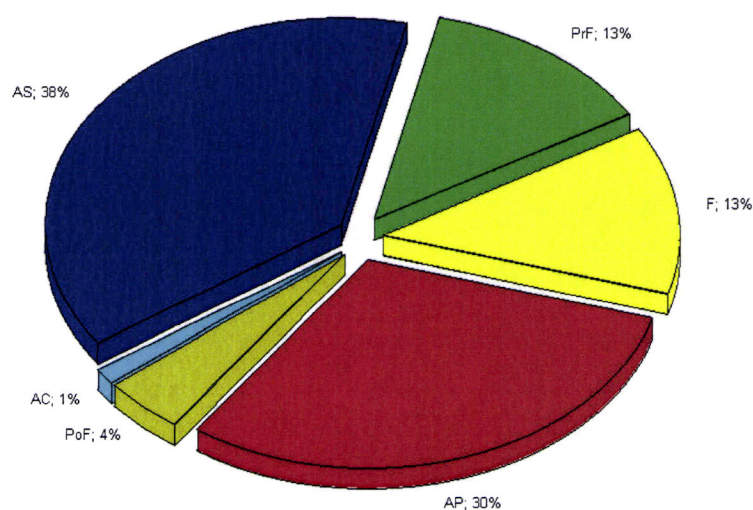


Figura 5.3 – Frequência de sistemas sinóticos para todo o período de estudo.

Em relação ao PM_{10} , figura 5.4a, verifica-se que as concentrações do PM_{10} ultrapassaram em 36 dias, o padrão primário para 24 horas (pp24h) de qualidade

do ar de $150 \mu\text{g}/\text{m}^3$, que não pode ser excedido por mais de uma vez ao ano. Estas ultrapassagens foram verificadas principalmente na estação de Inverno, período que apresenta maiores frequências de inversões térmicas, com pouca ou ausência de precipitação e vento.

Na figura 5.4b é apresentada a concentração do (O_3), onde foi verificada 126 ultrapassagens (23 no Inverno, 12 no Outono, 64 na Primavera e 28 no Verão) deste poluente em relação ao padrão primário de 1 hora (pp1h) de qualidade do ar de $160 \mu\text{g}/\text{m}^3$ (máxima diária) durante todo o período de estudo. Por outro lado, pode-se verificar que o poluente ozono apresenta uma tendência a aumentar a partir do ano de 1999.

A figura 5.4c mostra a concentração do CO que apresentou 104 ultrapassagens em relação ao padrão primário de 8 horas (9 ppm) durante o período. Estas ultrapassagens são verificadas principalmente no Inverno (70) e Outono (29). A partir do ano de 1996 verifica-se que o monóxido de carbono apresenta uma tendência a diminuir; esta diminuição pode está ligada aos monitoramentos da CETESB nas indústrias, programa de rodízios dos veículos, assim como uma frota de veículos mais modernas que apresentam catalizadores para redução de poluentes.

A figura 5.4d apresenta o comportamento do poluente SO_2 onde não foi observada nenhuma ultrapassagem deste poluente durante todo período em relação ao padrão primário de 24 horas ($365 \mu\text{g}/\text{m}^3$).

Conforme metodologia empregada para as medidas de concentrações dos poluentes neste estudo, é importante ressaltar que o SO_2 foi o único poluente que não ultrapassou os padrões de qualidade do ar em todas as 13 estações. Entretanto, os demais poluentes, dependendo da localização da estação, ultrapassaram em pelo menos uma estação.

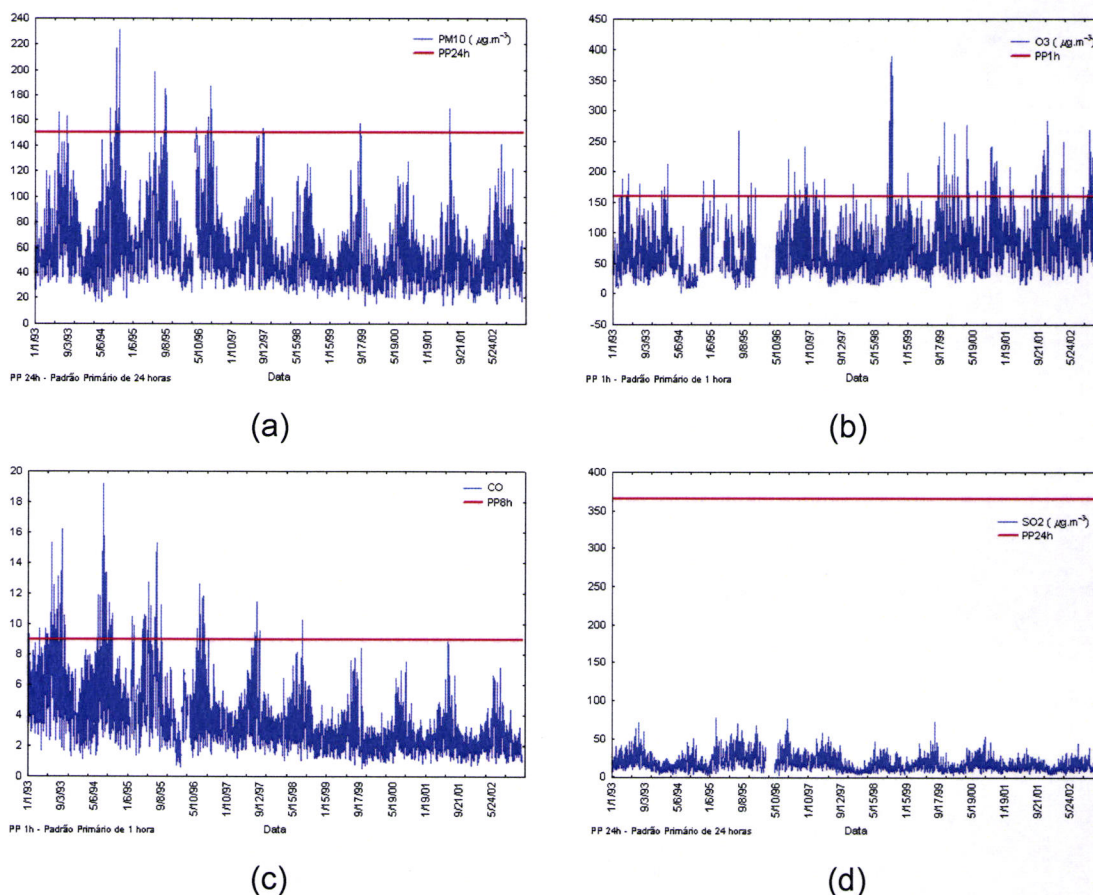


Figura 5.4 – Série Temporal dos Poluentes e Padrões Primários: (a) PM_{10} , (b) O_3 , (c) CO e (d) SO_2

5.2. Resultados para RESP1 (M, F e MF)

5.2.1. Análise em Componentes Principais (ACP) e Análise de “Cluster” (AC)

Para os resultados da ACP e AC, foram introduzidos os dados de morbidade respiratória RESP (M, F e MF) em conjunto com os parâmetros meteorológicos [Tmin (temperatura mínima), Uemin (Humidade específica mínima), Rad (Radiação solar) e vv (velocidade do vento)], concentrações de poluentes [CO, PM_{10} , SO_2 e O_3]. Deve-se ressaltar que durante as análises, optou-se em trabalhar com a humidade específica do ar. Isto se deve ao facto que nos testes realizados, a humidade específica do ar apresentou melhores resultados quando comparada com a humidade relativa, principalmente, em relação à variância explicada total.

Os resultados das análises mostradas nas tabelas a seguir apresentam os valores “em vermelho” dos pesos das variáveis em questão, após a rotação varimax, para os autovalores $\geq 0,70$.

Na tabela 5.1, Observou-se a extracção de 4 componentes após a rotação varimax para RESP1F e RESP1MF, com uma variância acumulada (Var. Acum.) em torno de 80%, enquanto que RESP1M apresentou a extracção de apenas 3 componentes, com uma variância acumulada de 70%.

Para CP1, verificou-se uma variância explicada (Var. Expl.) de aproximadamente 30%, com altos pesos positivos para os poluentes SO₂, MP₁₀ e CO em relação à RESP1F (0.60), RESP1M (0.55) e RESP1MF (0.62), assim como pesos negativos significativos para a velocidade do vento (vv) e pouco significativos para a temperatura mínima (Tmin) e humidade específica mínima (Uemin). Esta componente explica que uma diminuição da velocidade do vento ou ausência de vento, provavelmente, favorece um aumento da concentração dos níveis dos poluentes, e consequentemente, um aumento da morbidade de doenças respiratórias.

Na CP2, observou-se uma variância explicada de 22% com pesos positivos para as variáveis Tmin (0.94) e Uemin em oposição à RESP1F, RESP1F; RESP1M, RESP1MF, SO₂ e MP₁₀ com pesos pouco significativo. Esta componente indica que o aumento de temperatura e humidade específica do ar favorecem o decréscimo das doenças respiratórias.

Na CP3, verifica-se uma variância explicada na ordem de 16%, com pesos positivos para a radiação solar (Rad) e ozono (O₃). Este factor explica a formação fotoquímica do poluente ozono devido a presença da radiação solar, porém, não mostrou nenhuma relação com os internamentos hospitalares.

Na CP4, observa-se uma variância explicada de 11% com pesos positivos para a velocidade do vento (vv) em oposição aos poluentes MP₁₀ e O₃ com pesos significativos em torno de (-33.0 e -0.47), respectivamente. Esta componente pode ser explicada pela dispersão dos poluentes provocado pela duração e intensidade do vento, provavelmente, causado pela aproximação de sistemas pré-frontais ou entradas de brisas na Cidade de São Paulo. Esta condição favorece um pequeno aumento de morbidade respiratória.

Tabela 5.1 – Pesos das componentes principais após a rotação varimax para todo o período de estudo (01/01/1993 a 31/12/2002) – RESP1

Variáveis	CP1	CP2	CP3	CP4
Tmin	-0,062747	0,949134	0,123697	0,079355
Uemin	-0,076840	0,963811	-0,021045	-0,069386
Rad	0,049101	0,020690	0,915024	0,133516
vv	-0,174288	0,014865	0,004023	0,924571
MP ₁₀	0,834181	-0,134760	0,145099	-0,331449
SO ₂	0,770310	-0,270057	0,184085	-0,182900
O ₃	0,028517	0,158309	0,729884	-0,468099
CO	0,845693	0,025375	-0,156711	0,017392
RESP1F	0,603211	-0,149871	-0,006939	0,189670
Var. Expl. (%)	30,62514	21,79101	15,61658	11,18076
Var. Acum. (%)	30,62514	52,41616	68,03274	79,21350
Variáveis	CP1	CP2	CP3	
Tmin	-0,088947	0,950880	0,091472	
Uemin	-0,069642	0,949126	0,039068	
Rad	-0,010887	0,052022	0,733906	
vv	-0,364081	0,077290	-0,425663	
MP ₁₀	0,887054	-0,134179	0,219522	
SO ₂	0,780700	-0,267030	0,193187	
O ₃	0,101636	0,144404	0,861440	
CO	-0,824996	0,035521	-0,202797	
RESP1M	0,549766	-0,235402	-0,155300	
Var. Expl. (%)	30,47266	21,91764	15,81025	
Var. Acum. (%)	30,47266	52,39031	68,20056	
Variáveis	CP1	CP2	CP3	CP4
Tmin	-0,059158	0,948236	0,124584	0,080642
Uemin	-0,076682	0,961333	-0,020879	-0,070369
Rad	0,049544	0,021499	0,917443	0,133111
vv	-0,166699	0,017375	0,009334	0,931607
PM ₁₀	0,837581	-0,137810	0,147771	-0,326389
SO ₂	0,765042	-0,276242	0,187906	-0,184672
O ₃	0,027144	0,160984	0,724553	-0,475086
CO	0,845123	0,017627	-0,149270	0,022985
RESP1MF	0,625620	-0,187599	-0,034181	0,148821
Var. Expl. (%)	30,73416	21,90054	15,81551	11,13890
Var. Acum. (%)	30,73416	52,63470	68,45021	79,58911

Na figura 5.5, através do dendograma de árvore pelo método de Ward e distâncias euclidianas verificou-se a formação de um grupo entre as doenças respiratórias segundo o sexo RESP1 (M, F e MF) e os poluentes CO, MP₁₀ e SO₂, indicando uma boa similaridade de ligação sobre o comportamento da morbidade. Um segundo grupo formado pelas variáveis meteorológicas temperatura mínima (Tmin) e humidade específica mínima (Uemin), que é coerente do ponto vista meteorológico, com uma similaridade mais distante de RESP1 (M, F e MF), assim como a formação de um terceiro grupo composto pela radiação solar (Rad) e o poluente ozono (O₃).

A formação dos grupos descritos pela AC apresentou o mesmo padrão demonstrado nas extracções dos factores através da ACP.

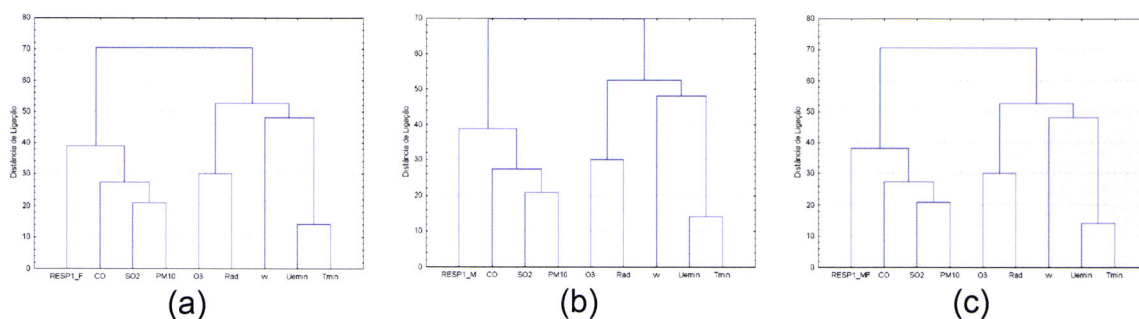


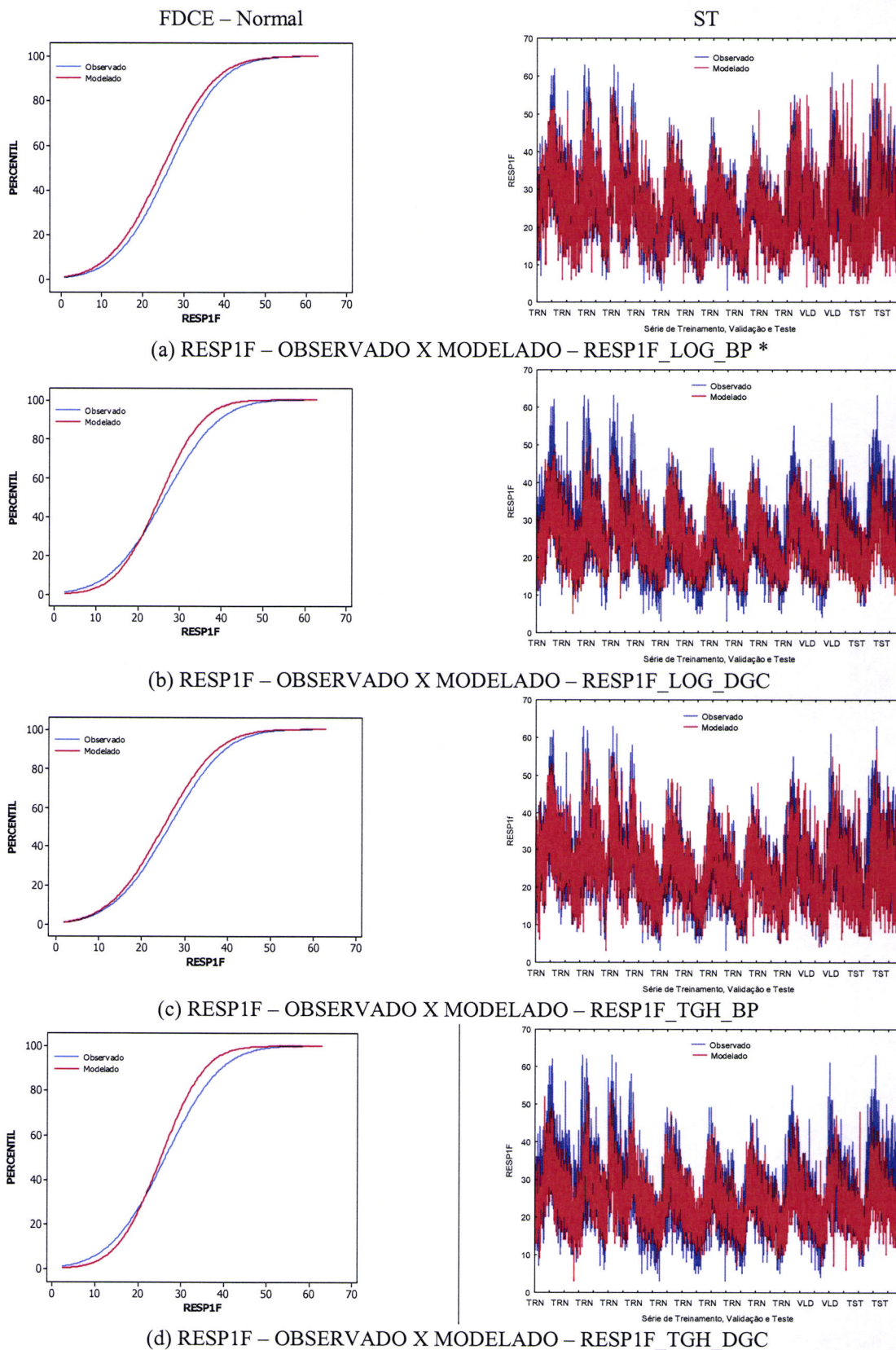
Figura 5.5 – Dendrograma de árvore através do método de Ward e distâncias euclidianas para todo o período de estudo: RESP1

5.2.2. Prognósticos Gerados pela RN – FDCE (Observado X Modelado)

A figura 5.6 (a - d), mostra a Função de Distribuição Cumulativa Empírica (FDCE) e a série temporal entre os valores observados e modelados através das redes neuronais para RESP1F.

Na figura 5.6^a, verificou-se que o melhor desempenho da rede neuronal foi através do algoritmo de treinamento Backpropagation (BP) e função de activação logística (LOG) para RESP1F_LOG_BP. Observa-se através das curvas de ajustes que os valores prognosticados pela rede detectam bem a variabilidade diária e sazonal das doenças respiratórias, porém, tendem a subestimar os valores observados, assim como uma instabilidade nos valores extremos que se apresentaram pouco robustos.

Na Tabela 5.2, verifica-se os parâmetros da rede e estatística descritiva entre os valores observados e modelados durante o aprendizado realizado pela rede. A rede RESP1F_LOG_BP apresentou a melhor resposta com arquitectura [62-130-1], taxa de aprendizagem (0,1), momento (1,75) e peso de $\pm 1,5$. Os parâmetros utilizados para uma melhor aferição da rede mostraram um comportamento semelhante ao observado em termos da média e desvio padrão, enquanto que foi verificado um erro absoluto médio de 5,3.



*Melhor rede treinada

Figura 5.6 – Comparação entre os valores observados de RESP1F e os valores modelados pela RNA através da Função de Distribuição Cumulativa Empírica (FDCE) e Série Temporal (ST).

Tabela 5.2 – Parâmetros da RNA e Estatística para RESP1F

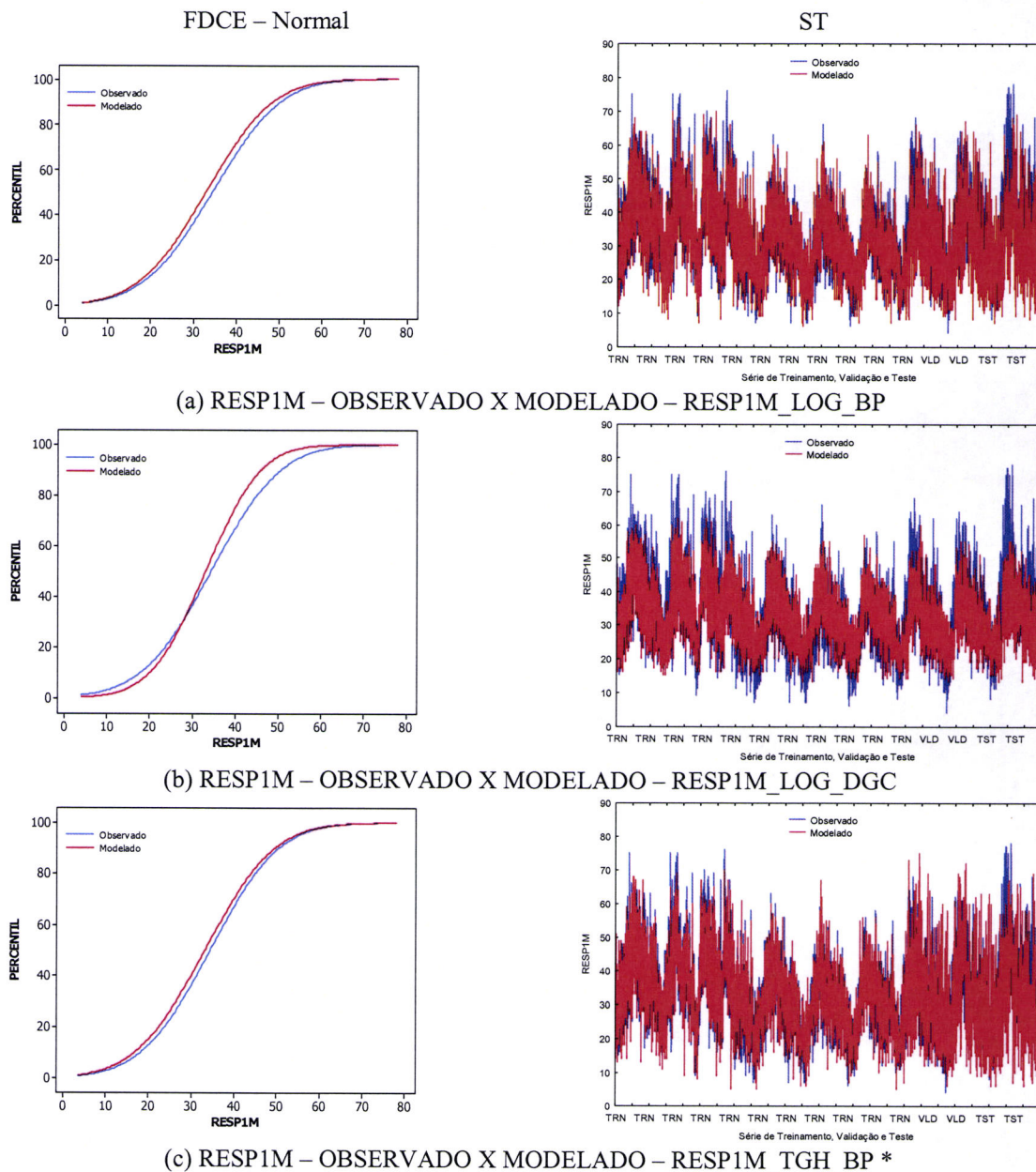
RESP1F LOG BP *					
Parâmetros		Estatística	Observado	Modelado	Erro Absoluto
Arquitectura	[62-101-1]	Média	26,5	24,9	5,3
W	± 1,5	Desv.Pad	10,3	10,3	5,5
TA	0,1	Min	3	4	0,007
Momento	1,75	Max	63	59	37
Iterações	500	R	0,74		
Retreinamentos	3	R ²	0,54		
RESP1F LOG DGC					
Parâmetros		Estatística	Observado	Modelado	Erro Absoluto
Arquitectura	[62-101-1]	Média	26,5	25,3	5,7
W	± 4	Desv.Pad	10,3	8,2	4,6
TA		Min	3	5	0,004
Momento		Max	63	50	29
Iterações	500	R	0,72		
Retreinamentos	3	R ²	0,52		
RESP1F TGH BP					
Parâmetros		Estatística	Observado	Modelado	Erro Absoluto
Arquitectura	[62-130-1]	Média	26,5	25	5,3
W	± 1,5	Desv.Pad	10,3	10	5
TA	0,1	Min	3	3,5	0,002
Momento	1,75	Max	63	57	43
Iterações	500	R	0,74		
Retreinamentos	3	R ²	0,54		
RESP1F TGH DGC					
Parâmetros		Estatística	Observado	Modelado	Erro Absoluto
Arquitectura	[62-130-1]	Média	26,5	25,4	5,6
W	± 0,5	Desv.Pad	10,3	8	4,6
TA		Min	3	3,5	0,0006
Momento		Max	63	55	29,8
Iterações	500	R	0,72		
Retreinamentos	3	R ²	0,52		

*Melhor rede treinada

A figura 5.7 (a - d) mostra a Função de Distribuição Cumulativa Empírica (FDCE) e a série temporal entre os valores observados e modelados através das redes neuronais para RESP1F.

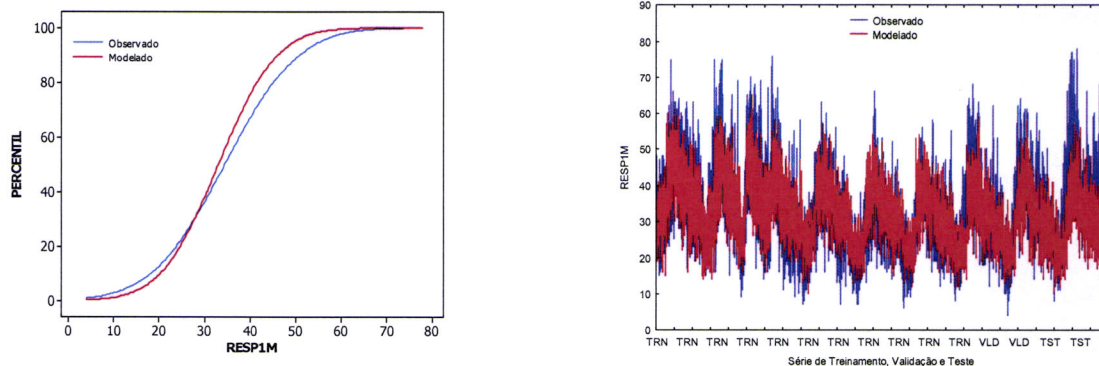
Na figura 5.7c, verificou-se que o melhor desempenho da rede neuronal foi através do algoritmo de treinamento Backpropagation (BP) e função de activação tangente hiperbólica (TGH) para RESP1F_LOG_BP. Observa-se através das curvas de ajustes que os valores prognosticados pela rede detectam bem a variabilidade diária e sazonal das doenças respiratórias, porém, tendem a subestimar os valores observados, assim como uma instabilidade nos valores extremos que se apresentaram pouco robustos, principalmente, na série testada.

Na Tabela 5.3, verifica-se os parâmetros da rede e estatística descritiva entre os valores observados e modelados durante o aprendizado realizado pela rede. A rede RESP1M_TGH_BP apresentou a melhor resposta com arquitectura [62-144-1], taxa de aprendizagem (0,1), momento (1,75) e peso de $\pm 0,5$. Os parâmetros utilizados para uma melhor aferição da rede mostraram um comportamento semelhante ao observado em termos da média e desvio padrão, enquanto que foi verificado um erro absoluto médio de 6,5.



*Melhor rede treinada

Figura 5.7 – Comparação entre os valores observados de RESP1M e os valores modelados pela RNA através da Função de Distribuição Cumulativa Empírica (FDCE) e Série Temporal (ST).



(d) RESP1M – OBSERVADO X MODELADO – RESP1M_TGH_DGC

Figura 5.7 – cont.

Tabela 5.3 – Parâmetros da RNA e Estatística para RESP1M

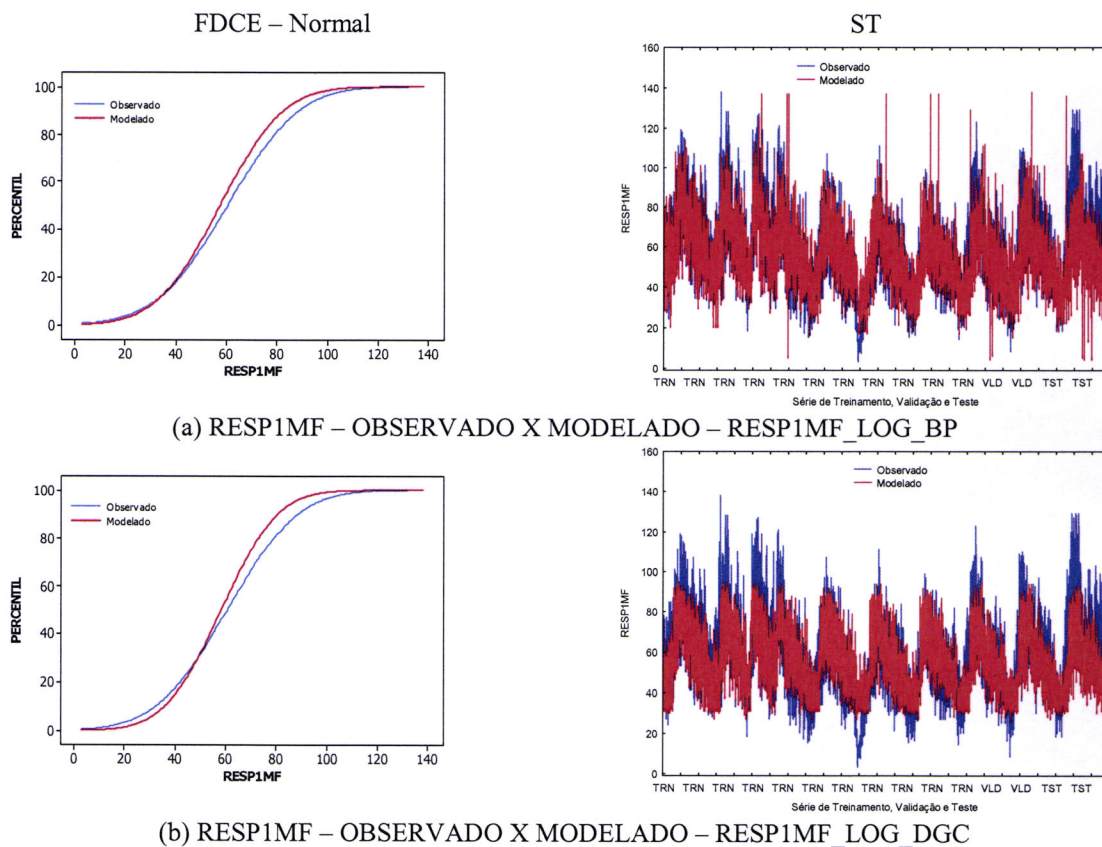
RESP1M LOG BP					
Parâmetros		Estatística	Observado	Modelado	Erro Absoluto
Arquitetura	[62-83-1]	Média	34,5	33	6,7
W	± 1,5	Desv.Pad	12,6	12,3	6,3
TA	0,1	Min	4	6	0,001
Momento	1,75	Max	78	70	49
Iterações	500	R	0,74		
Retreinamentos	3	R ²	0,54		
RESP1M LOG DGC					
Parâmetros		Estatística	Observado	Modelado	Erro Absoluto
Arquitetura	[62-83-1]	Média	34,5	33,2	6,5
W	± 3,5	Desv.Pad	12,6	10	5,5
TA		Min	4	12,7	0,001
Momento		Max	78	62,3	40,4
Iterações	500	R	0,74		
Retreinamentos	3	R ²	0,55		
RESP1M TGH BP *					
Parâmetros		Estatística	Observado	Modelado	Erro Absoluto
Arquitetura	[62-144-1]	Média	34,5	33,3	6,5
W	± 0,5	Desv.Pad	12,6	12,8	6,9
TA	0,1	Min	4	4,6	0,005
Momento	1,75	Max	78	74,8	47,1
Iterações	500	R	0,72		
Retreinamentos	3	R ²	0,52		
RESP1M TGH DGC					
Parâmetros		Estatística	Observado	Modelado	Erro Absoluto
Arquitetura	[62-144-1]	Média	34,5	33,2	6,6
W	± 0,3	Desv.Pad	12,6	9,9	5,6
TA		Min	4	9,6	0,002
Momento		Max	78	64,6	41,2
Iterações	500	R	0,74		
Retreinamentos	3	R ²	0,55		

*Melhor rede treinada

A figura 5.8 (a - d) mostra a Função de Distribuição Cumulativa Empírica (FDCE) e a série temporal entre os valores observados e modelados através das redes neurais para RESP1MF.

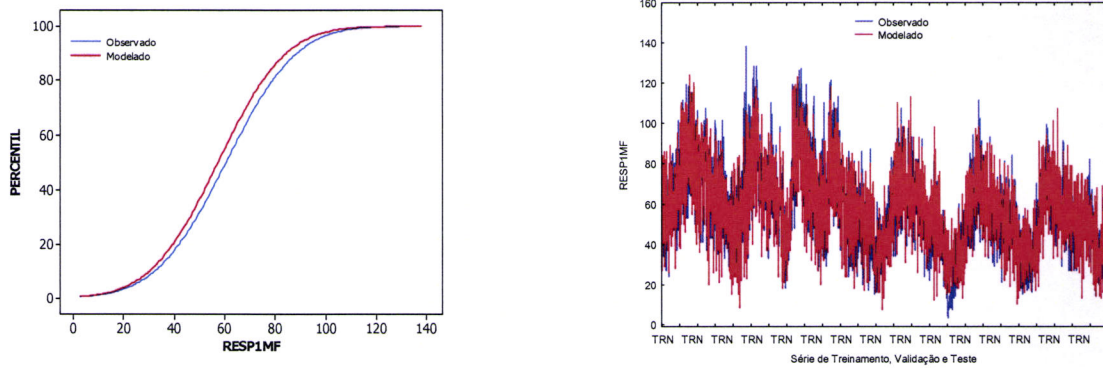
Na figura 5.8c, verificou-se que o melhor desempenho da rede neuronal foi através do algoritmo de treinamento Backpropagation (BP) e função de activação tangente hiperbólica (TGH) para RESP1MF_TGH_BP. Observa-se através das curvas de ajustes que os valores prognosticados pela rede detectam bem a variabilidade diária e sazonal das doenças respiratórias, porém, tendem a subestimar os valores observados, assim como uma instabilidade nos valores extremos que se apresentaram pouco robustos.

Na Tabela 5.4, verifica-se os parâmetros da rede e estatística descritiva entre os valores observados e modelados durante o aprendizado realizado pela rede. A rede RESP1MF_LOG_BP apresentou o melhor desempenho com arquitectura [62-117-1], taxa de aprendizagem (0,1), momento (1,75) e peso de $\pm 0,45$. Os parâmetros utilizados para uma melhor calibração da rede apresentaram um comportamento semelhante ao observado em termos da média e desvio padrão, enquanto que foi verificado um erro absoluto médio de 8,5.

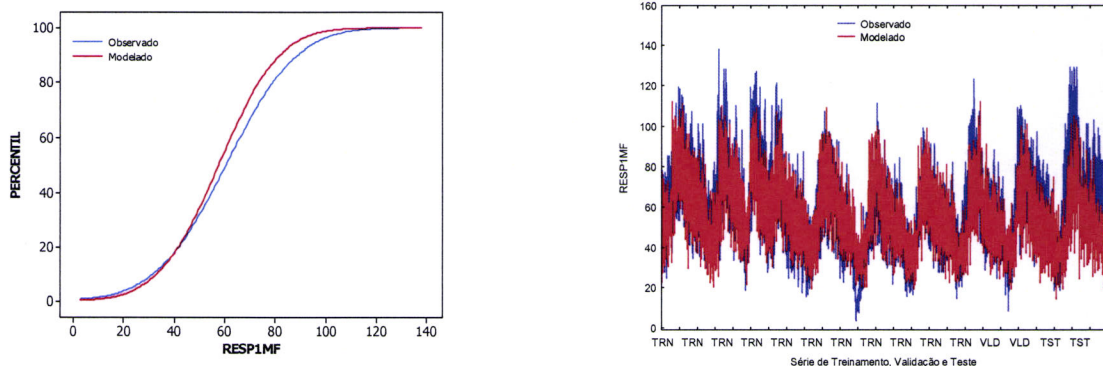


*Melhor rede treinada

Figura 5.8 – Comparação entre os valores observados de RESP1MF e os valores modelados pela RNA através da Função de Distribuição Cumulativa Empírica (FDCE) e Série Temporal (ST).



(c) RESP1MF – OBSERVADO X MODELADO – RESP1MF_TGH_BP *



(d) RESP1MF – OBSERVADO X MODELADO – RESP1MF_TGH_DGC

Figura 5.8 – cont.

Tabela 5.4 – Parâmetros da RNA e Estatística para RESP1MF

RESP1MF LOG BP					
Parâmetros		Estatística	Observado	Modelado	Erro Absoluto
Arquitetura	[62-18-1]	Média	60,8	57,6	11,5
W	± 6	Desv.Pad	22,1	19,5	10,3
TA	0,1	Mín	3	3,6	0,004
Momento	1,75	Max	138	137,6	77,7
Iterações	500	R		0,74	
Retreinamentos	3	R ²		0,55	
RESP1MF LOG DGC					
Parâmetros		Estatística	Observado	Modelado	Erro Absoluto
Arquitetura	[62-18-1]	Média	60,8	58,3	11,4
W	± 2	Desv.Pad	22,1	17,5	9,2
TA		Mín	3	26,4	0,003
Momento		Max	138	95,2	61,3
Iterações	500	R		0,76	
Retreinamentos	3	R ²		0,57	
RESP1MF TGH BP *					
Parâmetros		Estatística	Observado	Modelado	Erro Absoluto
Arquitetura	[62-117-1]	Média	60,8	59,1	8,5
W	± 0,45	Desv.Pad	22,1	21,6	6,9
TA	0,1	Mín	3	6,9	0,003
Momento	1,75	Max	138	124	47
Iterações	500	R		0,88	
Retreinamentos	3	R ²		0,77	
RESP1MF TGH DGC					
Parâmetros		Estatística	Observado	Modelado	Erro Absoluto

Arquitectura	[62-117-1]	Média	60,8	57,8	11,1
W	$\pm 0,6$	Desv.Pad	22,1	18,9	9,3
TA		Min	3	13,7	0,005
Momento		Max	138	112	64,5
Iterações	500	R	0,77		
Retreinamentos	3	R ²	0,60		

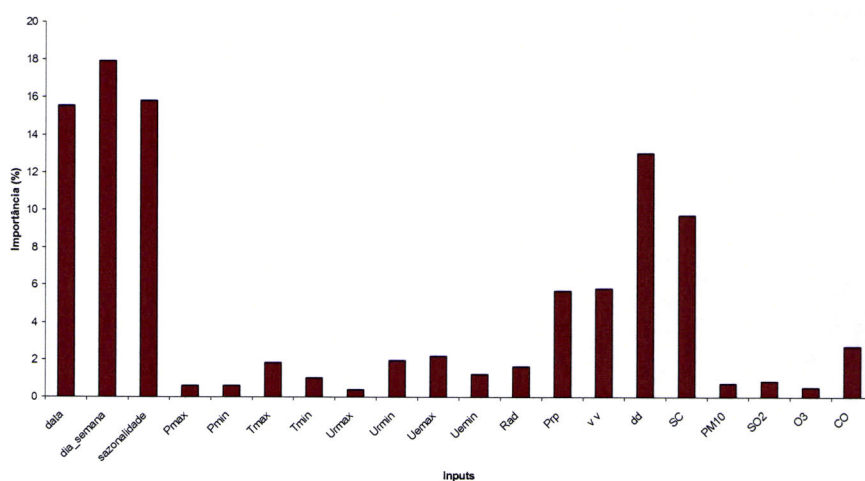
*Melhor rede treinada

Tabela 5.4 – cont.

5.2.3. Análise de Sensibilidade da RN – Importância dos “inputs”

A figura 5.9 (a - c) mostra a análise de sensibilidade dos “inputs” para os melhores desempenho da rede descritos para a morbilidade de doenças respiratórias em crianças até 12 anos segundo a faixa etária e sexo.

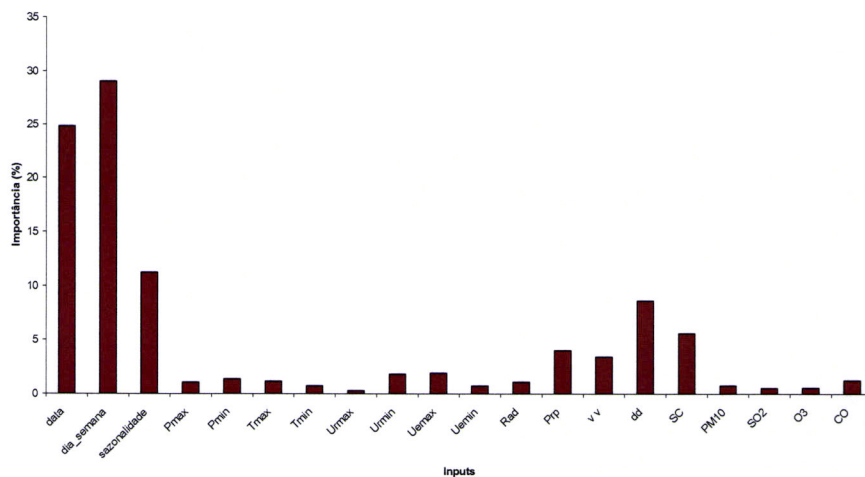
De um modo geral, observou-se que as variáveis de periodicidade apresentaram 60% de importância, principalmente, a variabilidade semanal (dias_semana – 25%), enquanto que 38% está relacionada com as variáveis meteorológicas, principalmente a direcção do vento (dd – 10 %) e condição sinóptica (SC – 8%). As variáveis poluentes mostraram uma importância em torno de 2%.



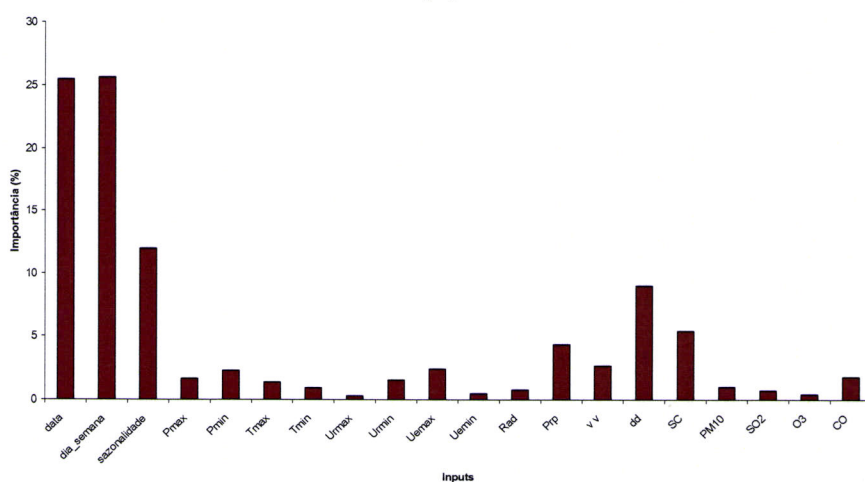
(a)

Figura 5.9 – Análise de Sensibilidade: (a) RESP1F_LOG_BP – [62-130-1]; (b) RESP1M_TGH_BP – [62-144-1] e (c) RESP1MF_TGH_BP – [62-117-1]





(b)



(c)

Figura 5.9 – cont.

5.2.4. Índice de Risco Meteorotrópico (IRM) – Matriz de Confusão

As tabelas 5.5, 5.6 e 5.7 mostram através da matriz de confusão o grau de acerto (valores em azul) do Índice de Risco Meteorotrópico (IRM) entre os valores preditos pelas melhores redes treinadas e os valores observados, durante a fase de aprendizado.

Os valores prognosticados pelas redes (RESP1F_LOG_BP – [62-130-1] e RESP1M_TGH_BP – [62-144-1]) apresentaram um grau de acerto em torno de 70% sobre o valor observado, enquanto que para RESP1MF_TGH_BP – [62-117-1] o grau de acerto foi em torno de 77%.

O IRM 2 ($q_{0.25} < \text{RESP1} \leq q_{0.50}$), faixa com maiores ocorrências de morbidades para RESP1, apresentou uma **Frequência Normal** em relação ao valor esperado pelas afecções respiratórias, enquanto que a classe de **Baixa Frequência** (IRM1) e **Frequência Acima do Normal** (IRM3) tenderam a sobrestimar e subestimar os valores observados, respectivamente. Para o IRM4 não foi prognosticado nenhuma **Frequência de Alerta**. A informação da matriz de confusão corrobora com a explicação descrita nas análises de Funções de Distribuição Cumulativa Empírica (FDCE).

Tabela 5.5 – Matriz de Confusão para RESP1F_LOG_BP – [62-130-1].

MODELADO	OBSERVADO			
	IRM 1	IRM 2	IRM 3	IRM 4
	$\text{RESP1F} \leq q_{0.25}$	$q_{0.25} < \text{RESP1F} \leq q_{0.50}$	$q_{0.50} < \text{RESP1F} \leq q_{0.75}$	$\text{RESP1F} \geq q_{0.75}$
	≥ 19	[20....33]	[34....63]	≤ 63
≥ 19	556	311	1	0
[20....33]	216	1457	128	0
[34....62]	1	240	225	0
≤ 63	0	14	26	0

Tabela 5.6 – Matriz de Confusão para RESP1M_TGH_BP – [62-144-1].

MODELADO	OBSERVADO			
	IRM 1	IRM 2	IRM 3	IRM 4
	$RESP1M \leq q_{0.25}$	$q_{0.25} < RESP1M \leq q_{0.50}$	$q_{0.50} < RESP1M \leq q_{0.75}$	$RESP1M \geq q_{0.75}$
	≥ 26	[27....33]	[34....77]	≤ 78
≥ 26	579	287	2	0
[27....33]	206	1454	135	0
[34....77]	1	243	232	0
≤ 78	0	14	26	0

Tabela 5.7 – Matriz de Confusão Para RESP1MF_TGH_BP – [62-117-1].

MODELADO	OBSERVADO			
	IRM 1	IRM 2	IRM 3	IRM 4
	$RESP1MF \leq q_{0.25}$	$q_{0.25} < RESP1MF \leq q_{0.50}$	$q_{0.50} < RESP1MF \leq q_{0.75}$	$RESP1MF \geq q_{0.75}$
	≥ 43	[44....57]	[58....137]	≤ 138
≥ 43	572	87	5	0
[44....57]	311	1548	126	0
[58....137]	23	146	341	0
≤ 138	0	2	15	0

5.3 – Resultados para RESP2 (M, F e MF)

5.3.1. Análise em Componentes Principais (ACP) e Análise de “Cluster”

Na tabela 5.8, Observou-se a extracção de 4 componentes após a rotação varimax para RESP2M e RESP2MF, com uma variância acumulada (Var. Acum.) em torno de 80%, enquanto que RESP2F apresentou a extracção de apenas 3 componentes, com uma variância acumulada de 70%.

Para CP1, verificou-se uma variância explicada (Var. Expl.) de aproximadamente 30%, com pesos positivos para os poluentes SO₂, MP₁₀ e CO, em relação à RESP2F (0.35), RESP2M (0.30) e RESP2MF (0.45), assim como pesos negativos significativos para a velocidade do vento (vv) e pouco significativos para a temperatura mínima (Tmin) e humidade específica mínima (Uemin). Este factor tem o mesmo comportamento verificado na extracção da CP1 para crianças com morbilidades respiratórias até 13 anos segundo o sexo – RESP1 (M, F e MF).

Na CP2, observou-se uma variância explicada de 22% com pesos positivos para as variáveis Tmin (0.94) e Uemin em oposição à RESP2F, RESP2M, RESP2MF, SO₂ e MP₁₀ com pesos pouco significativo. Esta componente indica que o aumento de temperatura e humidade específica do ar favorecem o decréscimo das doenças respiratórias, conforme CP2 de RESP1 (M, F e MF).

Na CP3, verifica-se uma variância explicada na ordem de 16%, com pesos positivos para a radiação solar (Rad) e ozono (O₃). Como verificado para RESP1 (M, F e MF), esta componente explicou a formação fotoquímica do poluente ozono devido a presença da radiação solar, porém, não foi observado nenhuma relação com os internamentos hospitalares.

Na CP4, observa-se uma variância explicada de 11% com pesos positivos para a velocidade do vento (vv) em oposição aos poluentes MP₁₀ e O₃ com pesos significativos. Este factor pode ser explicado pela dispersão dos poluentes provocado pela duração e intensidade do vento, conforme denotado pela extracção da CP4 na RESP1 (M, F e MF).

Tabela 5.8 – Pesos das componentes principais após a rotação varimax para todo o período de estudo (01/01/1993 a 31/12/2002) – RESP2

Variáveis	CP1	CP2	CP3	CP4
Tmin	-0,097820	0,952402	0,092200	
Uemin	-0,069049	0,959804	0,031203	
Rad	0,010723	0,050951	0,725247	
vv	-0,408708	0,043383	-0,389223	
PM10	0,899213	-0,115862	0,194430	
SO2	0,795846	-0,250008	0,172028	
O3	0,125811	0,148675	0,848289	
CO	0,824564	0,053536	-0,219013	
RESP2_F	0,352029	-0,120029	-0,331820	
Var. Expl. (%)	29,73200	21,55391	15,63855	
Var. Acum. (%)	29,73200	51,28591	66,92446	
Variáveis	CP1	CP2	CP3	CP4
Tmin	-0,091807	0,949787	0,117837	0,056520
Uemin	-0,068819	0,969229	-0,013469	-0,051003
Rad	0,028460	0,020260	0,901110	0,142836
vv	-0,354084	-0,002497	-0,051634	0,752652
PM10	0,890792	-0,109610	0,143117	-0,191838
SO2	0,800621	-0,249645	0,175054	-0,070652
O3	0,094575	0,153425	0,753233	-0,416491
CO	0,842818	0,053354	-0,181347	0,108008
RESP2_M	0,310861	-0,049449	0,033622	0,618401
Var. Expl. (%)	29,58805	21,55537	15,30695	11,52879
Var. Acum. (%)	29,58805	51,14341	66,45036	77,97915
Variáveis	CP1	CP2	CP3	CP4
Tmin	-0,081213	0,949035	0,126936	0,063654
Uemin	-0,074160	0,967650	-0,022712	-0,062658
Rad	0,053076	0,027767	0,919314	0,129888
vv	-0,256983	0,002733	0,028627	0,860004
PM10	0,862181	-0,118520	0,142508	-0,284060
SO2	0,783847	-0,257650	0,184843	-0,151574
O3	0,051276	-0,157048	0,715274	-0,472860
CO	0,844679	0,040742	-0,147943	0,048898
RESP2_MF	0,459367	-0,088561	-0,112064	0,430052
Var. Expl. (%)	29,80915	21,55613	15,83864	11,24522
Var. Acum. (%)	29,80915	51,36528	67,20392	78,44914

A figura 5.10, mostra o dendograma de árvore através do método de Ward e distâncias euclidianas, onde foi verificado a formação de um grupo conectado apenas pelos poluentes CO, MP₁₀ e SO₂. Por outro lado, foi observado a formação de um segundo grupo composto pela morbidade de doenças respiratórias em adolescentes [RESP2 (M, F e MF)] e a velocidade do vento (vv). Um terceiro grupo formado pelas variáveis meteorológicas temperatura mínima (Tmin) e umidade específica mínima (Uemin), assim como a formação de um quarto grupo composto pela radiação solar (Rad) e o poluente ozônio (O₃), conforme observado em RESP1 (M, F e MF).

A formação dos grupos descritos pela AC apresentou o mesmo padrão demonstrado nas extracções dos factores através da ACP, com excepção ao segundo grupo, onde se verificou o deslocamento de RESP2 (M, F e MF) em relação aos poluentes quando comparado com RESP1 (M, F e MF)

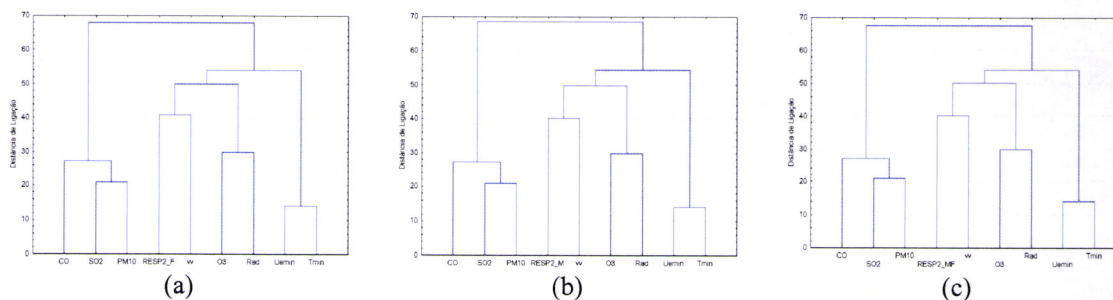


Figura 5.10 – Dendrograma de árvore através do método de Ward e distâncias euclidianas para todo o período de estudo: RESP2

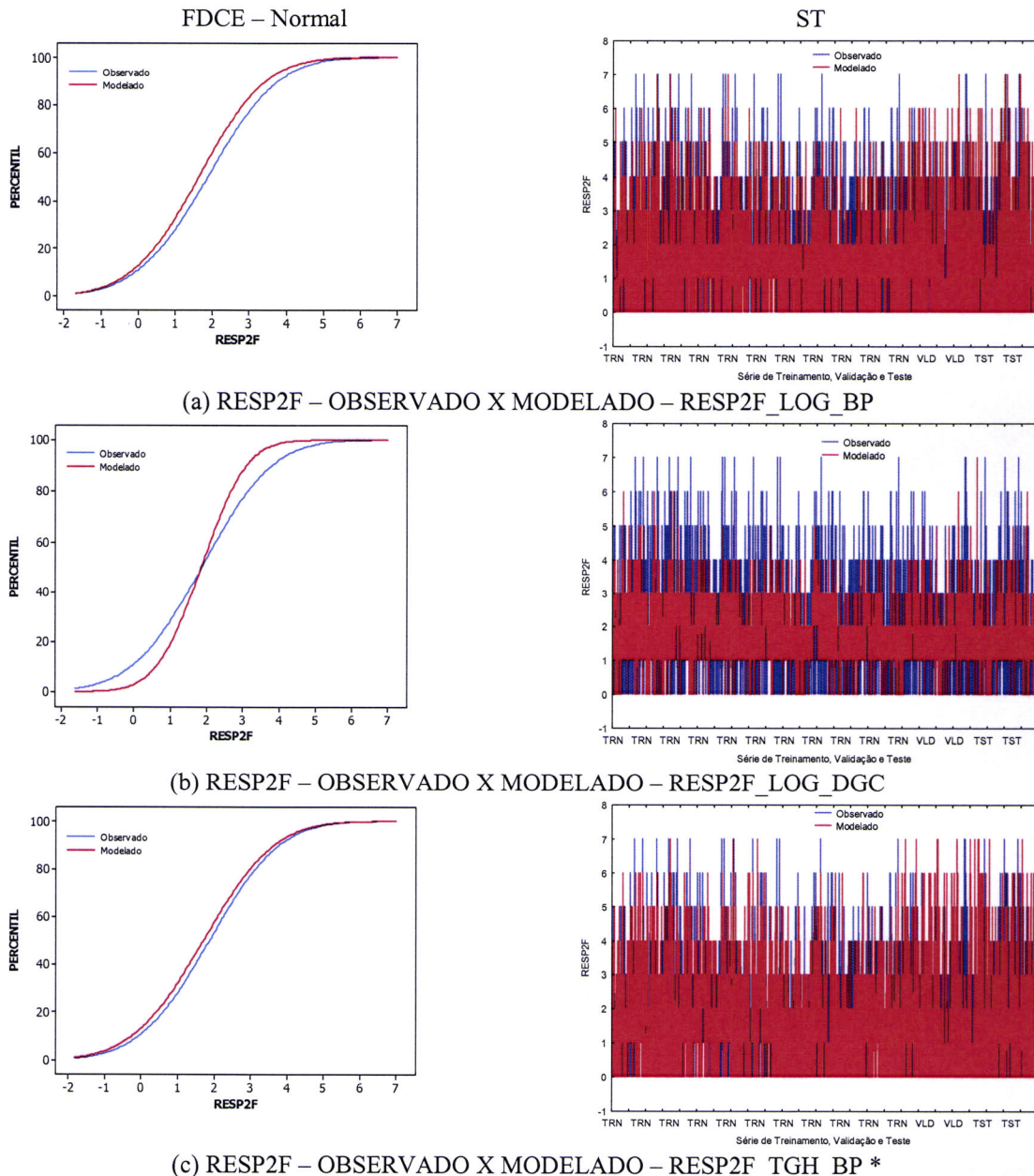
5.3.2. Prognósticos Gerados pela RN – FDCE (Observado X Modelado)

A figura 5.11 (a - d) mostra a Função de Distribuição Cumulativa Empírica (FDCE) e a série temporal entre os valores observados e modelados através das redes neuronais para RESP2F.

Na figura 5.11c, verificou-se que o melhor desempenho da rede neuronal foi através do algoritmo de treinamento Backpropagation (BP) e função de activação tangente hiperbólica (TGH) para RESP2F_TGH_BP. Observa-se através das curvas de ajustes que os valores prognosticados pela rede detectam bem a variabilidade diária e sazonal das doenças respiratórias, porém, tendem a subestimar os valores observados durante a fase de treinamento e sobrestimam na fase testada. A inclusão de mais uma camada escondida poderia melhorar a performance do aprendizado durante a fase de treinamento e, conseqüentemente, tornariam mais robustos os prognósticos gerados pela rede sobre as admissões hospitalares em relação a população adolescente do sexo feminino (14 a 19 anos).

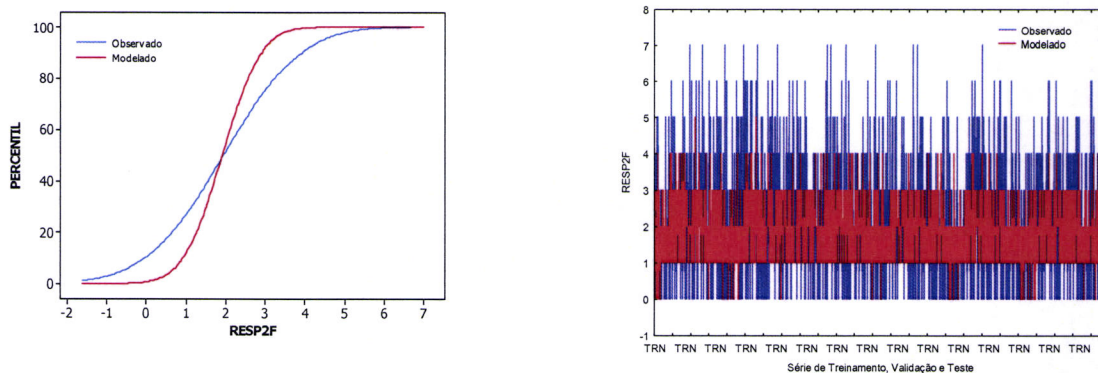
Na Tabela 5.9, verifica-se os parâmetros da rede e estatística descritiva entre os valores observados e modelados durante o aprendizado realizado pela

rede. A rede RESP2F_TGH_BP apresentou o melhor desempenho com arquitetura [62-32-1], taxa de aprendizagem (0,1), momento (1,75) e peso de $\pm 0,7$. Os parâmetros utilizados para uma melhor calibração da rede apresentaram um comportamento semelhante ao observado em termos da média e desvio padrão, enquanto que foi verificado um erro absoluto médio de 0,9.



*Melhor rede treinada

Figura 5.11 – Comparação entre os valores observados de RESP2F e os valores modelados pela RNA através da Função de Distribuição Cumulativa Empírica (FDCE) e Série Temporal (ST). RESP2F



(d) RESP2F – OBSERVADO X MODELADO – RESP2F_TGH_DGC

Figura 5.11 – cont.

Tabela 5.9 – Parâmetros da RNA e Estatística para RESP2F

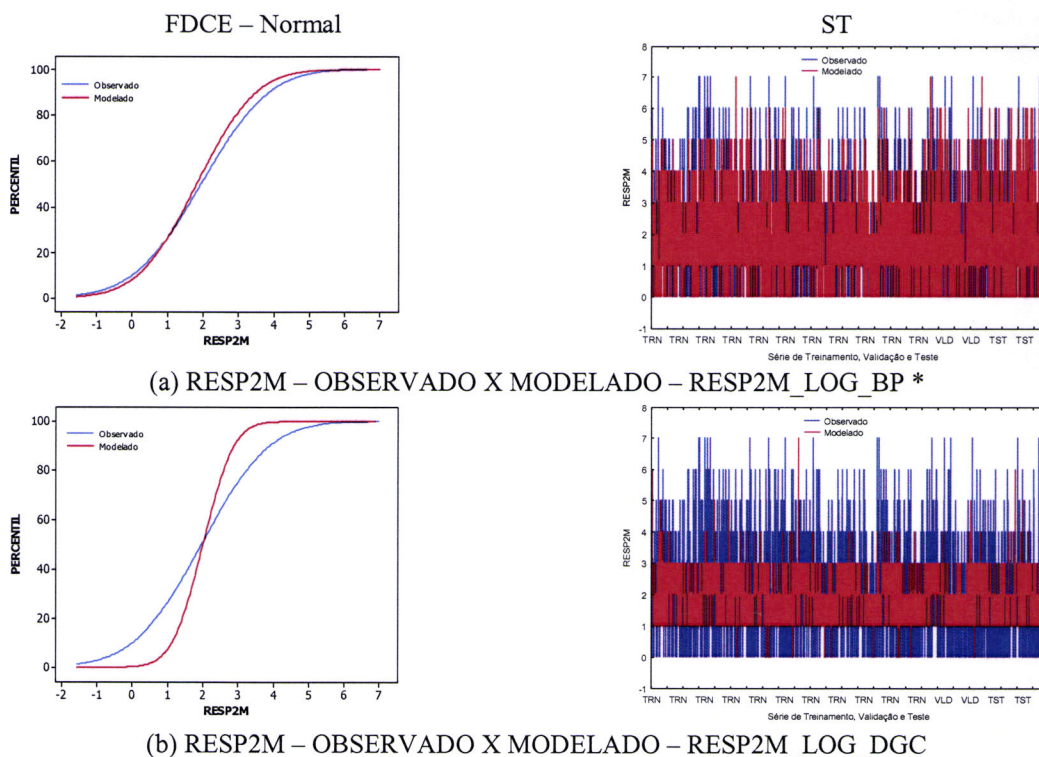
RESP2F LOG BP					
Parâmetros		Estatística	Observado	Modelado	Erro Absoluto
Arquitetura	[62-101-1]	Média	1,9	1,6	1
W	± 2	Desv.Pad	1,5	1,4	1
TA	0,9	Min	0	0,001	0,0001
Momento	1,75	Max	7	6,8	6,6
Iterações	500	R	0,54		
Retreinamentos	3	R ²	0,29		
RESP2F LOG DGC					
Parâmetros		Estatística	Observado	Modelado	Erro Absoluto
Arquitetura	[62-101-1]	Média	1,9	1,9	1
W	± 5	Desv.Pad	1,5	1	0,85
TA		Min	0	0,02	0,0001
Momento		Max	7	6,7	6,9
Iterações	500	R	0,47		
Retreinamentos	3	R ²	0,22		
RESP2F TGH BP *					
Parâmetros		Estatística	Observado	Modelado	Erro Absoluto
Arquitetura	[62-108-1]	Média	1,9	1,7	0,9
W	± 0,7	Desv.Pad	1,5	1,5	1
TA	0,1	Min	0	0,0002	0,0002
Momento	1,75	Max	7	6,9	6,9
Iterações	500	R	0,56		
Retreinamentos	3	R ²	0,31		
RESP2F TGH DGC					
Parâmetros		Estatística	Observado	Modelado	Erro Absoluto
Arquitetura	[62-108-1]	Média	1,9	3,9	1,6
W	± 1,5	Desv.Pad	1,5	1,3	1,3
TA		Min	0	0,73	0,0007
Momento		Max	7	8,3	8,3
Iterações	500	R	0,49		
Retreinamentos	3	R ²	0,24		

*Melhor rede treinada.

A figura 5.12 (a - d) mostra a Função de Distribuição Cumulativa Empírica (FDCE) e a série temporal entre os valores observados e modelados através das redes neuronais para RESP2M.

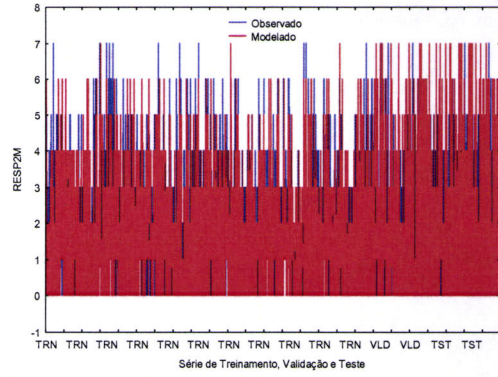
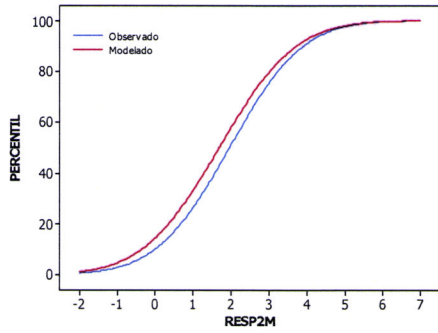
Na figura 5.12^a, verificou-se que o melhor desempenho da rede neuronal foi através do algoritmo de treinamento Backpropagation (BP) e função de activação logística (LOG) para RESP2M_LOG_BP. Observa-se através das curvas de ajustes que os valores prognosticados pela rede detectam bem a variabilidade diária e sazonal das doenças respiratórias, porém, tendem a subestimar os valores observados durante a fase de treinamento e sobrestimam na fase testada, conforme verificado em RESP2F_TGH_BP. A inclusão de mais uma camada escondida poderia melhorar a performance do aprendizado durante a fase de treinamento.

Na Tabela 5.10, verifica-se os parâmetros da rede e estatística descritiva entre os valores observados e modelados durante o aprendizado realizado pela rede. A rede RESP2M_LOG_BP apresentou o melhor desempenho com arquitectura [62-32-1], taxa de aprendizagem (0,1), momento (1,75) e peso de $\pm 3,8$. Os parâmetros utilizados para uma melhor calibração da rede apresentaram um comportamento semelhante ao observado em termos da média e desvio padrão, enquanto que foi verificado um erro absoluto médio de 1.

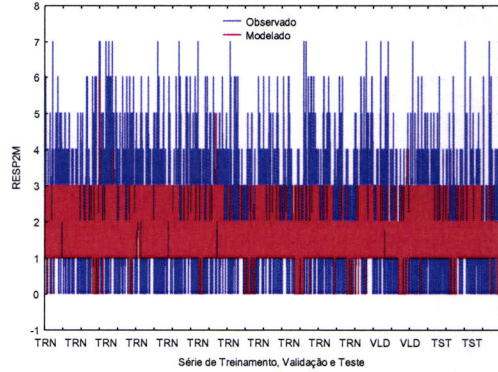
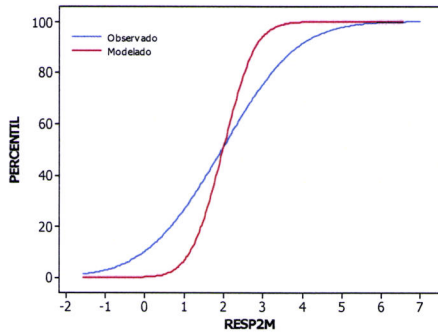


*Melhor rede treinada

Figura 5.12 – Comparação entre os valores observados de RESP2M e os valores modelados pela RNA através da Função de Distribuição Cumulativa Empírica (FDCE) e Série Temporal (ST).



(c) RESP2M – OBSERVADO X MODELADO – RESP2M_TGH_BP



(d) RESP2M – OBSERVADO X MODELADO – RESP2M_TGH_DGC

Figura 5.12 – cont.

Tabela 5.10 – Parâmetros da RNA e Estatística para RESP2M

RESP2M LOG BP *					
Parâmetros		Estatística	Observado	Modelado	Erro Absoluto
Arquitetura	[62-32-1]	Média	1,9	1,8	1
W	$\pm 3,8$	Desv.Pad	1,5	1,3	0,9
TA	0,1	Min	0	0,007	0,0001
Momento	1,75	Max	7	6,6	6,2
Iterações	500	R		0,53	
Retreinamentos	3	R ²		0,28	
RESP2M LOG DGC					
Parâmetros		Estatística	Observado	Modelado	Erro Absoluto
Arquitetura	[62-32-1]	Média	1,9	2	1,1
W	± 5	Desv.Pad	1,5	0,6	0,9
TA		Min	0	0,2	0,00004
Momento		Max	7	6,6	5,2
Iterações	500	R		0,38	
Retreinamentos	3	R ²		0,14	
RESP2M TGH BP					
Parâmetros		Estatística	Observado	Modelado	Erro Absoluto
Arquitetura	[62-91-1]	Média	1,9	2	1,7
W	$\pm 0,9$	Desv.Pad	1,5	1,5	1,6
TA	0,1	Min	0	0	0,0002
Momento	1,75	Max	7	7	6,9
Iterações	500	R		0,51	
Retreinamentos	3	R ²		0,26	
RESP2M TGH DGC					
Parâmetros		Estatística	Observado	Modelado	Erro Absoluto
Arquitetura	[62-91-1]	Média	1,9	2	1,1
W	$\pm 0,55$	Desv.Pad	1,5	0,7	0,8
TA		Min	0	0,1	0

Momento		Max	7	6	5
Iterações	500	R	0,36		
Retreinamentos	3	R ²	0,12		

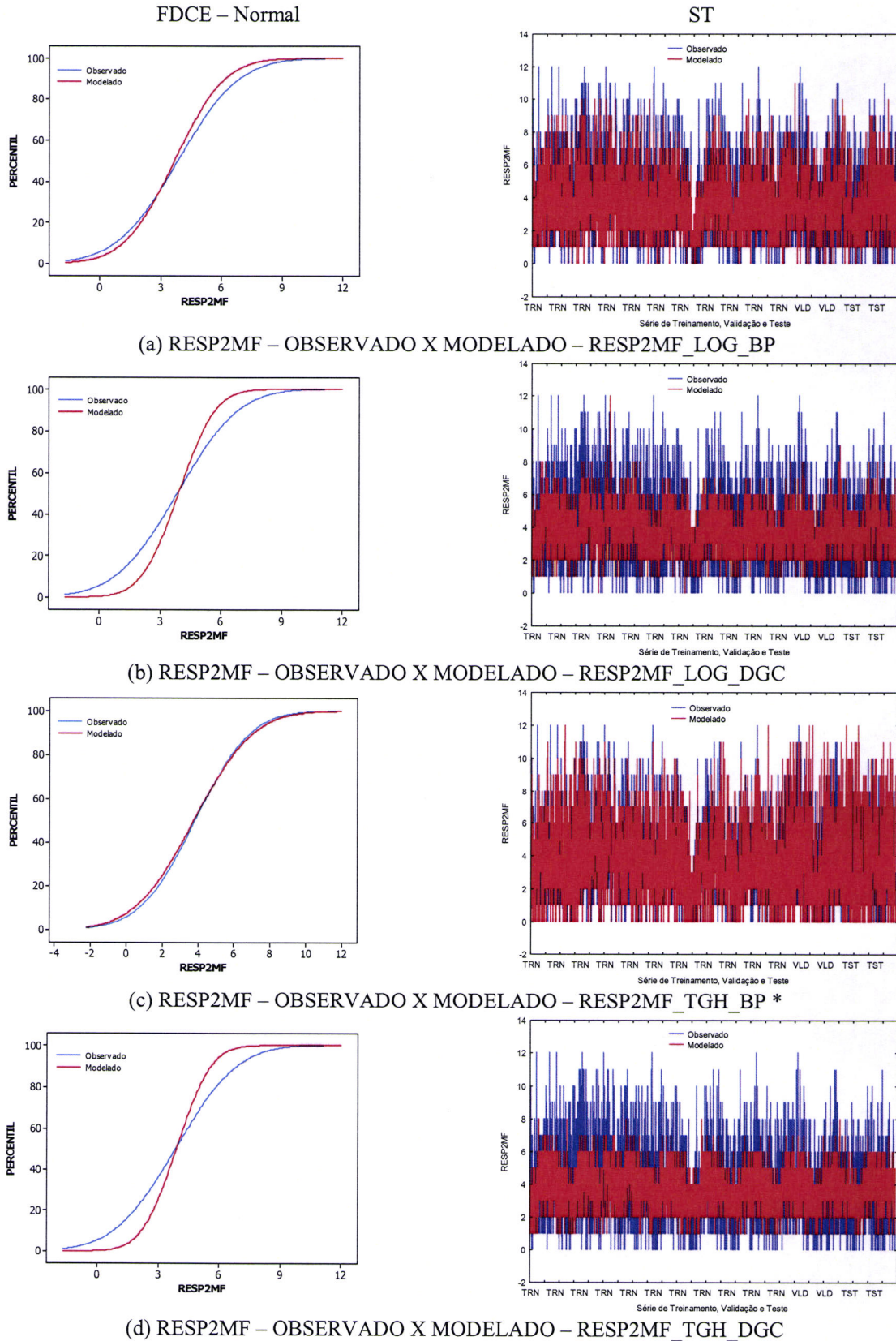
*Melhor rede treinada

Tabela 5.10 – cont.

A figura 5.13 (a - d) mostra a Função de Distribuição Cumulativa Empírica (FDCE) e a série temporal entre os valores observados e modelados através das redes neuronais para RESP2MF.

Na figura 13.c, verificou-se que o melhor desempenho da rede neuronal foi através do algoritmo de treinamento Backpropagation (BP) e função de activação tangente hiperbólica (TGH) para RESP2MF_TGH_BP. Observa-se através das curvas de ajustes que os valores prognosticados pela rede detectam bem a variabilidade diária e sazonal das doenças respiratórias, porém, tendem a subestimar os valores observados durante a fase de treinamento e sobrestimam na fase de testes. Deve-se ressaltar que a performance da rede melhora sensivelmente o reconhecimento do padrão de RESP2 com a inclusão de ambos os sexos durante o aprendizado.

Na Tabela 5.11, verifica-se os parâmetros da rede e estatística descritiva entre os valores observados e modelados durante o aprendizado realizado pela rede. A rede RESP2MF_TGH_BP apresentou o melhor desempenho com arquitectura [62-86-1], taxa de aprendizagem (0,1), momento (1,75) e peso de $\pm 0,9$. Os parâmetros utilizados para uma melhor calibração da rede demonstraram um comportamento semelhante ao observado em termos da média e desvio padrão, enquanto que foi verificado um erro absoluto médio de 1,5.



*Melhor rede treinada.

Figura 5.13 – Comparação entre os valores observados de RESP2MF e os valores modelados pela RNA através da Função de Distribuição Cumulativa Empírica (FDCE) e Série Temporal (ST).

Tabela 5.11 – Parâmetros da RNA e Estatística para RESP2MF

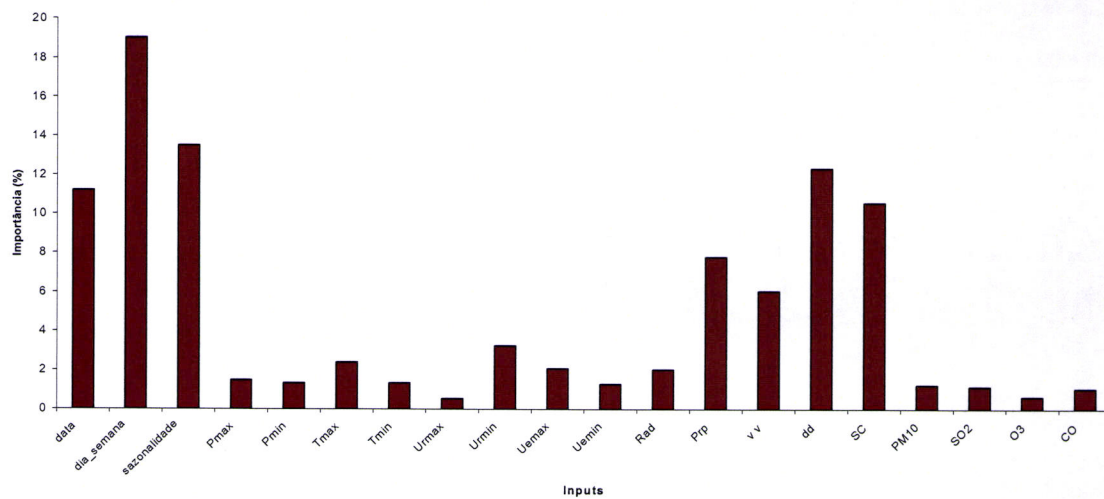
RESP2MF LOG BP					
Parâmetros		Estatística	Observado	Modelado	Erro Absoluto
Arquitetura	[62-75-1]	Média	3,9	3,7	1,3
W	± 1,5	Desv.Pad	2,4	2	1,2
TA	0,1	Min	0	0,1	0,0007
Momento	1,75	Max	12	10,6	9,7
Iterações	500	R	0,69		
Retreinamentos	3	R ²	0,47		
RESP2MF LOG DGC					
Parâmetros		Estatística	Observado	Modelado	Erro Absoluto
Arquitetura	[62-75-1]	Média	3,9	3,9	1,6
W	± 4	Desv.Pad	2,4	1,4	1,3
TA		Min	0	0,18	0,0009
Momento		Max	12	11,9	8
Iterações	500	R	0,51		
Retreinamentos	3	R ²	0,28		
RESP2MF TGH BP *					
Parâmetros		Estatística	Observado	Modelado	Erro Absoluto
Arquitetura	[62-86-1]	Média	3,9	3,8	1,5
W	± 0,9	Desv.Pad	2,4	2,6	1,6
TA	0,1	Min	0	0,001	0,0003
Momento	1,75	Max	12	11,9	9,8
Iterações	500	R	0,60		
Retreinamentos	3	R ²	0,35		
RESP2MF TGH DGC					
Parâmetros		Estatística	Observado	Modelado	Erro Absoluto
Arquitetura	[62-86-1]	Média	3,9	3,9	1,6
W	± 0,65	Desv.Pad	2,4	1,3	1,3
TA		Min	0	0,7	0,0007
Momento		Max	12	8,3	8,3
Iterações	500	R	0,49		
Retreinamentos	3	R ²	0,24		

*Melhor rede treinada.

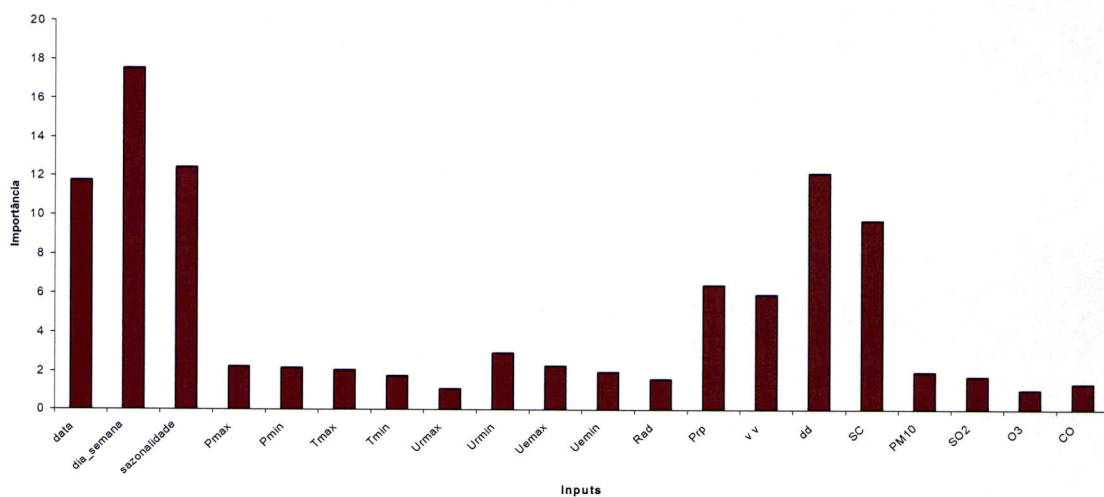
5.3.3. Análise de Sensibilidade da RN – Importância dos “inputs”

A figura 5.14 (a - c) mostra a análise de sensibilidade dos “inputs”, para os melhores desempenho da rede para a morbidade de doenças respiratórias, em adolescentes, segundo a faixa etária e sexo.

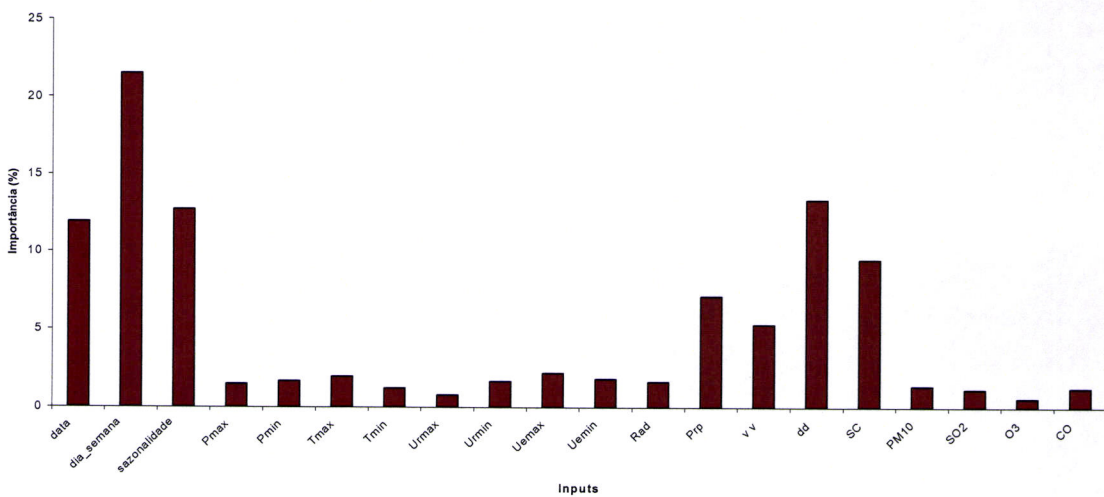
De um modo geral, observou-se que as variáveis de periodicidade denotaram uma importância de 40%, principalmente, a variabilidade semanal (dias_semana – 18%), enquanto que 55% está relacionada com as variáveis meteorológicas, principalmente a direcção do vento (dd – 14 %) e condição sinóptica (SC – 10%). As variáveis poluentes mostraram uma importância em torno de 5 %.



(a)



(b)



(c)

Figura 5.14: Análise de Sensibilidade: (a) RESP2F_TGH_BP – [62-108-1]; (b) RESP2M_LOG_BP – [62-32-1] e (c) RESP2MF_TGH_BP – [62-86-1]

5.3.4. Índice de Risco Meteorotrópico (IRM) – Matriz de Confusão

As tabelas 5.12, 5.13 e 5.14 mostram através da matriz de confusão o grau de acerto (valores em azul) do Índice de Risco Meteorotrópico (IRM) entre os valores modelados pelas melhores redes treinadas e os valores observados, durante a fase de aprendizado.

Os valores prognosticados pelas redes (RESP2F_TGH_BP – [62-108-1] e RESP2M_LOG_BP – [62-32-1]) apresentaram um grau de acerto em torno de 56% sobre o valor observado, enquanto que para RESP2MF_TGH_BP – [62-86-1] o grau de acerto foi em torno de 64%.

O IRM 1 ($RESP2 \leq q_{0.25}$), faixa com maiores ocorrências de morbidades para RESP2, apresentou uma **Baixa Frequência** em relação ao valor esperado pelas afecções respiratórias, enquanto que a classe de **Frequência Normal** (IRM2) e **Frequência Acima do Normal** (IRM3) tenderam a sobrestimar e subestimar os valores observados, respectivamente. Para o IRM 4 não foi prognosticado nenhuma **Frequência de Alerta**. As informações obtidas pela matriz de confusão corroboram com a explicação descrita nas análises de Funções de Distribuição Cumulativa Empírica (FDCE). Deve-se ressaltar que a pouca robustez na generalização dos valores preditos pela rede poderia ser otimizado através da inclusão de mais uma camada escondida.

Tabela 5.12 – Matriz de Confusão para RESP2F_TGH_BP – [62-108-1].

MODELADO	OBSERVADO			
	IRM 1	IRM 2	IRM 3	IRM 4
	$RESP2F \leq q_{0.25}$	$q_{0.25} < RESP2F \leq q_{0.50}$	$q_{0.50} < RESP2F \leq q_{0.75}$	$RESP2F \geq q_{0.75}$
	≥ 1	[2....3]	[4....6]	≤ 7
≥ 1	1166	349	55	0
[2....3]	745	551	82	0
[4....6]	81	85	29	0
≤ 7	9	7	10	0

Tabela 5.13 – Matriz de Confusão para RESP2M_LOG_BP – [62-32-1].

MODELADO	OBSERVADO			
	IRM 1	IRM 2	IRM 3	IRM 4
	$RESP2M \leq q_{0.25}$	$q_{0.25} < RESP2M \leq q_{0.50}$	$q_{0.50} < RESP2M \leq q_{0.75}$	$RESP2M \geq q_{0.75}$
	≥ 1	[2...3]	[4...6]	≤ 7
≥ 1	1066	439	65	0
[2...3]	725	595	68	0
[4...6]	51	75	69	0
≤ 7	5	4	17	0

Tabela 5.14 – Matriz de Confusão Para RESP2MF_TGH_BP – [62-86-1].

MODELADO	OBSERVADO			
	IRM 1	IRM 2	IRM 3	IRM 4
	$RESP2MF \leq q_{0.25}$	$q_{0.25} < RESP2MF \leq q_{0.50}$	$q_{0.50} < RESP2MF \leq q_{0.75}$	$RESP2MF \geq q_{0.75}$
	≥ 2	[3...4]	[5...11]	≤ 12
≥ 2	1218	300	83	0
[3...4]	503	635	107	0
[5...11]	81	171	171	0
≤ 12	4	2	22	0

5.4. Resultados para RESP3 (M, F e MF)

5.4.1. Análise em Componentes Principais (ACP) e Análise de “Cluster”

Na tabela 5.15, Observou-se a extracção de 3 componentes após a rotação varimax para RESP3M, RESP3M RESP3MF, com uma variância acumulada (Var. Acum.) em torno de 70%.

Para CP1, verificou-se uma variância explicada (Var. Expl.) de aproximadamente 31%, com pesos positivos para os poluentes SO₂, MP₁₀ e CO, em relação à RESP2F (0.59), RESP2M (0.63) e RESP2MF (0.67), assim como pesos negativos significativos para a velocidade do vento (vv) e pouco significativos para a temperatura mínima (Tmin) e humidade específica mínima (Uemin). Esta componente tem o mesmo padrão verificado para a extracção da CP1 para RESP1 e RESP2 (M, F e MF).

Na CP2, observou-se uma variância explicada de 22% com pesos positivos para as variáveis Tmin (0.95) e Uemin (0.96) em oposição à RESP3F, RESP3M, RESP3MF, SO₂ e MP₁₀. Esta componente indica que o aumento de temperatura e humidade específica do ar favorecem o decréscimo das doenças respiratórias, conforme CP2 de RESP1 (M, F e MF).

Na CP3, verifica-se uma variância explicada na ordem de 17%, com pesos positivos para a radiação solar (Rad) e ozono (O₃), apresentando a mesma variabilidade explicada quando comparada com RESP1 e RESP2 (M, F e MF).

Tabela 5.15 – Pesos das componentes principais após a rotação varimax para todo o período de estudo (01/01/1993 a 31/12/2002) – RESP3

Variáveis	CP1	CP2	CP3
Tmin	-0,090926	0,950443	0,092235
Uemin	-0,056927	0,957644	0,030702
Rad	0,008746	0,070414	0,701668
vv	-0,377477	0,062128	-0,424301
PM10	0,877071	-0,138632	0,239907
SO2	0,774531	-0,269133	0,210762
O3	0,093334	0,146781	0,858032
CO	0,834133	0,039275	-0,193296
RESP3_F	0,585672	-0,172958	-0,372954
Var. Expl. (%)	30,62403	21,58100	17,09378
Var. Acum. (%)	30,62403	52,20504	69,29881
Variáveis	CP1	CP2	CP3
Tmin	-0,092398	0,949085	0,095583
Uemin	-0,048165	0,960976	0,026485
Rad	-0,008771	0,063554	0,708101
vv	-0,365188	0,066795	-0,433714
PM10	0,867160	-0,144998	0,257563
SO2	0,772020	-0,271037	0,219601
O3	0,084742	0,146757	0,855794
CO	0,839158	0,037905	-0,182407
RESP3_M	0,631756	-0,155291	-0,337885
Var. Expl. (%)	31,00154	-21,59426	17,07994
Var. Acum. (%)	31,00154	52,59580	69,67573
Variáveis	CP1	CP2	CP3
Tmin	-0,087343	0,950492	0,093044
Uemin	-0,044833	0,960616	0,027047
Rad	-0,011516	0,066875	0,702061
vv	-0,358267	0,069973	-0,440453
PM10	0,865576	-0,147900	0,265773
SO2	0,766728	-0,275285	0,229749
O3	0,071048	-0,143815	0,860885
CO	0,843770	0,035433	-0,174952
RESP3_MF	0,666565	-0,156987	-0,347981
Var. Expl. (%)	31,26167	21,59872	17,46040
Var. Acum. (%)	31,26167	52,86039	70,32079

A figura 5.15, mostra o dendograma de árvore através do método de Ward e distâncias euclidianas, onde foi verificado a formação de um grupo conectado apenas pelos poluentes CO, MP₁₀ e SO₂ e RESP3 (M, F e MF), demonstrando uma boa similaridade em relação a RESP3 (M, F e MF). Um segundo grupo formado pelas variáveis meteorológicas temperatura mínima (Tmin), humidade específica mínima (Uemin) e velocidade do vento (vv), assim como a formação de um terceiro grupo composto pela radiação solar (Rad) e o ozono (O₃), conforme observado em RESP1 e RESP2 (M, F e MF).

Ressalta-se que os grupos descritos pela AC apresentaram o mesmo padrão demonstrado nas extracções das componentes através da ACP.

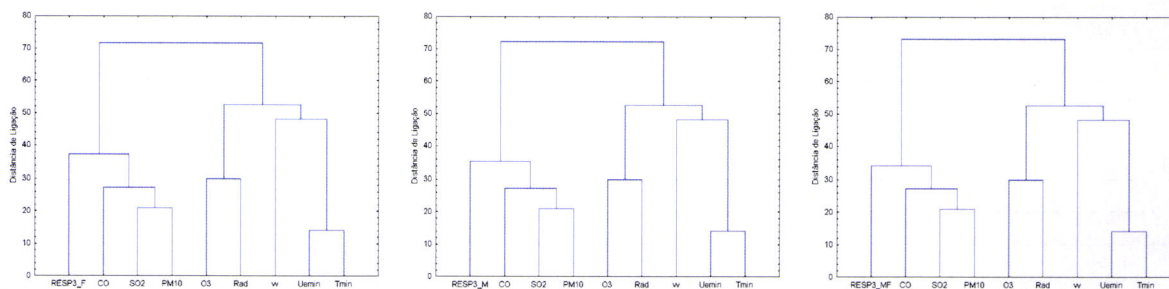


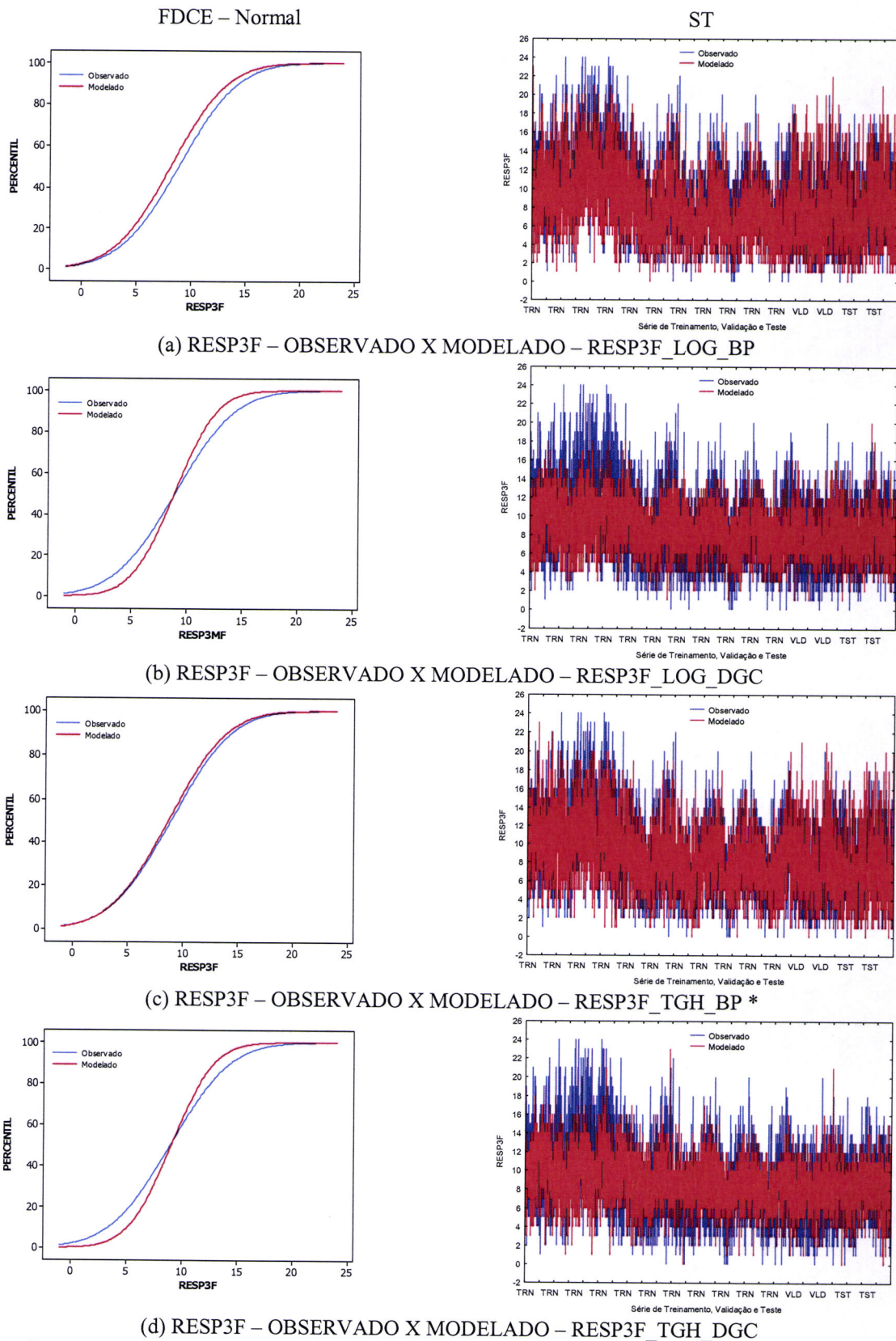
Figura 5.15 – Dendrograma de árvore através do método de Ward e distâncias euclidianas para todo o período de estudo: RESP3

5.4.2. Prognósticos Gerados pela RN – FDCE (Observado X Modelado)

A figura 5.16 (a - d) mostra a Função de Distribuição Cumulativa Empírica (FDCE) e a série temporal entre os valores observados e modelados através das redes neuronais para RESP3F.

Na figura 5.16c, verificou-se que o melhor desempenho da rede neuronal foi através do algoritmo de treinamento Backpropagation (BP) e função de activação tangente hiperbólica (TGH) para RESP3F_TGH_BP. Observa-se através das curvas de ajustes que os valores prognosticados pela rede detectam bem a variabilidade diária e sazonal das doenças respiratórias, porém, tendem a subestimar os valores observados durante a fase de treinamento e sobrestimam na fase de testes.

Na Tabela 5.16, verifica-se os parâmetros da rede e estatística descritiva entre os valores observados e modelados durante o aprendizado realizado pela rede. A rede RESP3F_TGH_BP apresentou o melhor desempenho com arquitectura [62-141-1], taxa de aprendizagem (0,1), momento (1,75) e peso de $\pm 0,5$. Os parâmetros utilizados para uma melhor calibração da rede demonstraram um comportamento semelhante ao observado em termos da média e desvio padrão, enquanto que foi verificado um erro absoluto médio de 2,6.



* Melhor rede treinada

Figura 5.16 – Comparação entre os valores observados de RESP3F e os valores modelados pela RNA através da Função de Distribuição Cumulativa Empírica (FDCE) e Série Temporal (ST).

Tabela 5.16 – Parâmetros da RNA e Estatística para RESP3F

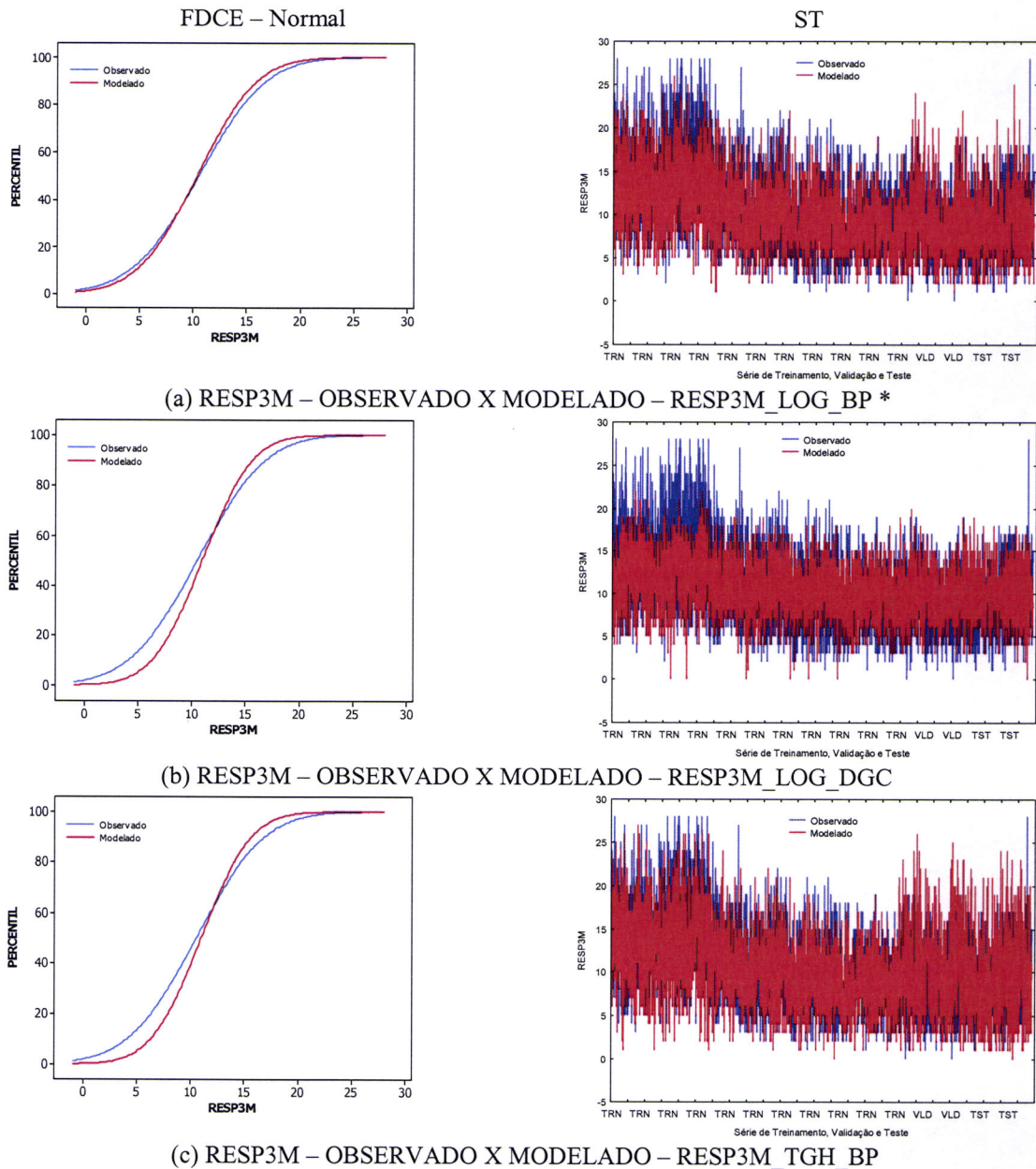
RESP3F LOG BP					
Parâmetros		Estatística	Observado	Modelado	Erro Absoluto
Arquitectura	[62-134-1]	Média	8,9	8,2	2,3
W	$\pm 1,3$	Desv.Pad	4,3	4,1	2,2
TA	0,1	Min	0	0,4	0,0008
Momento	1,75	Max	24	23,2	17,9
Iterações	500	R		0,71	
Retreinamentos	3	R ²		0,54	
RESP3F LOG DGC					
Parâmetros		Estatística	Observado	Modelado	Erro Absoluto
Arquitectura	[62-134-1]	Média	8,9	8,9	2,7
W	± 3	Desv.Pad	4,3	3	2,1
TA		Min	0	0,7	0,0003
Momento		Max	24	20,2	12,9
Iterações	500	R		0,61	
Retreinamentos	3	R ²		0,37	
RESP3F TGH BP *					
Parâmetros		Estatística	Observado	Modelado	Erro Absoluto
Arquitectura	[62-141-1]	Média	8,9	8,7	2,6
W	$\pm 0,5$	Desv.Pad	4,3	4,1	2,3
TA	0,1	Min	0	0,1	0,002
Momento	1,75	Max	24	22,90	14,9
Iterações	500	R		0,67	
Retreinamentos	3	R ²		0,45	
RESP3F TGH DGC					
Parâmetros		Estatística	Observado	Modelado	Erro Absoluto
Arquitectura	[62-141-1]	Média	8,9	9	2,7
W	$\pm 1,3$	Desv.Pad	4,3	3,1	2,1
TA		Min	0	0,3	0,0004
Momento		Max	24	23	13,4
Iterações	500	R		0,61	
Retreinamentos	3	R ²		0,37	

* Melhor rede treinada

A figura 5.17 (a - d) mostra a Função de Distribuição Cumulativa Empírica (FDCE) e a série temporal entre os valores observados e modelados através das redes neuronais para RESP3M.

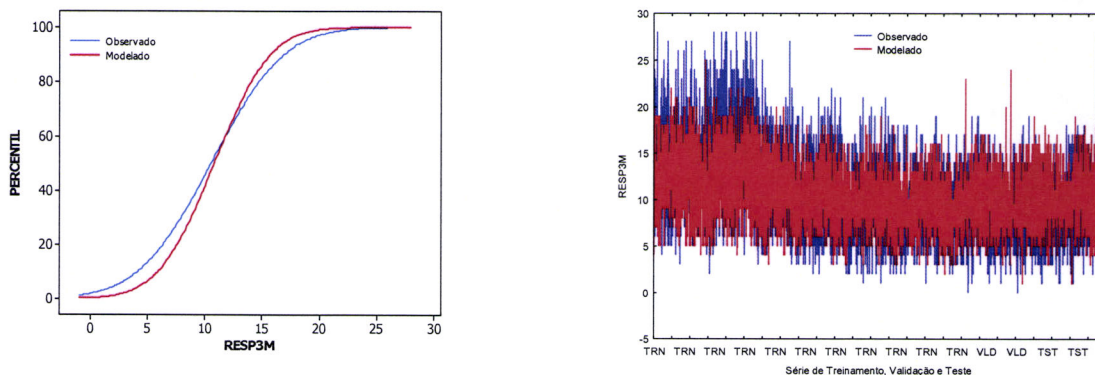
Na figura 5.17a verificou-se que o melhor desempenho da rede neuronal foi através do algoritmo de treinamento Backpropagation (BP) e função de activação logística (LOG) para RESP3M_LOG_BP. Observa-se através das curvas de ajustes que os valores prognosticados pela rede detectam bem a variabilidade diária e sazonal das doenças respiratórias, porém, tendem a subestimar os valores observados durante a fase de treinamento e sobrestimam na fase de validação.

Na Tabela 5.17 verifica-se os parâmetros da rede e estatística descritiva entre os valores observados e modelados durante o aprendizado realizado pela rede. A rede RESP3M_LOG_BP apresentou o melhor desempenho com arquitectura [62-135-1], taxa de aprendizagem (0,1), momento (1,75) e peso de ± 1 . Os parâmetros utilizados para uma melhor calibração da rede demonstraram um comportamento semelhante ao observado em termos da média e desvio padrão, enquanto que foi verificado um erro absoluto médio de 2,6.



*Melhor rede treinada

Figura 5.17 – Comparação entre os valores observados de RESP3M e os valores modelados pela RNA através da Função de Distribuição Cumulativa Empírica (FDCE) e Série Temporal (ST).



(d) RESP3M – OBSERVADO X MODELADO – RESP3M_TGH_DGC

* Melhor rede treinada

Figura 5.17 – cont.

Tabela 5.17 – Parâmetros da RNA e Estatística para RESP3M

RESP3M LOG BP *					
Parâmetros		Estatística	Observado	Modelado	Erro Absoluto
Arquitetura	[62-135-1]	Média	10,6	10,4	2,6
W	± 1	Desv.Pad	5	4,4	2,3
TA	0,1	Min	0	0,9	0,001
Momento	1,75	Max	28	26,2	20,6
Iterações	500	R		0,73	
Retreinamentos	3	R ²		0,54	
RESP3M LOG DGC					
Parâmetros		Estatística	Observado	Modelado	Erro Absoluto
Arquitetura	[62-135-1]	Média	10,6	11	3,1
W	± 5	Desv.Pad	5	3,6	2,5
TA		Min	0	0,1	0,002
Momento		Max	28	22,2	14,9
Iterações	500	R		0,62	
Retreinamentos	3	R ²		0,38	
RESP3M TGH BP					
Parâmetros		Estatística	Observado	Modelado	Erro Absoluto
Arquitetura	[62-105-1]	Média	10,6	10,7	3
W	± 0,6	Desv.Pad	5	5	2,7
TA	0,1	Min	0	0,4	0,0009
Momento	1,75	Max	28	26,8	19,8
Iterações	500	R		0,67	
Retreinamentos	3	R ²		0,45	
RESP3M TGH DGC					
Parâmetros		Estatística	Observado	Modelado	Erro Absoluto
Arquitetura	[62-65-1]	Média	10,6	10,9	3
W	± 2	Desv.Pad	5	3,8	2,3
TA		Min	0	0,7	0,002
Momento		Max	28	25,3	18,2
Iterações	500	R		0,66	
Retreinamentos	3	R ²		0,43	

*Melhor rede treinada

A figura 5.18 (a - d) mostra a Função de Distribuição Cumulativa Empírica (FDCE) e a série temporal entre os valores observados e modelados através das redes neuronais para RESP3MF.

Na figura 5.18a, verificou-se que o melhor desempenho da rede neuronal foi através do algoritmo de treinamento Backpropagation (BP) e função de activação logística (LOG) para RESP3MF_LOG_BP. Observa-se através das curvas de ajustes que os valores prognosticados pela rede detectam bem a variabilidade diária e sazonal das doenças respiratórias, porém, tendem a subestimar os valores observados durante a fase de treinamento e sobrestimaram na fase de testes e validação. Deve-se ressaltar que a performance da rede melhora sensivelmente o reconhecimento do padrão de RESP3 com a inclusão de ambos os sexos durante o aprendizado.

Na Tabela 5.18, verifica-se os parâmetros da rede e estatística descritiva entre os valores observados e modelados durante o aprendizado realizado pela rede. A rede RESP3MF_LOG_BP apresentou o melhor desempenho com arquitectura [62-92-1], taxa de aprendizagem (0,1), momento (1,75) e peso de $\pm 1,5$. Os parâmetros utilizados para uma melhor calibração da rede demonstraram um comportamento semelhante ao observado em termos da média e desvio padrão, enquanto que foi verificado um erro absoluto médio de 3,8.

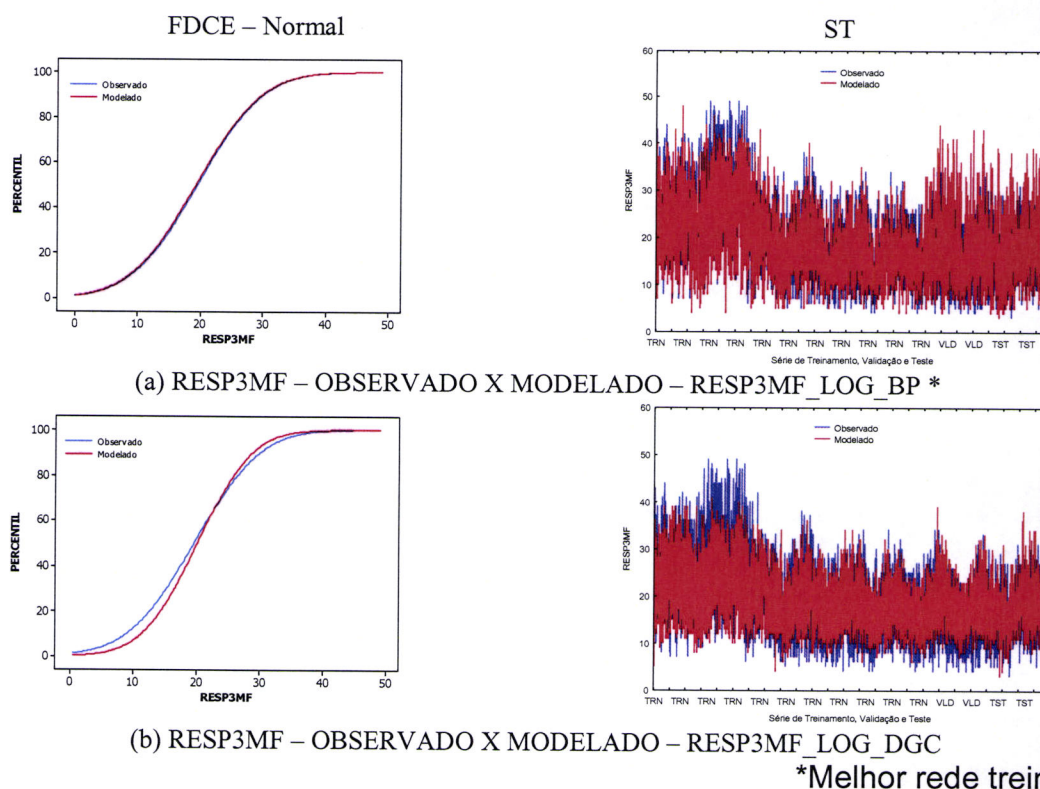
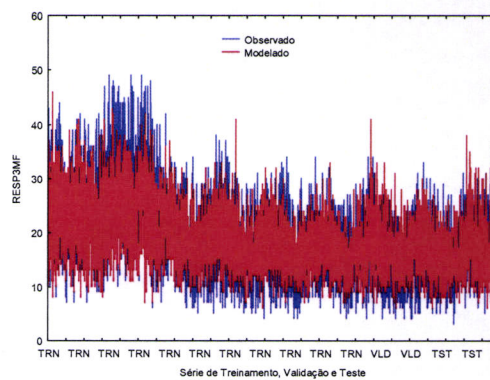
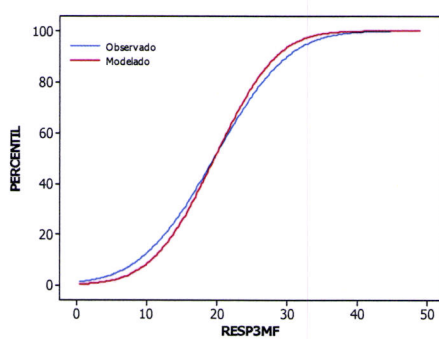
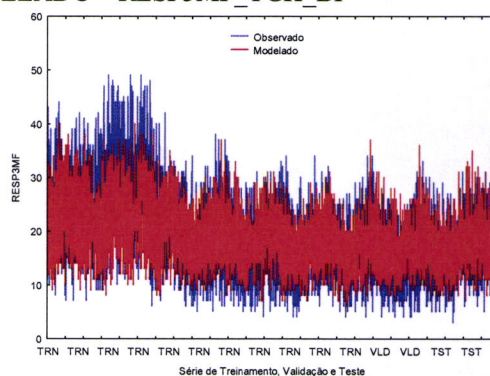
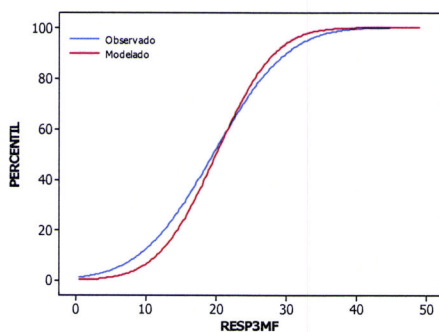


Figura 5.18 – Comparação entre os valores observados de RESP3MF e os valores modelados pela RNA através da Função de Distribuição Cumulativa Empírica (FDCE) e Série Temporal (ST).



(c) RESP3MF – OBSERVADO X MODELADO – RESP3MF_TGH_BP



(d) RESP3MF – OBSERVADO X MODELADO – RESP3MF_TGH_DGC

*Melhor rede treinada

Figura 5.18 – cont.

Tabela 5.18 – Parâmetros da RNA e Estatística para RESP3MF

RESP3MF LOG BP *					
Parâmetros		Estatística	Observado	Modelado	Erro Absoluto
Arquitetura	[62-92-1]	Média	19,7	19,4	3,8
W	± 1,5	Desv.Pad	8,2	8,3	3,8
TA	0,1	Mín	3	3,5	0,001
Momento	1,75	Max	49	47,6	34,9
Iterações	500	R	0,79		
Retreinamentos	3	R ²	0,63		
RESP3MF LOG DGC					
Parâmetros		Estatística	Observado	Modelado	Erro Absoluto
Arquitetura	[62-92-1]	Média	19,7	20,1	4,3
W	± 4	Desv.Pad	8,2	6,8	3,5
TA		Mín	3	3,1	0,0001
Momento		Max	49	41,3	24,1
Iterações	500	R	0,75		
Retreinamentos	3	R ²	0,56		
RESP3MF TGH BP					
Parâmetros		Estatística	Observado	Modelado	Erro Absoluto
Arquitetura	[62-106-1]	Média	19,7	19,7	4
W	± 0,3	Desv.Pad	8,2	6,9	3,2
TA	0,1	Mín	3	6	0,0003
Momento	1,75	Max	49	45,9	23,9
Iterações	500	R	0,79		
Retreinamentos	3	R ²	0,62		
RESP3MF TGH_DGC					

Parâmetros		Estadística	Observado	Modelado	Erro Absoluto
Arquitectura	[62-106-1]	Média	19,7	20,1	4,3
W	$\pm 0,6$	Desv.Pad	8,2	6,5	3,4
TA		Min	3	5,6	0,0009
Momento		Max	49	40,2	27,5
Iterações	500	R	0,75		
Retreinamentos	3	R ²	0,55		

*Melhor rede treinada

Tabela 5.18 – cont.

5.4.3. Análise de Sensibilidade da RN – Importância dos “inputs”

A figura 5.19 (a - c) demonstra a análise de sensibilidade dos “inputs”, para os melhores desempenho da rede para a morbilidade de doenças respiratórias, em adultos jovens, segundo a faixa etária e sexo.

De um modo geral, observou-se que as variáveis de periodicidade denotaram uma importância de 45%, principalmente, a variabilidade semanal (dias_semana – 20%), enquanto que 50% está relacionada com as variáveis meteorológicas, principalmente a direcção do vento (dd – 14 %) e condição sinóptica (SC – 10%). As variáveis poluentes mostraram uma importância em torno de 5 %, principalmente, em relação ao monóxido de carbono (CO).

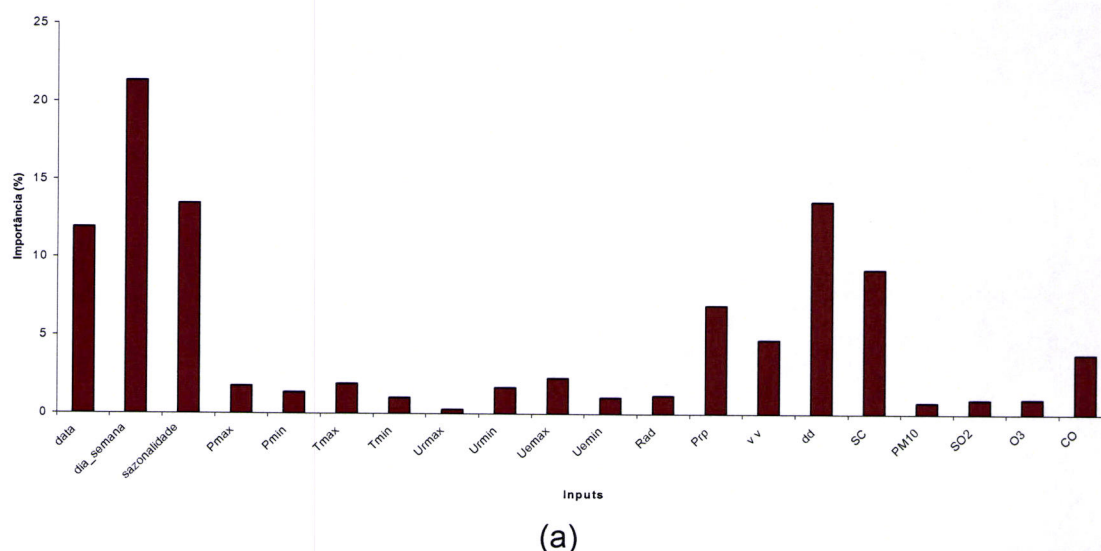
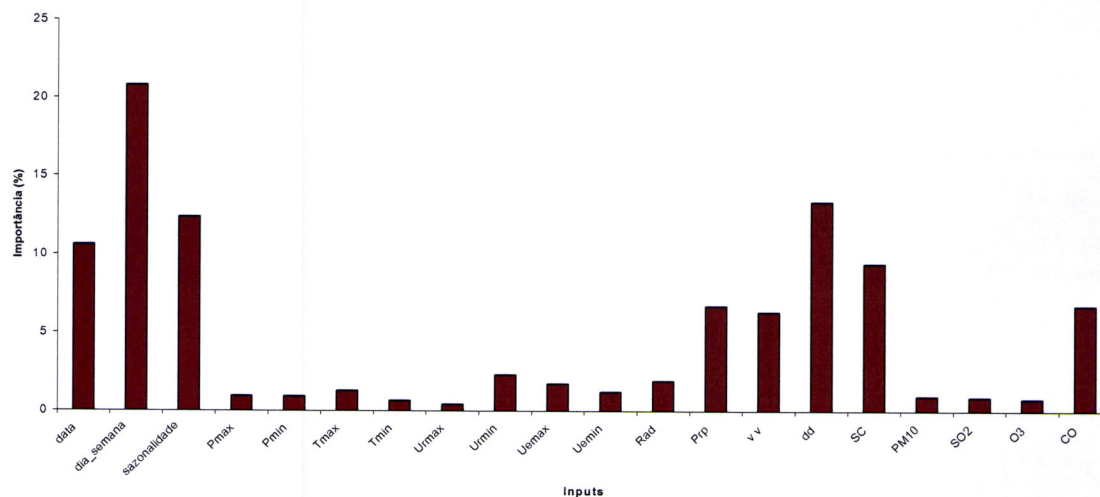
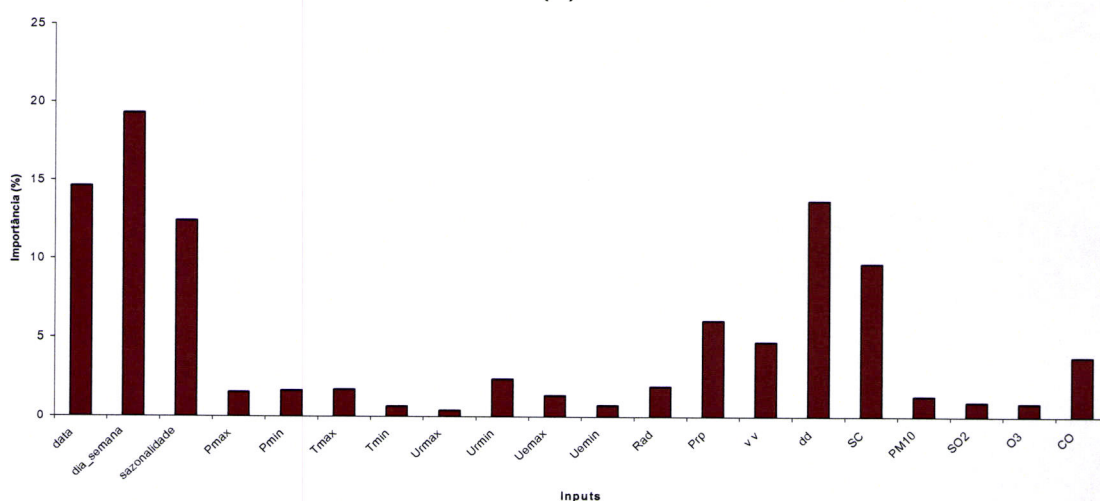


Figura 5.19 – Análise de Sensibilidade: (a) RESP3F_TGH_BP – [62-141-1]; (b) RESP3M_LOG_BP – [62-135-1] e (c) RESP3MF_LOG_BP – [62-92-1]



(b)



(c)

Figura 5.19 – cont.

5.4.4. Índice de Risco Meteorotrópico (IRM) – Matriz de Confusão

As tabelas 5.19, 5.20 e 5.21 mostram através da matriz de confusão o grau de acerto (valores em azul) do Índice de Risco Meteorotrópico (IRM) entre os valores preditos pelas melhores redes treinadas e os valores observados, durante a fase de aprendizado.

Os valores prognosticados pelas redes (RESP3F_TGH_BP – [62-141-1] e RESP3M_LOG_BP – [62-135-1]) apresentaram um grau de acerto em torno de 72% sobre o valor observado, enquanto que para RESP3MF_LOG_BP – [62-92-1] o grau de acerto foi em torno de 78%.

O IRM 2 ($q_{0.25} < \text{RESP3} \leq q_{0.50}$), faixa com maiores ocorrências de morbidades para RESP1, apresentou uma **Frequência Normal** em relação ao valor esperado pelas afecções respiratórias, enquanto que a classe de **Baixa Frequência** (IRM1) e **Frequência Acima do Normal** (IRM3) tenderam a sobrestimar e subestimar os valores observados, respectivamente. Para o IRM 4 não foi prognosticado nenhuma **Frequência de Alerta**. A informação da matriz de confusão corrobora com a explicação descrita nas análises de Funções de Distribuição Cumulativa Empírica (FDCE).

Tabela 5.19 – Matriz de Confusão para RESP3F_TGH_BP – [62-141-1].

MODELADO	OBSERVADO			
	IRM 1	IRM 2	IRM 3	IRM 4
	$\text{RESP3F} \leq q_{0.25}$	$q_{0.25} < \text{RESP3F} \leq q_{0.50}$	$q_{0.50} < \text{RESP3F} \leq q_{0.75}$	$\text{RESP3F} \geq q_{0.75}$
	≥ 6	[7....11]	[12....23]	≤ 24
≥ 6	995	210	53	0
[7....11]	311	1090	149	0
[12....23]	18	100	225	0
≤ 24	0	5	23	0

Tabela 5.20 – Matriz de Confusão para RESP3M_LOG_BP – [62-135-1].

MODELADO	OBSERVADO			
	IRM 1	IRM 2	IRM 3	IRM 4
	$RESP3M \leq q_{0.25}$	$q_{0.25} < RESP3M \leq q_{0.50}$	$q_{0.50} < RESP3M \leq q_{0.75}$	$RESP3M \geq q_{0.75}$
	≥ 7	[8....13]	[14....27]	≤ 28
≥ 7	1015	213	46	0
[8....13]	328	1052	169	0
[14....27]	12	109	205	0
≤ 28	0	9	21	0

Tabela 5.21 – Matriz de Confusão Para RESP3MF_LOG_BP – [62-92-1].

MODELADO	OBSERVADO			
	IRM 1	IRM 2	IRM 3	IRM 4
	$RESP3MF \leq q_{0.25}$	$q_{0.25} < RESP3MF \leq q_{0.50}$	$q_{0.50} < RESP3MF \leq q_{0.75}$	$RESP3MF \geq q_{0.75}$
	≥ 14	[15....25]	[26....48]	≤ 49
≥ 14	1032	165	21	0
[15....25]	275	1133	189	0
[26....48]	15	89	231	0
≤ 49	0	6	23	0

5.5. Resultados para RESP4 (M, F e MF)

5.5.1. Análise em Componentes Principais (ACP) e Análise de “Cluster”

Na tabela 5.22, Observou-se a extracção de 4 componentes após a rotação varimax para RESP4M, RESP4M e RESP4MF, com uma variância acumulada em torno de 80%.

Para a CP1, verificou-se uma variância explicada de aproximadamente 30%, com pesos positivos para os poluentes SO₂, MP₁₀ e CO, em relação à RESP4F (0.59), RESP4M (0.63) e RESP4MF (0.67), assim como pesos negativos significativos para a velocidade do vento (vv) e pouco significativos para a temperatura mínima (Tmin) e humidade específica mínima (Uemin).

Na CP2, observou-se uma variância explicada de 22% com pesos positivos para as variáveis Tmin (0.95) e Uemin (0.96) em oposição à RESP3 (M, F e MF), SO₂ e MP₁₀. Esta componente explica a mesma variabilidade encontrada na CP2 de RESP1, RESP2 e RESP3 (M, F e MF).

Na CP3, verifica-se uma variância explicada na ordem de 17%, com pesos positivos para a radiação solar (Rad) e ozono (O₃), apresentando a mesma variabilidade explicada quando comparada com RESP1 e RESP2 e RESP3 (M, F e MF).

Para CP4 foi observado o mesmo padrão denotado pela extracção da CP4 em RESP1 e RESP2 e RESP3 (M, F e MF).

Tabela 5.22 – Pesos das componentes principais após a rotação varimax para todo o período de estudo (01/01/1993 a 31/12/2002) – RESP4

Variáveis	CP1	CP2	CP3	CP4
Tmin	-0,065800	0,952307	0,125466	-0,076619
Uemin	-0,078568	0,965527	-0,022297	0,071192
Rad	0,054655	0,029362	0,918749	-0,133561
vv	-0,206812	0,011074	0,025670	0,893955
PM10	0,838315	-0,127177	0,152630	0,328048
SO2	0,772847	-0,262813	0,193254	0,180990
O3	0,025530	0,155467	0,715439	0,477565
CO	0,845883	0,035410	-0,137729	-0,018717
RESP4_F	0,569012	-0,080026	-0,132719	-0,330852
Var. Expl. (%)	30,29235	21,55258	16,12091	11,25330
Var. Acum. (%)	30,29235	51,84494	67,96585	79,21914
Variáveis	CP1	CP2	CP3	CP4
Tmin	0,063590	0,945837	0,130422	-0,070238
Uemin	0,065768	0,965016	-0,021887	0,065469
Rad	-0,058358	0,032920	0,917652	-0,136907
vv	-0,176636	0,008171	0,032848	0,909632
PM10	0,819734	-0,144774	0,165738	0,352960
SO2	0,752919	-0,280696	0,206445	0,208761
O3	0,002664	0,157337	0,710226	0,494301
CO	0,856292	0,018063	-0,121111	-0,012375
RESP4_M	0,632837	-0,175013	-0,156008	-0,254582
Var. Expl. (%)	30,43249	21,64576	16,77130	11,20656
Var. Acum. (%)	30,43249	52,07825	68,84955	80,05611
Variáveis	CP1	CP2	CP3	CP4
Tmin	0,058398	0,949805	0,127914	-0,074922
Uemin	0,067785	0,965440	-0,022592	0,066726
Rad	-0,055262	0,032638	0,917801	-0,135709
vv	-0,168668	0,012737	0,030810	0,913914
PM10	0,818410	-0,141976	0,168069	0,355894
SO2	0,755091	-0,276766	0,208062	0,208480
O3	0,006211	0,155902	0,710471	0,493031
CO	0,854008	0,021178	-0,118542	-0,005700
RESP4_MF	0,662346	-0,132617	-0,167510	-0,247360
Var. Expl. (%)	30,74038	21,59145	16,87084	11,22854
Var. Acum. (%)	30,74038	52,33184	69,20267	80,43122

Na figura 5.15 foi verificado o mesmo padrão de conexão dos grupos obtidos pela AC quando comparado com a RESP3 (M, F e MF). Ressalta-se que os grupos descritos pela AC apresentaram o mesmo modelo demonstrado nas extracções das componentes através da ACP.

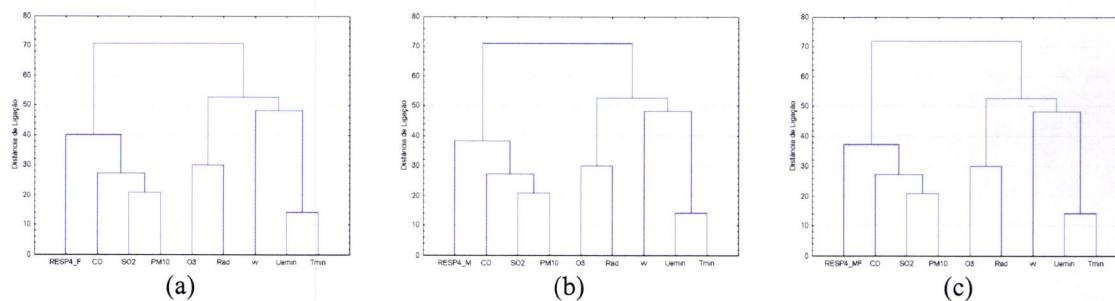


Figura 5.20 – Dendrograma de árvore através do método de Ward e distâncias euclidianas para todo o período de estudo: RESP4

5.5.2. Prognósticos Gerados pela RN – FDCE (Observado X Modelado)

A figura 5.21 (a - d) mostra a Função de Distribuição Cumulativa Empírica (FDCE) e a série temporal entre os valores observados e modelados através das redes neurais para RESP4F.

Na figura 5.21c verificou-se que o melhor desempenho da rede neuronal foi através do algoritmo de treinamento Backpropagation (BP) e função de activação tangente hiperbólica (TGH) para RESP4F_TGH_BP. Observa-se através das curvas de ajustes que os valores prognosticados pela rede detectam bem a variabilidade diária e sazonal das doenças respiratórias, porém, tendem a subestimar os valores observados durante a fase de treinamento e sobrestimam na fase de testes.

Na Tabela 5.23 verifica-se os parâmetros da rede e estatística descritiva entre os valores observados e modelados durante o aprendizado realizado pela rede. A rede RESP4F_TGH_BP apresentou o melhor desempenho com arquitetura [62-95-1], taxa de aprendizagem (0,1), momento (1,75) e peso de $\pm 0,7$. Os parâmetros utilizados para uma melhor calibração da rede demonstraram um comportamento semelhante ao observado em termos da média e desvio padrão, enquanto que foi verificado um erro absoluto médio de 2,1.

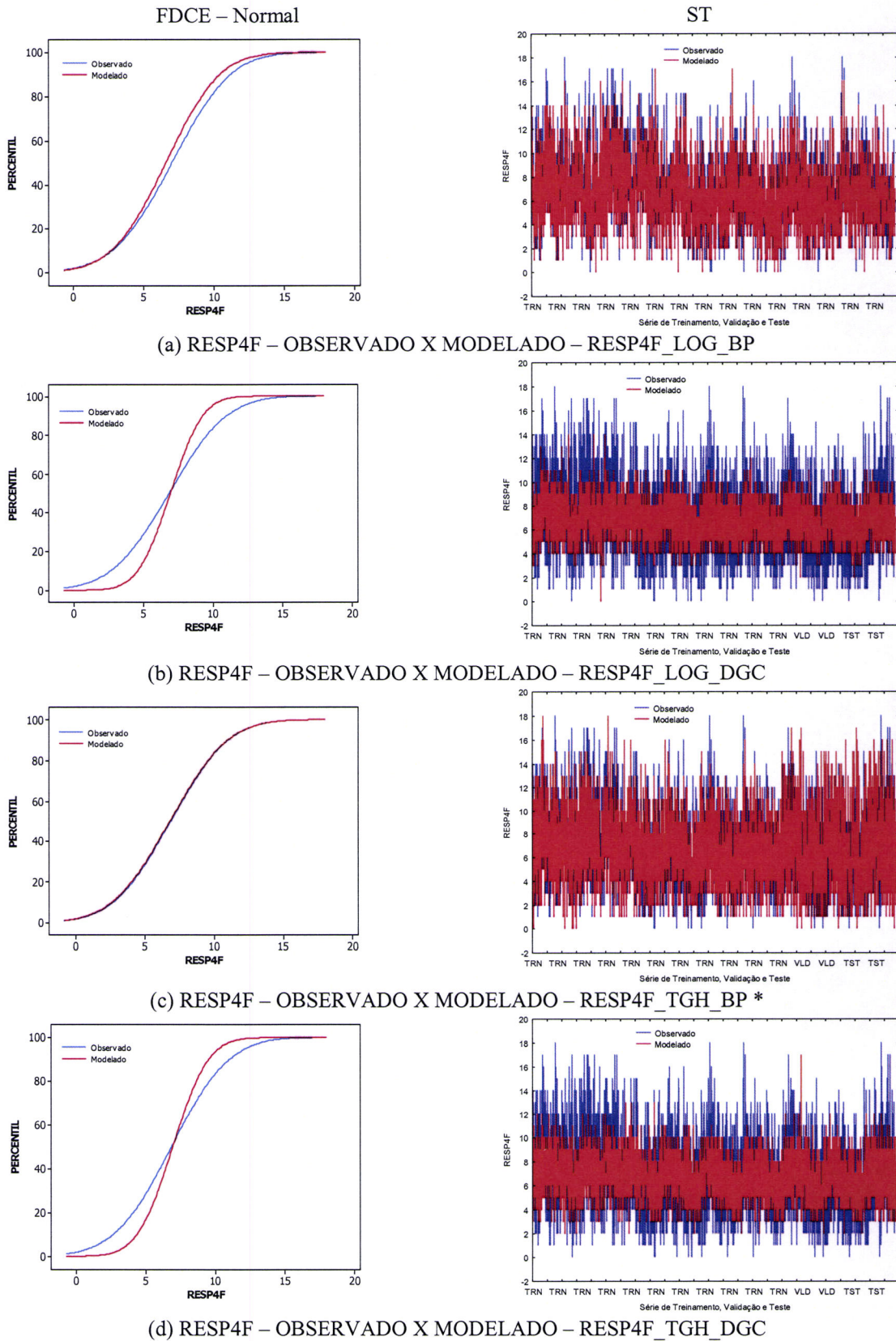


Figura 5.21 – Comparação entre os valores observados de RESP4F e os valores modelados pela RNA através da Função de Distribuição Cumulativa Empírica (FDCE) e Série Temporal (ST).

Tabela 5.23 – Parâmetros da RNA e Estatística para RESP4F

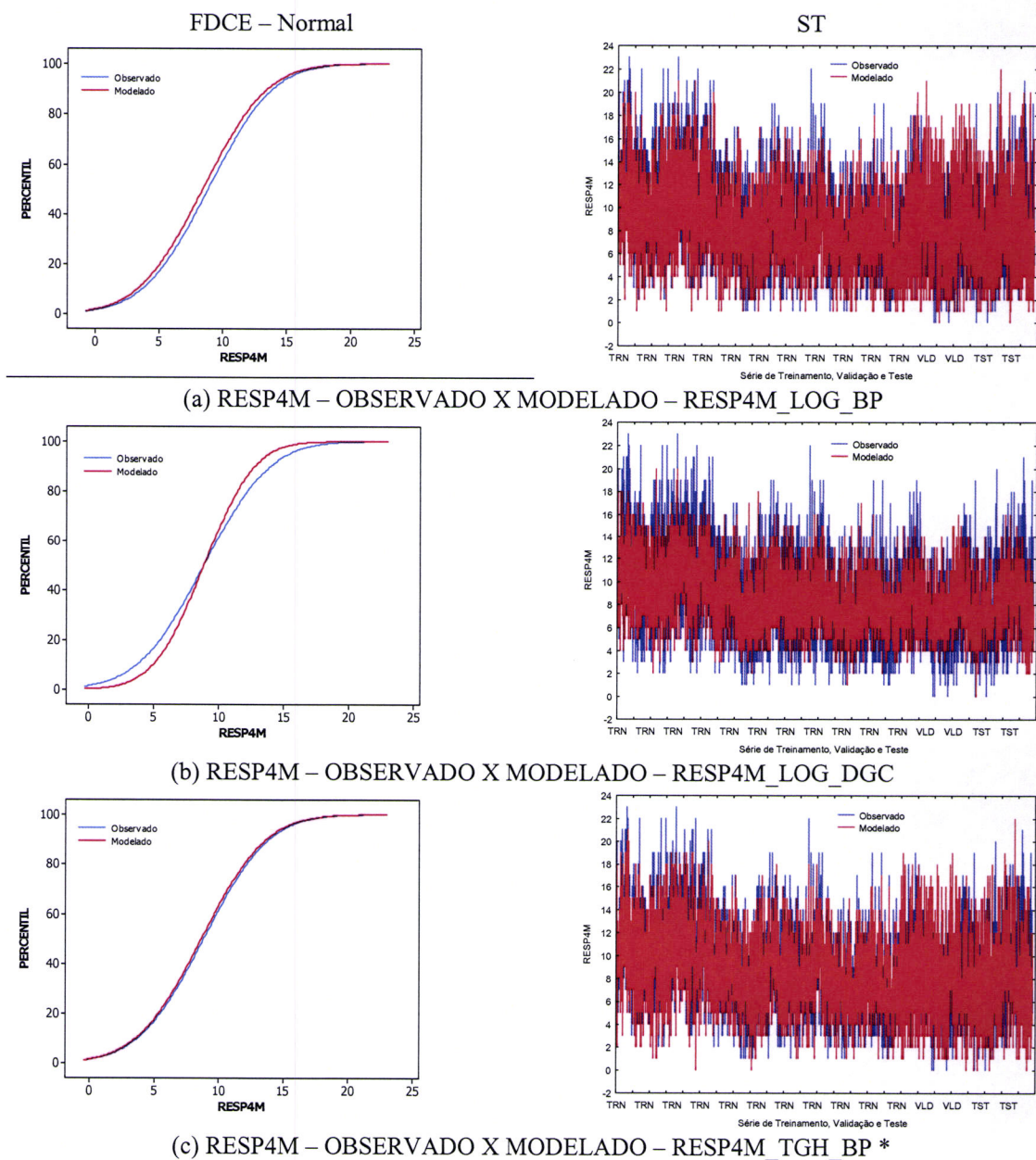
RESP4F LOG BP					
Parâmetros		Estatística	Observado	Modelado	Erro Absoluto
Arquitectura	[62-82-1]	Média	7,1	6,6	1
W	± 5	Desv.Pad	3,3	3	0,8
TA	0,1	Min	0	0,07	0,0004
Momento	1,5	Max	18	17,4	7,8
Iterações	500	R		0,61	
Retreinamentos	3	R ²		0,38	
RESP4F LOG DGC					
Parâmetros		Estatística	Observado	Modelado	Erro Absoluto
Arquitectura	[62-82-1]	Média	7,1	7	2,2
W	± 11	Desv.Pad	3,3	1,8	1,7
TA		Min	0	0,0005	0,003
Momento		Max	18	14,5	10,9
Iterações	500	R		0,53	
Retreinamentos	3	R ²		0,28	
RESP4F TGH BP *					
Parâmetros		Estatística	Observado	Modelado	Erro Absoluto
Arquitectura	[62-95-1]	Média	7,1	6,8	2,1
W	± 0,7	Desv.Pad	3,3	3,3	2
TA	0,1	Min	0	0,2	0,0007
Momento	1,75	Max	18	17,9	15,2
Iterações	500	R		0,62	
Retreinamentos	3	R ²		0,39	
RESP4F TGH DGC					
Parâmetros		Estatística	Observado	Modelado	Erro Absoluto
Arquitectura	[62-95-1]	Média	7,1	7	2,2
W	± 1,3	Desv.Pad	3,3	2	1,7
TA		Min	0	1,8	0,002
Momento		Max	18	17,1	12,1
Iterações	500	R		0,54	
Retreinamentos	3	R ²		0,29	

*Melhor rede treinada.

A figura 5.22 (a - d) mostra a Função de Distribuição Cumulativa Empírica (FDCE) e a série temporal entre os valores observados e modelados através das redes neuronais para RESP4M.

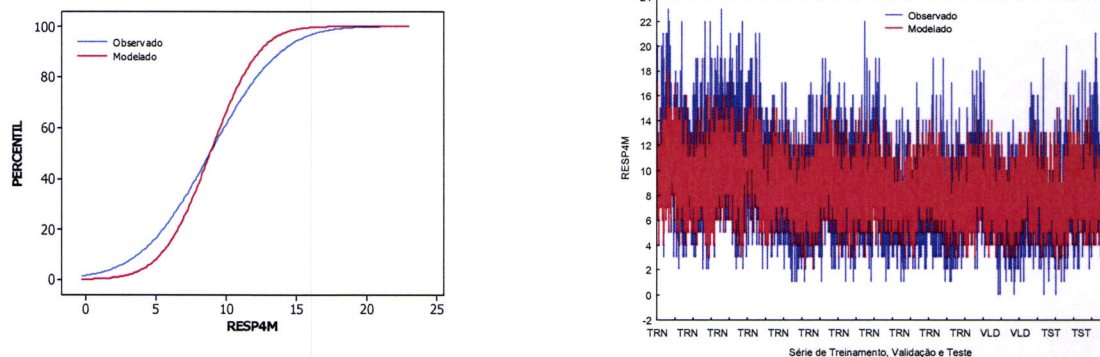
Na figura 5.22c verificou-se que o melhor desempenho da rede neuronal foi através do algoritmo de treinamento Backpropagation (BP) e função de activação tangente hiperbólica (TGH) para RESP4M_TGH_BP. Observa-se através das curvas de ajustes que os valores prognosticados pela rede detectam bem a variabilidade diária e sazonal das doenças respiratórias, porém, tendem a subestimar os valores observados durante a fase de treinamento e sobrestimam na fase de testes.

Na Tabela 5.24 verifica-se os parâmetros da rede e estatística descritiva entre os valores observados e modelados durante o aprendizado realizado pela rede. A rede RESP4M_TGH_BP apresentou o melhor desempenho com arquitetura [62-118-1], taxa de aprendizagem (2), momento (1,75) e peso de $\pm 0,6$. Os parâmetros utilizados para uma melhor calibração da rede demonstraram um comportamento semelhante ao observado em termos da média e desvio padrão, enquanto que foi verificado um erro absoluto médio de 2,2.



*Melhor rede treinada.

Figura 5.22 – Comparação entre os valores observados de RESP4M e os valores modelados pela RNA através da Função de Distribuição Cumulativa Empírica (FDCE) e Série Temporal (ST).



(d) RESP4M – OBSERVADO X MODELADO – RESP4M_TGH_DGC

Figura 5.22 – cont.

Tabela 5.24 – Parâmetros da RNA e Estatística para RESP4M

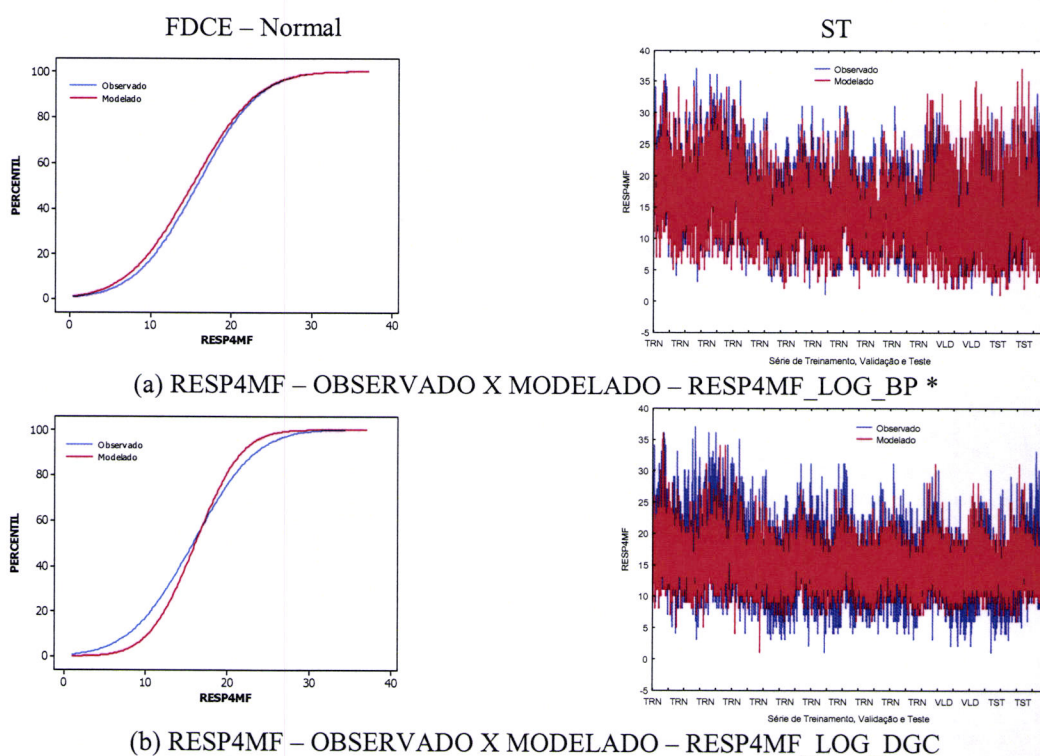
RESP4M LOG BP					
Parâmetros		Estatística	Observado	Modelado	Erro Absoluto
Arquitectura	[62-85-1]	Média	8,9	8,5	2,4
W	± 1,8	Desv.Pad	4	3,9	2,3
TA	0,1	Mín	0	0,3	0,0005
Momento	1,75	Max	23	22,1	16,2
Iterações	500	R	0,65		
Retreinamentos	3	R ²	0,42		
RESP4M LOG DGC					
Parâmetros		Estatística	Observado	Modelado	Erro Absoluto
Arquitectura	[62-85-1]	Média	8,9	8,9	2,4
W	± 0,7	Desv.Pad	4	3	1,9
TA		Mín	0	0,4	0,0005
Momento		Max	23	20	12,5
Iterações	500	R	0,64		
Retreinamentos	3	R ²	0,41		
RESP4M TGH BP *					
Parâmetros		Estatística	Observado	Modelado	Erro Absoluto
Arquitectura	[62-118-1]	Média	8,9	8,7	2,2
W	± 0,6	Desv.Pad	4	3,9	2,3
TA	0,1	Mín	0	0,1	0,0001
Momento	2	Max	23	21,6	16,5
Iterações	500	R	0,68		
Retreinamentos	3	R ²	0,46		
RESP4M TGH DGC					
Parâmetros		Estatística	Observado	Modelado	Erro Absoluto
Arquitectura	[62-118-1]	Média	8,9	8,9	2,6
W	± 1	Desv.Pad	4	2,7	2
TA		Mín	0	2	0,0002
Momento		Max	23	18,4	11,2
Iterações	500	R	0,59		
Retreinamentos	3	R ²	0,35		

*Melhor rede treinada.

A figura 5.23 (a - d) mostra a Função de Distribuição Cumulativa Empírica (FDCE) e a série temporal entre os valores observados e modelados através das redes neuronais para RESP4MF.

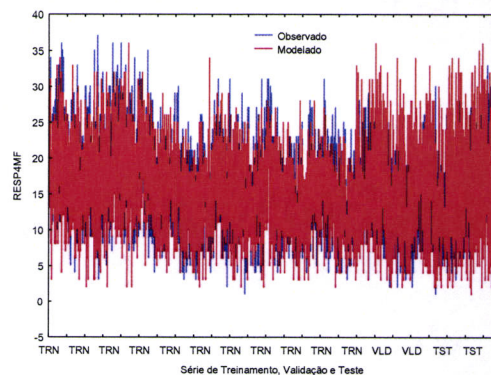
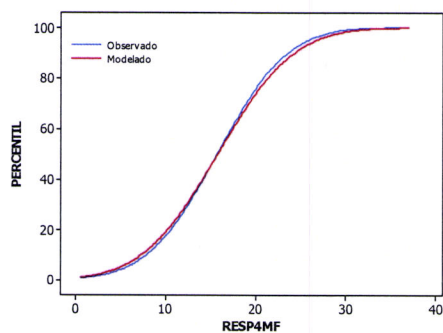
Na figura 5.23c verificou-se que o melhor desempenho da rede neuronal foi através do algoritmo de treinamento Backpropagation (BP) e função de activação logística (LOG) para RESP4MF_LOG_BP. Observa-se através das curvas de ajustes que os valores prognosticados pela rede detectam bem a variabilidade diária e sazonal das doenças respiratórias, porém, tendem a subestimar os valores observados durante a fase de treinamento e sobrestimam na fase de validação e teste. Deve-se ressaltar que a performance da rede melhora sensivelmente o reconhecimento do padrão de RESP4 com a inclusão de ambos os sexos durante o aprendizado.

Na Tabela 5.25 verifica-se os parâmetros da rede e estatística descritiva entre os valores observados e modelados durante o aprendizado realizado pela rede. A rede RESP4MF_LOG_BP apresentou o melhor desempenho com arquitectura [62-138-1], taxa de aprendizagem (0,1), momento (1,75) e peso de $\pm 1,3$. Os parâmetros utilizados para um melhor ajuste da rede demonstraram um comportamento semelhante ao observado em termos da média e desvio padrão, enquanto que foi verificado um erro absoluto médio de 3,1.

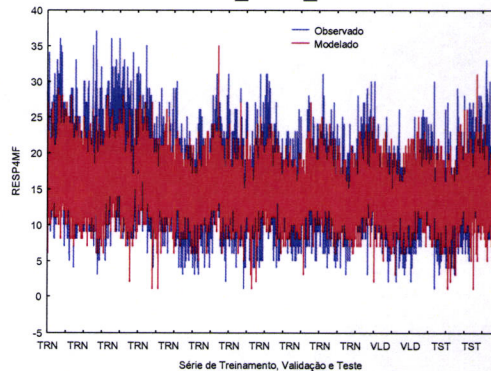
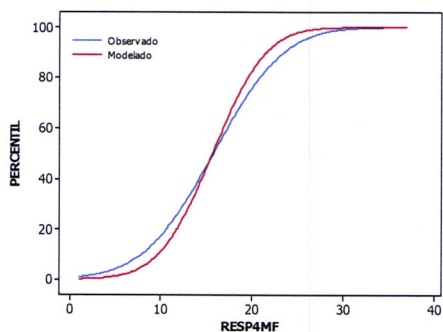


*Melhor rede treinada.

Figura 5.23 – Comparação entre os valores observados de RESP4MF e os valores modelados pela RNA através da Função de Distribuição Cumulativa Empírica (FDCE) e Série Temporal (ST).



(c) RESP4MF – OBSERVADO X MODELADO – RESP4MF_TGH_BP



(d) RESP4MF – OBSERVADO X MODELADO – RESP4MF_TGH_DGC

Figura 5.23 – cont.

Tabela 5.25 – Parâmetros da RNA e Estatística para RESP4MF

RESP4MF LOG BP *					
Parâmetros		Estatística	Observado	Modelado	Erro Absoluto
Arquitetura	[62-127-1]	Média	15,8	15,1	3,1
W	$\pm 1,3$	Desv.Pad	6,1	6,3	3,7
TA	0,1	Min	1	1,5	0,001
Momento	1,75	Max	37	36,7	29,7
Iterações	500	R	0,71		
Retreinamentos	3	R ²	0,51		
RESP4MF LOG DGC					
Parâmetros		Estatística	Observado	Modelado	Erro Absoluto
Arquitetura	[62-127-1]	Média	15,8	16,1	3,4
W	$\pm 3,5$	Desv.Pad	6,1	4,5	2,9
TA		Min	1	1	0,0006
Momento		Max	37	35,8	18,7
Iterações	500	R	0,66		
Retreinamentos	3	R ²	0,43		
RESP4MF TGH BP					
Parâmetros		Estatística	Observado	Modelado	Erro Absoluto
Arquitetura	[62-138-1]	Média	15,8	15,9	3,6
W	$\pm 0,6$	Desv.Pad	6,1	6,6	3,7
TA	0,1	Min	1	1,2	0,001
Momento	1,75	Max	37	36	28,4
Iterações	500	R	0,67		
Retreinamentos	3	R ²	0,45		
RESP4MF TGH DGC					
Parâmetros		Estatística	Observado	Modelado	Erro Absoluto
Arquitetura	[62-138-1]	Média	15,8	15,7	3,6
W	$\pm 1,3$	Desv.Pad	6,1	4,6	2,9
TA		Min	1	1	0,0006

Momento		Max	37	34,6	17,3
Iterações	500	R	0,65		
Retreinamentos	3	R ²	0,43		

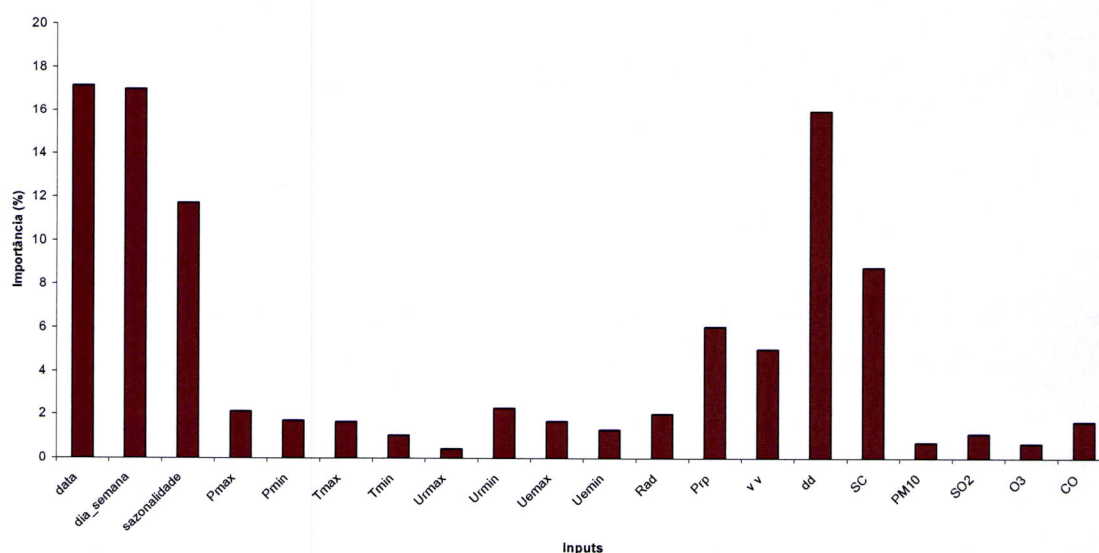
*Melhor rede treinada.

Tabela 5.25 – cont.

5.5.3. Análise de Sensibilidade da RN – Importância dos “inputs”

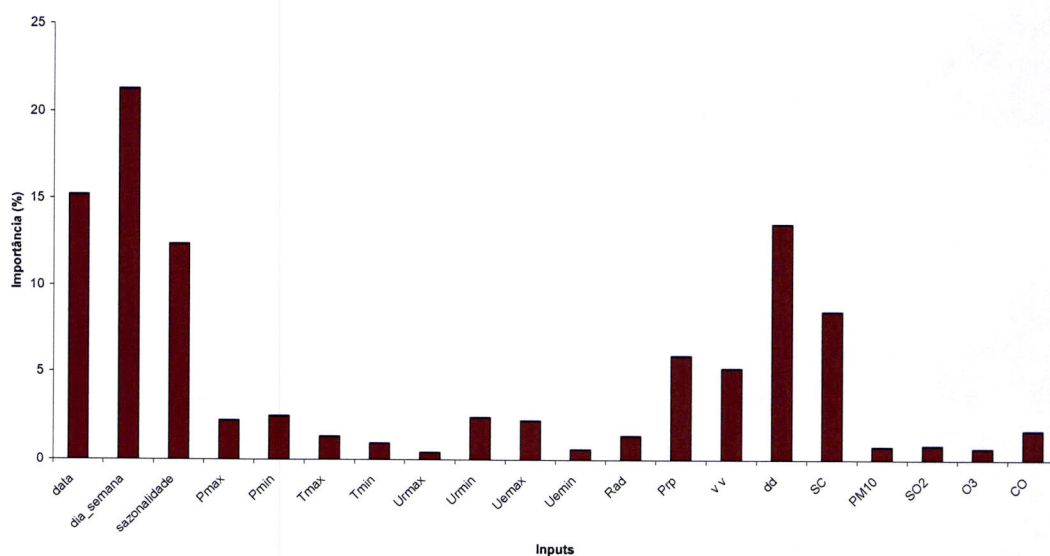
A figura 5.24 (a - c) demonstra a análise de sensibilidade dos “inputs”, para os melhores desempenho da rede para a morbilidade de doenças respiratórias, em adultos maduros, segundo a faixa etária e sexo.

De um modo geral, observou-se que as variáveis de periodicidade denotaram uma importância de 45%, principalmente, a variabilidade semanal (dias_semana – 20%), enquanto que 51% está relacionada com as variáveis meteorológicas, principalmente a direcção do vento (dd – 14 %) e condição sinóptica (SC – 10%). As variáveis poluentes mostraram uma importância em torno de 4 %, principalmente, em relação ao monóxido de carbono (CO).

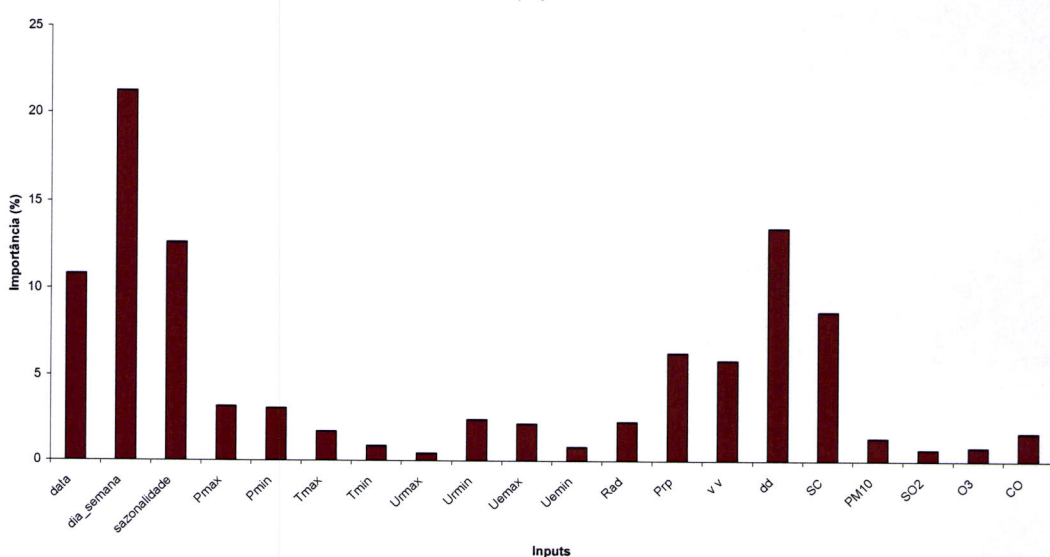


(a)

Figura 5.24 – Análise de Sensibilidade: (a) RESP4F_TGH_BP – [62-95-1]; (b) RESP4M_TGH_BP – [62-118-1] e (c) RESP4MF_TGH_BP – [62-138-1].



(b)



(c)

Figura 5.24 – cont.

5.5.4. Índice de Risco Meteorotrópico (IRM) – Matriz de Confusão

As tabelas 5.26, 5.27 e 5.28 mostram através da matriz de confusão o grau de acerto (valores em azul) do Índice de Risco Meteorotrópico (IRM) entre os valores preditos pelas melhores redes treinadas e os valores observados, durante a fase de aprendizado.

Os valores prognosticados pelas redes (RESP4F_TGH_BP – [62-95-1] e RESP4M_TGH_BP – [62-118-1]) apresentaram um grau de acerto em torno de

72% sobre o valor observado, enquanto que para RESP4MF_TGH_BP – [62-138-1] o grau de acerto foi em torno de 75%.

O IRM 2 ($q_{0.25} < \text{RESP4} \leq q_{0.50}$), faixa com maiores ocorrências de morbidades para RESP1, apresentou uma **Frequência Normal** em relação ao valor esperado pelas afecções respiratórias, enquanto que a classe de **Baixa Frequência** (IRM1) e **Frequência Acima do Normal** (IRM3) tenderam a sobrestimar e subestimar os valores observados, respectivamente. Para o IRM 4 não foi prognosticado nenhuma **Frequência de Alerta**. A informação da matriz de confusão corrobora com a explicação descrita nas análises de Funções de Distribuição Cumulativa Empírica (FDCE).

Tabela 5.26 – Matriz de Confusão para RESP4F_TGH_BP – [62-95-1].

MODELADO	OBSERVADO			
	IRM 1	IRM 2	IRM 3	IRM 4
	$\text{RESP4F} \leq q_{0.25}$	$q_{0.25} < \text{RESP4F} \leq q_{0.50}$	$q_{0.50} < \text{RESP4F} \leq q_{0.75}$	$\text{RESP4F} \geq q_{0.75}$
	≥ 5	[6....9]	[10....17]	≤ 18
≥ 5	963	288	31	0
[6....9]	431	1203	107	0
[10....17]	2	28	116	0
≤ 18	0	2	8	0

Tabela 5.27 – Matriz de Confusão para RESP4M_TGH_BP – [62-118-1].

MODELADO	OBSERVADO			
	IRM 1	IRM 2	IRM 3	IRM 4
	$RESP4M \leq q_{0.25}$	$q_{0.25} < RESP4M \leq q_{0.50}$	$q_{0.50} < RESP4M \leq q_{0.75}$	$RESP4M \geq q_{0.75}$
	≥ 6	[7....11]	[12....22]	≤ 23
≥ 6	923	289	49	0
[7....11]	421	1286	79	0
[12....22]	1	9	112	0
≤ 23	0	3	7	0

Tabela 5.28 – Matriz de Confusão Para RESP4MF_TGH_BP – [62-138-1].

MODELADO	OBSERVADO			
	IRM 1	IRM 2	IRM 3	IRM 4
	$RESP4MF \leq q_{0.25}$	$q_{0.25} < RESP4MF \leq q_{0.50}$	$q_{0.50} < RESP4MF \leq q_{0.75}$	$RESP4MF \geq q_{0.75}$
	≥ 11	[12....20]	[21....36]	≤ 37
≥ 11	953	249	39	0
[12....20]	401	1316	69	0
[21....36]	1	6	115	0
≤ 37	0	4	6	0

5.6. Resultados para RESP5 (M, F e MF)

5.6.1. Análise em Componentes Principais (ACP) e Análise de “Cluster”

Na tabela 5.29, Observou-se a extracção de 4 componentes após a rotação varimax para RESP5M e RESP5MF, com uma variância acumulada (Var. Acum.) em torno de 81%, enquanto que RESP5F apresentou a extracção de apenas 3 componentes, com uma variância acumulada de 70%.

Para a CP1, verificou-se uma variância explicada de aproximadamente 31%, com pesos construtivos para os poluentes SO₂, MP₁₀ e CO e RESP5F, RESP5M e RESP5MF, assim como uma oposição significativa para a velocidade do vento (vv) e pouco significativa para a temperatura mínima (Tmin) e humidade específica mínima (Uemin). Esta componente tem o mesmo comportamento verificado na extracção da CP1 para crianças com morbilidades respiratórias até 13 anos segundo o sexo – RESP1 (M, F e MF).

Na CP2, observou-se uma variância explicada de 22% com pesos positivos para as variáveis Tmin (0.95) e Uemin (0.96) em oposição à RESP5F, RESP5M, RESP5MF, SO₂ e MP₁₀ com pesos pouco significativo. Esta componente indica que o aumento de temperatura e humidade específica do ar favorecem o decréscimo das doenças respiratórias, conforme CP2 de RESP1, RESP2, RESP3 e RESP4 (M, F e MF).

Na CP3, verifica-se uma variância explicada na ordem de 17%, com pesos positivos para a radiação solar (Rad) e ozono (O₃), Conforme verificado para RESP1, RESP2, RESP3 e RESP4 (M, F e MF).

Na CP4, observa-se uma variância explicada de 11% com pesos positivos para a velocidade do vento (vv) em oposição aos poluentes MP₁₀ e O₃ com pesos significativos. Esta componente pode ser explicada pela dispersão dos poluentes provocado pela duração e intensidade do vento, conforme denotado pela extracção do factor 4 na RESP1 (M, F e MF).

Tabela 5.29 – Pesos das componentes principais após a rotação varimax para todo o período de estudo (01/01/1993 a 31/12/2002) – RESP5

Variáveis	CP1	CP2	CP3	CP4
Tmin	0,089914	0,955160	0,086912	
Uemin	0,060127	0,960565	0,030006	
Rad	0,006088	0,055900	0,712142	
vv	0,364393	0,064753	-0,438906	
PM10	0,876700	-0,130702	0,242304	
SO2	0,783401	-0,258002	0,203470	
O3	0,103724	0,148997	0,844409	
CO	0,831760	0,047881	-0,188572	
RESP5_F	0,593708	0,071934	-0,367655	
Var. Expl. (%)	30,97884	21,55844	16,68120	
Var. Acum. (%)	30,97884	52,53728	69,21849	
Variáveis	CP1	CP2	CP3	CP4
Tmin	-0,059969	0,948840	0,128759	-0,073677
Uemin	-0,062556	0,966844	-0,023727	0,062577
Rad	0,047414	0,031150	0,918500	-0,132965
vv	-0,156541	0,014148	0,028636	0,920112
PM10	0,815732	-0,142638	0,173811	0,360723
SO2	0,740391	-0,281804	0,216347	0,223947
O3	-0,004161	0,159848	0,707742	0,483562
CO	0,864203	0,022806	-0,114098	-0,007964
RESP5_M	0,675445	-0,107238	-0,176001	0,239889
Var. Expl. (%)	30,87831	21,56882	16,91967	11,22649
Var. Acum. (%)	30,87831	52,44713	69,36680	80,59329
Variáveis	CP1	CP2	CP3	CP4
Tmin	-0,059272	0,950937	0,127975	0,077233
Uemin	-0,062700	0,967487	-0,025670	-0,061889
Rad	0,039825	0,032299	0,918255	0,129057
vv	-0,142250	0,020841	0,027121	0,932963
PM10	0,814530	-0,140092	0,182352	0,358138
SO2	0,746677	-0,276547	0,223575	0,214033
O3	-0,007988	0,159706	0,701938	-0,487604
CO	0,865801	0,026501	-0,104038	0,011988
RESP5_MF	0,715298	-0,072602	-0,224805	0,178837
Var. Expl. (%)	31,46287	21,55339	17,17663	11,14109
Var. Acum. (%)	31,46287	53,01626	70,19289	81,33399

Na figura 5.25 foi verificado o mesmo padrão de conexão dos grupos obtidos pela AC quando comparado com a RESP1, RESP3 e RESP4 (M, F e MF). Ressalta-se que os grupos descritos pela AC apresentaram o mesmo modelo demonstrado nas extracções das componentes na ACP.

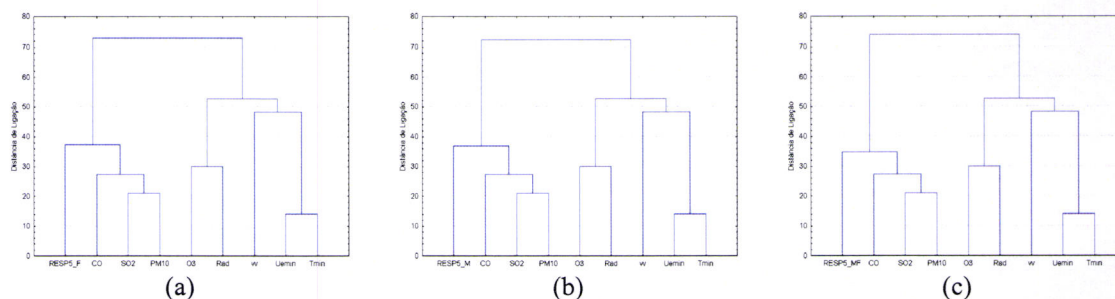


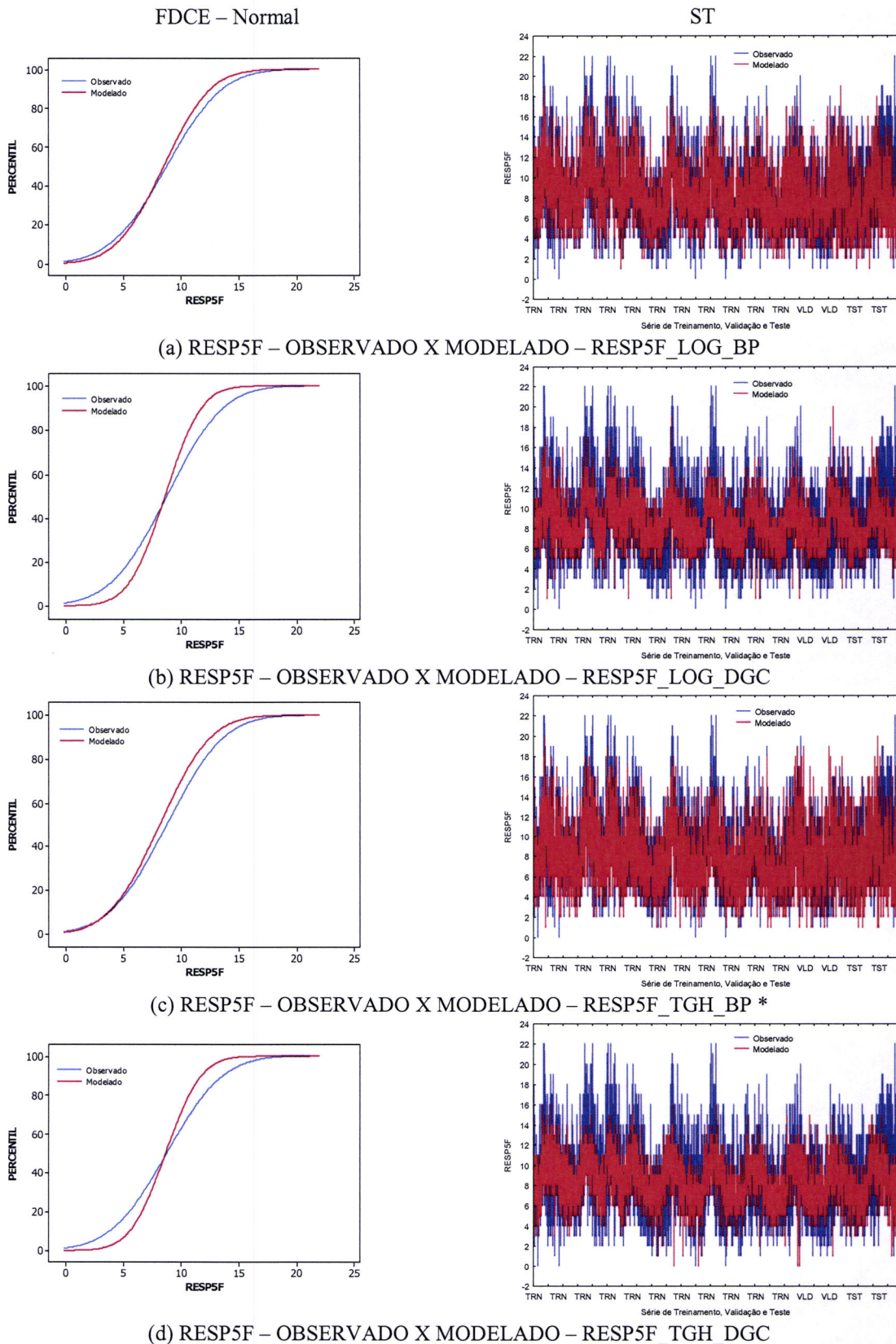
Figura 5.25 – Dendrograma de árvore através do método de Ward e distâncias euclidianas para todo o período de estudo: RESP5

5.6.2. Prognósticos Gerados pela RN – FDCE (Observado X Modelado)

A figura 5.26 (a - d) mostra a Função de Distribuição Cumulativa Empírica (FDCE) e a série temporal entre os valores observados e modelados através das redes neuronais para RESP5F.

Na figura 5.26c, verificou-se que o melhor desempenho da rede neuronal foi através do algoritmo de treinamento Backpropagation (BP) e função de activação tangente hiperbólica (TGH) para RESP5F_TGH_BP. Observa-se através das curvas de ajustes que os valores prognosticados pela rede detectam bem a variabilidade diária e sazonal das doenças respiratórias, porém, tendem a subestimar os valores observados durante a fase de treinamento e sobrestimam na fase de validação e teste.

Na Tabela 5.30, verifica-se os parâmetros da rede e estatística descritiva entre os valores observados e modelados durante o aprendizado realizado pela rede. A rede RESP5F_TGH_BP apresentou o melhor desempenho com arquitectura [62-118-1], taxa de aprendizagem (0,1), momento (4,0) e peso de $\pm 0,5$. Os parâmetros utilizados para um melhor ajuste da rede demonstraram um comportamento semelhante ao observado em termos da média e desvio padrão, enquanto que foi verificado um erro absoluto médio de 2,4.



*Melhor rede treinada.

Figura 5.26 – Comparação entre os valores observados de RESP5F e os valores modelados pela RNA através da Função de Distribuição Cumulativa Empírica (FDCE) e Série Temporal (ST).

Tabela 5.30 – Parâmetros da RNA e Estatística para RESP5F

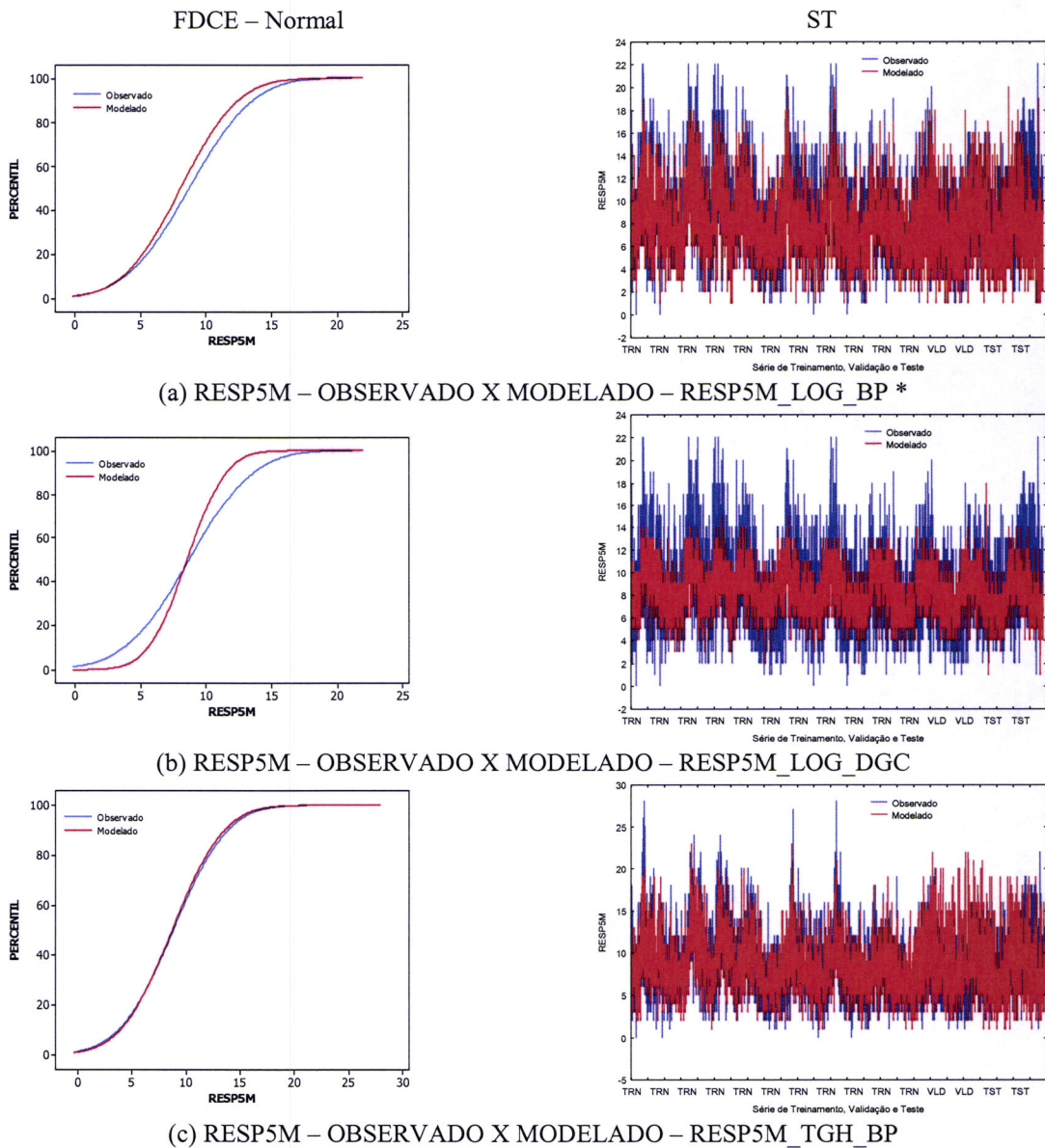
RESP5F LOG BP					
Parâmetros		Estatística	Observado	Modelado	Erro Absoluto
Arquitectura	[62-115-1]	Média	8,8	8,5	2,3
W	$\pm 1,3$	Desv.Pad	3,8	3,2	2
TA	2	Min	0	0,8	0,0008
Momento	1	Max	22	19,4	17
Iterações	500	R		0,62	
Retreinamentos	3	R ²		0,39	
RESP5F LOG DGC					
Parâmetros		Estatística	Observado	Modelado	Erro Absoluto
Arquitectura	[62-115-1]	Média	8,8	8,6	2,6
W	± 4	Desv.Pad	3,8	2,5	2,1
TA		Min	0	0,6	0,00008
Momento		Max	22	19,6	13,6
Iterações	500	R		0,50	
Retreinamentos	3	R ²		0,25	
RESP5F TGH BP *					
Parâmetros		Estatística	Observado	Modelado	Erro Absoluto
Arquitectura	[62-118-1]	Média	8,8	8,1	2,4
W	$\pm 0,5$	Desv.Pad	3,8	3,4	2,2
TA	0,1	Min	0	0,6	0,0003
Momento	4	Max	22	20,2	15,8
Iterações	500	R		0,63	
Retreinamentos	3	R ²		0,39	
RESP5F TGH DGC					
Parâmetros		Estatística	Observado	Modelado	Erro Absoluto
Arquitectura	[62-118-1]	Média	8,8	18	4,1
W	$\pm 1,5$	Desv.Pad	3,8	4,9	3,2
TA		Min	0	5,5	0,0006
Momento		Max	22	33,9	20,5
Iterações	500	R		0,62	
Retreinamentos	3	R ²		0,38	

*Melhor rede treinada.

A figura 5.27 (a - d) mostra a Função de Distribuição Cumulativa Empírica (FDCE) e a série temporal entre os valores observados e modelados através das redes neurais para RESP5M.

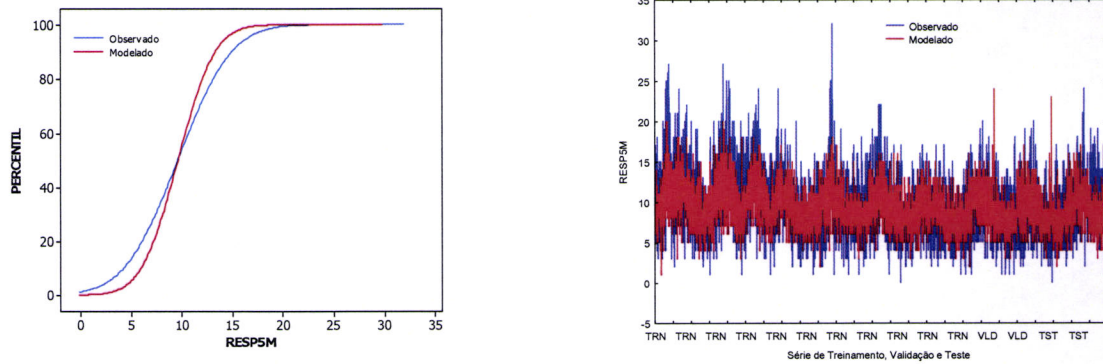
Na figura 5.27a, verificou-se que o melhor desempenho da rede neuronal foi através do algoritmo de treinamento Backpropagation (BP) e função de activação logística (LOG) para RESP5M_LOG_BP. Observa-se através das curvas de ajustes que os valores prognosticados pela rede detectam bem a variabilidade diária e sazonal das doenças respiratórias, porém, tendem a subestimar os valores observados durante a fase de treinamento e sobrestimam na fase de validação e teste.

Na Tabela 5.31 verifica-se os parâmetros da rede e estatística descritiva entre os valores observados e modelados durante o aprendizado realizado pela rede. A rede RESP5M_LOG_BP apresentou o melhor desempenho com arquitectura [62-99-1], taxa de aprendizagem (2,0), momento (1,0) e peso de $\pm 1,3$. Os parâmetros utilizados para um melhor ajuste da rede demonstraram um comportamento semelhante ao observado em termos da média e desvio padrão, enquanto que foi verificado um erro absoluto médio de 2,4.



*Melhor rede treinada.

Figura 5.27 – Comparação entre os valores observados de RESP5M e os valores modelados pela RNA através da Função de Distribuição Cumulativa Empírica (FDCE) e Série Temporal (ST).



(d) RESP5M – OBSERVADO X MODELADO – RESP5M_TGH_DGC

Figura 5.27 – cont.

Tabela 5.31 – Parâmetros da RNA e Estatística para RESP5M

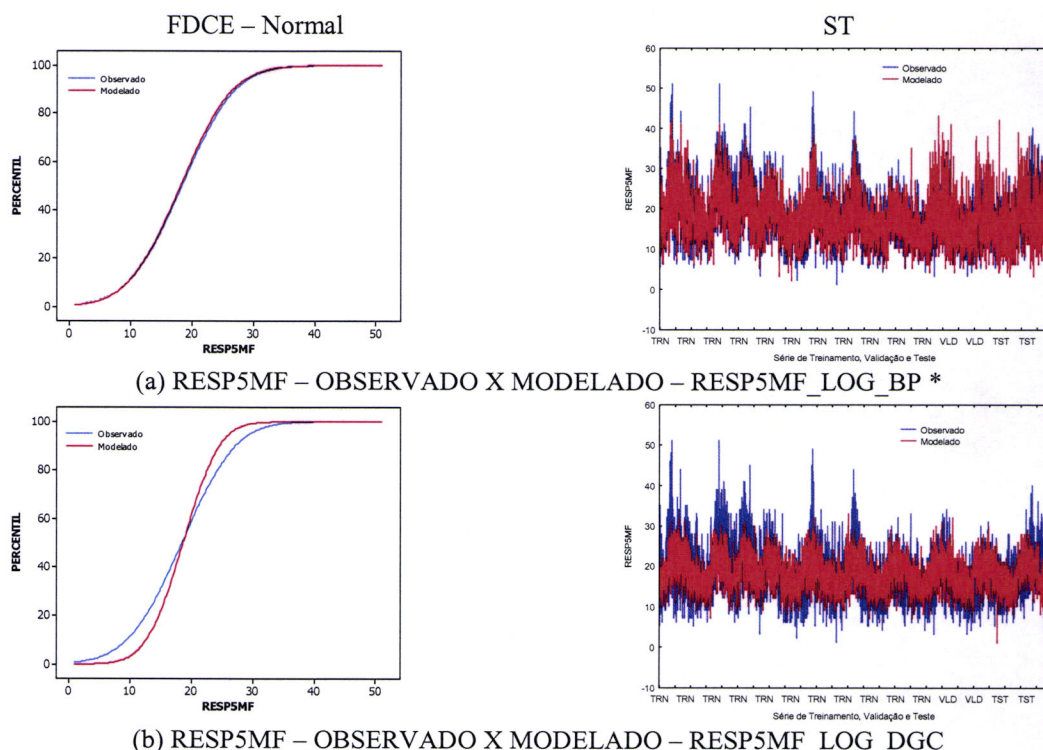
RESP5M LOG BP *					
Parâmetros		Estatística	Observado	Modelado	Erro Absoluto
Arquitetura	[62-99-1]	Média	8,8	8,1	2,4
W	± 1,3	Desv.Pad	3,8	3,5	2,2
TA	1	Min	0	0,8	0,0004
Momento	2	Max	22	20,3	15
Iterações	500	R	0,62		
Retreinamentos	3	R ²	0,39		
RESP5M LOG DGC					
Parâmetros		Estatística	Observado	Modelado	Erro Absoluto
Arquitetura	[62-99-1]	Média	8,8	8,6	2,7
W	± 4	Desv.Pad	3,8	2,3	2,1
TA		Min	0	0,6	0,0004
Momento		Max	22	18,3	14,7
Iterações	500	R	0,49		
Retreinamentos	3	R ²	0,24		
RESP5M TGH BP					
Parâmetros		Estatística	Observado	Modelado	Erro Absoluto
Arquitetura	[62-108-1]	Média	8,8	8,8	2,6
W	± 0,5	Desv.Pad	3,8	3,7	2,4
TA	0,1	Min	0	0,8	0,00002
Momento	1,75	Max	22	22,7	17,6
Iterações	500	R	0,58		
Retreinamentos	3	R ²	0,33		
RESP5M TGH DGC					
Parâmetros		Estatística	Observado	Modelado	Erro Absoluto
Arquitetura	[62-108-1]	Média	8,8	9,7	2,8
W	± 1,1	Desv.Pad	3,8	2,9	2,2
TA		Min	0	1,4	0,001
Momento		Max	22	24,4	16,2
Iterações	500	R	0,54		
Retreinamentos	3	R ²	0,29		

*Melhor rede treinada.

A figura 5.28 (a - d) mostra a Função de Distribuição Cumulativa Empírica (FDCE) e a série temporal entre os valores observados e modelados através das redes neuronais para RESP5MF.

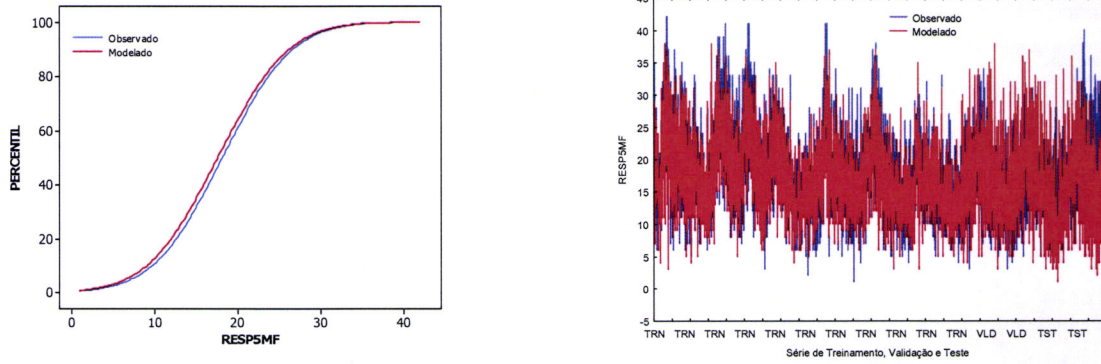
Na figura 5.28a, verificou-se que o melhor desempenho da rede neuronal foi através do algoritmo de treinamento Backpropagation (BP) e função de activação logística (LOG) para RESP5MF_LOG_BP. Observa-se através das curvas de ajustes que os valores prognosticados pela rede detectam bem a variabilidade diária e sazonal das doenças respiratórias, porém, tendem a subestimar os valores observados durante a fase de treinamento e sobrestimam na fase de validação e teste. Deve-se ressaltar que a performance da rede melhora sensivelmente o reconhecimento do padrão de RESP5 com a inclusão de ambos os sexos durante o aprendizado.

Na Tabela 5.31, verifica-se os parâmetros da rede e estatística descritiva entre os valores observados e modelados durante o aprendizado realizado pela rede. A rede RESP5MF_LOG_BP apresentou o melhor desempenho com arquitectura [62-71-1], taxa de aprendizagem (0,1), momento (1,75) e peso de ± 2 . Os parâmetros utilizados para um melhor ajuste da rede demonstraram um comportamento semelhante ao observado em termos da média e desvio padrão, enquanto que foi verificado um erro absoluto médio de 3,8.

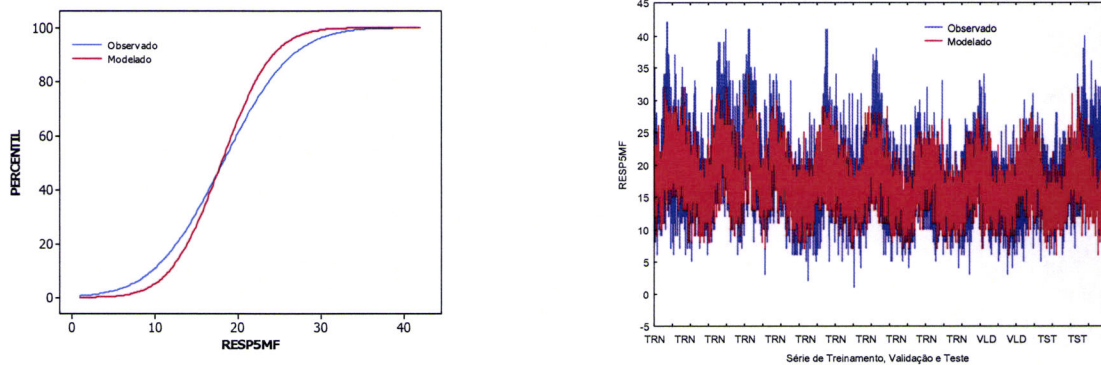


*Melhor rede treinada.

Figura 5.28 – Comparação entre os valores observados de RESP5MF e os valores modelados pela RNA através da Função de Distribuição Cumulativa Empírica (FDCE) e Série Temporal (ST).



(c) RESP5MF – OBSERVADO X MODELADO – RESP5MF_TGH_BP



(d) RESP5MF – OBSERVADO X MODELADO – RESP5MF_TGH_DGC

Figura 5.28 – cont.

Tabela 5.32 – Parâmetros da RNA e Estatística para RESP5MF

RESP5MF LOG BP *					
Parâmetros		Estatística	Observado	Modelado	Erro Absoluto
Arquitetura	[62-71-1]	Média	18,5	18,2	3,8
W	± 2	Desv.Pad	6,9	6,7	3,7
TA	0,9	Min	1	2	0,0006
Momento	1,3	Max	51	42,9	29,2
Iterações	500	R	0,70		
Retreinamentos	3	R ²	0,49		
RESP5MF LOG DGC					
Parâmetros		Estatística	Observado	Modelado	Erro Absoluto
Arquitetura	[62-71-1]	Média	18,5	18,7	4,5
W	± 6	Desv.Pad	6,9	4,6	3,6
TA		Min	1	1,2	0,001
Momento		Max	51	33,1	27,5
Iterações	500	R	0,56		
Retreinamentos	3	R ²	0,32		
RESP5MF TGH BP					
Parâmetros		Estatística	Observado	Modelado	Erro Absoluto
Arquitetura	[62-130-1]	Média	18,5	17,7	3,9
W	± 0,5	Desv.Pad	6,9	6,7	3,8
TA	0,1	Min	1	1,5	0,001
Momento	2	Max	51	38,1	24,1
Iterações	500	R	0,67		
Retreinamentos	3	R ²	0,45		
RESP5MF TGH DGC					
Parâmetros		Estatística	Observado	Modelado	Erro Absoluto
Arquitetura	[62-130-1]	Média	18,5	18	4,1
W	± 0,5	Desv.Pad	6,9	4,9	3,2
TA		Min	1	5,5	0,0006

Momento		Max	51	33,9	20,5
Iterações	500	R		0,62	
Retreinamentos	3	R ²		0,38	

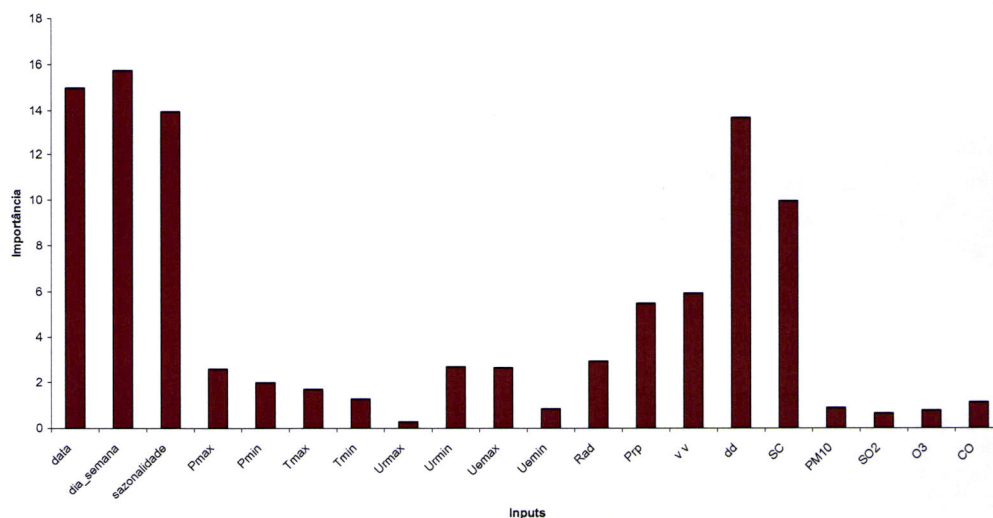
*Melhor rede treinada.

Tabela 5.32 – cont.

5.6.3. Análise de Sensibilidade da RN – Importância dos “inputs”

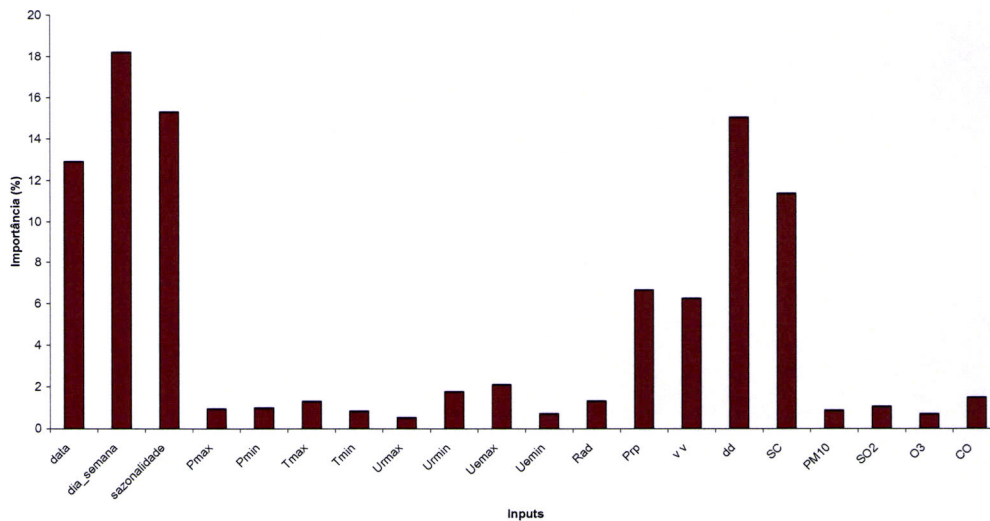
A figura 5.29 (a - c) demonstra a análise de sensibilidade dos “inputs”, para os melhores desempenho da rede para a morbilidade de doenças respiratórias, em idosos com ou mais de 65 anos, segundo a faixa etária e sexo.

De um modo geral, observou-se que as variáveis de periodicidade denotaram uma importância de 45%, principalmente, a variabilidade semanal (dias_semana – 18%), enquanto que 51% está relacionada com as variáveis meteorológicas, principalmente a direcção do vento (dd – 14 %) e condição sinóptica (SC – 10%). As variáveis poluentes mostraram uma importância em torno de 4 %, principalmente, em relação ao monóxido de carbono (CO).

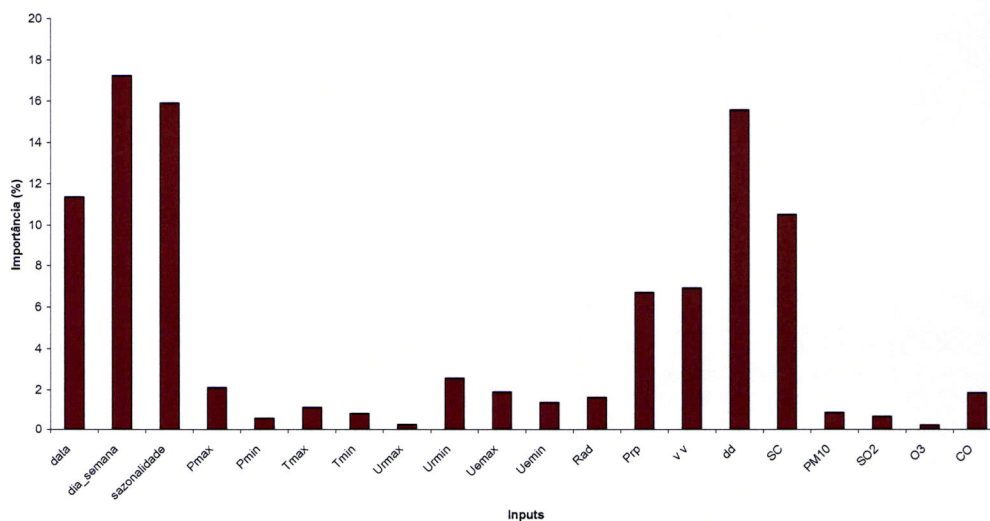


(a)

Figura 5.29 – Análise de Sensibilidade: (a) RESP5F_TGH_BP – [62-115-1]; (b) RESP5M_LOG_BP – [62-99-1] e (c) RESP5MF_LOG_BP – [62-71-1]



(b)



(c)

Figura 5.29 – cont.

5.6.4. Índice de Risco Meteorotrópico (IRM) – Matriz de Confusão

As tabelas 5.33, 5.34 e 5.35 mostram através da matriz de confusão o grau de acerto (valores em azul) do Índice de Risco Meteorotrópico (IRM) entre os valores preditos pelas melhores redes treinadas e os valores observados, durante a fase de aprendizado.

Os valores prognosticados pelas redes (RESP5F_TGH_BP – [62-115-1] e RESP5M_LOG_BP – [62-99-1]) apresentaram um grau de acerto em torno de 68% sobre o valor observado, enquanto que para RESP5MF_LOG_BP – [62-71-1] o grau de acerto foi em torno de 74%.

O IRM 2 ($q_{0.25} < \text{RESP5} \leq q_{0.50}$), faixa com maiores ocorrências de morbidades para RESP1, apresentou uma **Frequência Normal** em relação ao valor esperado pelas afecções respiratórias, enquanto que a classe de **Baixa Frequência** (IRM1) e **Frequência Acima do Normal** (IRM3) tenderam a sobrestimar e subestimar os valores observados, respectivamente. Para o IRM 4 não foi prognosticado nenhuma **Frequência de Alerta**. A informação da matriz de confusão corrobora com a explicação descrita nas análises de Funções de Distribuição Cumulativa Empírica (FDCE).

Tabela 5.33 – Matriz de Confusão para RESP5F_TGH_BP – [62-115-1].

MODELADO	OBSERVADO			
	IRM 1	IRM 2	IRM 3	IRM 4
	$\text{RESP5F} \leq q_{0.25}$	$q_{0.25} < \text{RESP5F} \leq q_{0.50}$	$q_{0.50} < \text{RESP5F} \leq q_{0.75}$	$\text{RESP5F} \geq q_{0.75}$
	≥ 6	[7....11]	[12....21]	≤ 22
≥ 6	642	136	35	0
[7....11]	476	1280	51	0
[12....21]	90	240	219	0
≤ 22	0	2	8	0

Tabela 5.34 – Matriz de Confusão para RESP5M_LOG_BP – [62-99-1].

MODELADO	OBSERVADO			
	IRM 1	IRM 2	IRM 3	IRM 4
	$RESP5M \leq q_{0.25}$	$q_{0.25} < RESP5M \leq q_{0.50}$	$q_{0.50} < RESP5M \leq q_{0.75}$	$RESP5M \geq q_{0.75}$
	≥ 6	[7....11]	[12....21]	≤ 22
≥ 6	653	106	31	0
[7....11]	531	1358	67	0
[12....21]	64	210	149	0
≤ 22	0	4	6	0

Tabela 5.35 – Matriz de Confusão Para RESP5MF_LOG_BP – [62-71-1].

MODELADO	OBSERVADO			
	IRM 1	IRM 2	IRM 3	IRM 4
	$RESP5MF \leq q_{0.25}$	$q_{0.25} < RESP5MF \leq q_{0.50}$	$q_{0.50} < RESP5MF \leq q_{0.75}$	$RESP5MF \geq q_{0.75}$
	≥ 13	[14....18]	[19....50]	≤ 51
≥ 13	665	132	10	0
[14....18]	430	1407	131	0
[19....50]	16	171	198	0
≤ 51	0	1	18	0

5.7. Resultados para RESP6 (M, F e MF)

5.7.1. Análise em Componentes Principais (ACP) e Análise de “Cluster”

Na tabela 5.36, Observou-se a extracção de 4 componentes após a rotação varimax para RESP6M, RESP6F e RESP6MF, com uma variância acumulada em torno de 81%.

Na CP1, verificou-se uma variância explicada de aproximadamente 32%, com pesos positivos para os poluentes SO₂, MP₁₀ e CO e RESP6F (0.71), RESP6M (0.72) e RESP6MF (0.74), assim como pesos negativos significativos para a velocidade do vento (vv) e pouco significativos para a temperatura mínima (Tmin) e humidade específica mínima (Uemin).

Na CP2, observou-se uma variância explicada de 22% com pesos positivos para as variáveis Tmin (0.95) e Uemin (0.96) em oposição à RESP3 (M, F e MF), SO₂ e MP₁₀. Esta componente explica a mesma variabilidade encontrada no factor 2 de RESP1, RESP2 e RESP3, RESP4 e RESP5 (M, F e MF).

Na CP3, verifica-se uma variância explicada na ordem de 17%, com pesos positivos para a radiação solar (Rad) e ozono (O₃), apresentando a mesma variabilidade explicada quando comparada com RESP1, RESP2 e RESP3, RESP4 e RESP5 (M, F e MF).

Na CP4 foi observado o mesmo padrão denotado pela extracção da CP4 em RESP4 (M, F e MF).

Tabela 5.36 – Pesos das componentes principais após a rotação varimax para todo o período de estudo (01/01/1993 a 31/12/2002) – RESP6

Variáveis	CP1	CP2	CP3	CP4
Tmin	-0,054710	0,951070	0,126432	0,079881
Uemin	-0,066061	0,965542	-0,023080	-0,065259
Rad	0,044526	0,029952	0,919491	0,129854
vv	-0,149190	0,021658	0,022136	0,934456
PM10	0,820968	-0,142374	0,170667	-0,346120
SO2	0,757298	-0,276786	0,210231	-0,198338
O3	0,001262	0,158888	0,709611	-0,489827
CO	0,854265	0,019996	-0,116217	0,016927
RESP6_F	0,713184	-0,117191	-0,158586	0,136011
Var. Expl. (%)	31,49599	21,64697	16,71412	11,12604
Var. Acum. (%)	31,49599	53,14296	69,85708	80,98313
Variáveis	CP1	CP2	CP3	CP4
Tmin	-0,051541	0,948072	0,128266	0,078424
Uemin	-0,058802	0,964849	-0,022428	-0,064320
Rad	0,044414	0,029556	0,919327	0,130429
vv	-0,146963	0,020533	0,022264	0,934536
PM10	0,822390	-0,149439	0,170859	-0,345897
SO2	0,746843	-0,287966	0,213307	-0,205333
O3	-0,002377	0,160332	0,710404	-0,491608
CO	0,858226	0,010581	-0,115685	0,017724
RESP6_M	0,723219	-0,170562	-0,144422	0,132946
Var. Expl. (%)	31,41376	21,79473	16,90946	11,13730
Var. Acum. (%)	31,41376	53,20849	70,11795	81,25525
Variáveis	CP1	CP2	CP3	CP4
Tmin	-0,051011	0,949853	0,127241	0,079615
Uemin	-0,059962	0,965489	-0,023090	-0,064245
Rad	0,042054	0,029672	0,919652	0,129314
vv	-0,144740	0,022122	0,021996	0,935926
PM10	0,820987	-0,147530	0,172752	-0,346626
SO2	0,750529	-0,283979	0,213947	-0,202735
O3	-0,003672	0,159488	0,709642	-0,492201
CO	0,858049	0,013843	-0,113689	0,017730
RESP6_MF	0,737149	-0,144245	-0,151995	0,123833
Var. Expl. (%)	31,66303	21,73202	16,92800	11,13243
Var. Acum. (%)	31,66303	53,39505	70,32305	81,45548

Na figura 5.15 foi verificado o mesmo padrão de conexão dos grupos obtidos pela AC quando comparado com a RESP1, RESP3, RESP4 e RESP5 (M, F e MF). Ressalta-se que os grupos descritos pela AC apresentaram o mesmo modelo demonstrado nas extracções das componentes através da ACP.

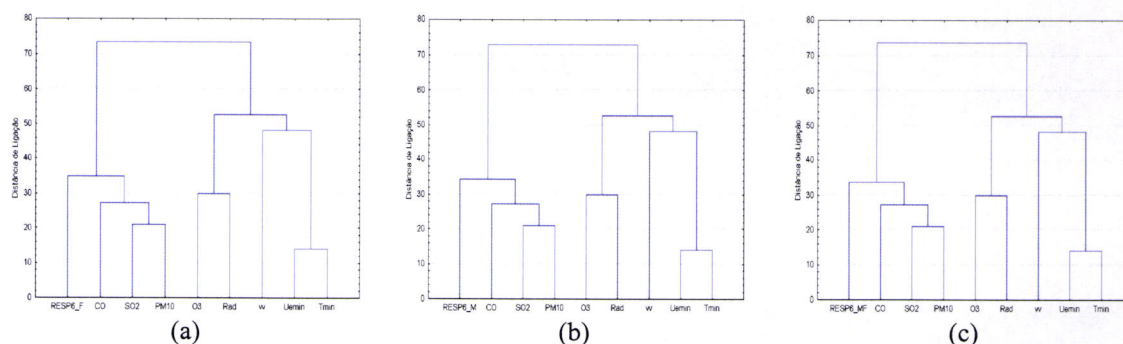


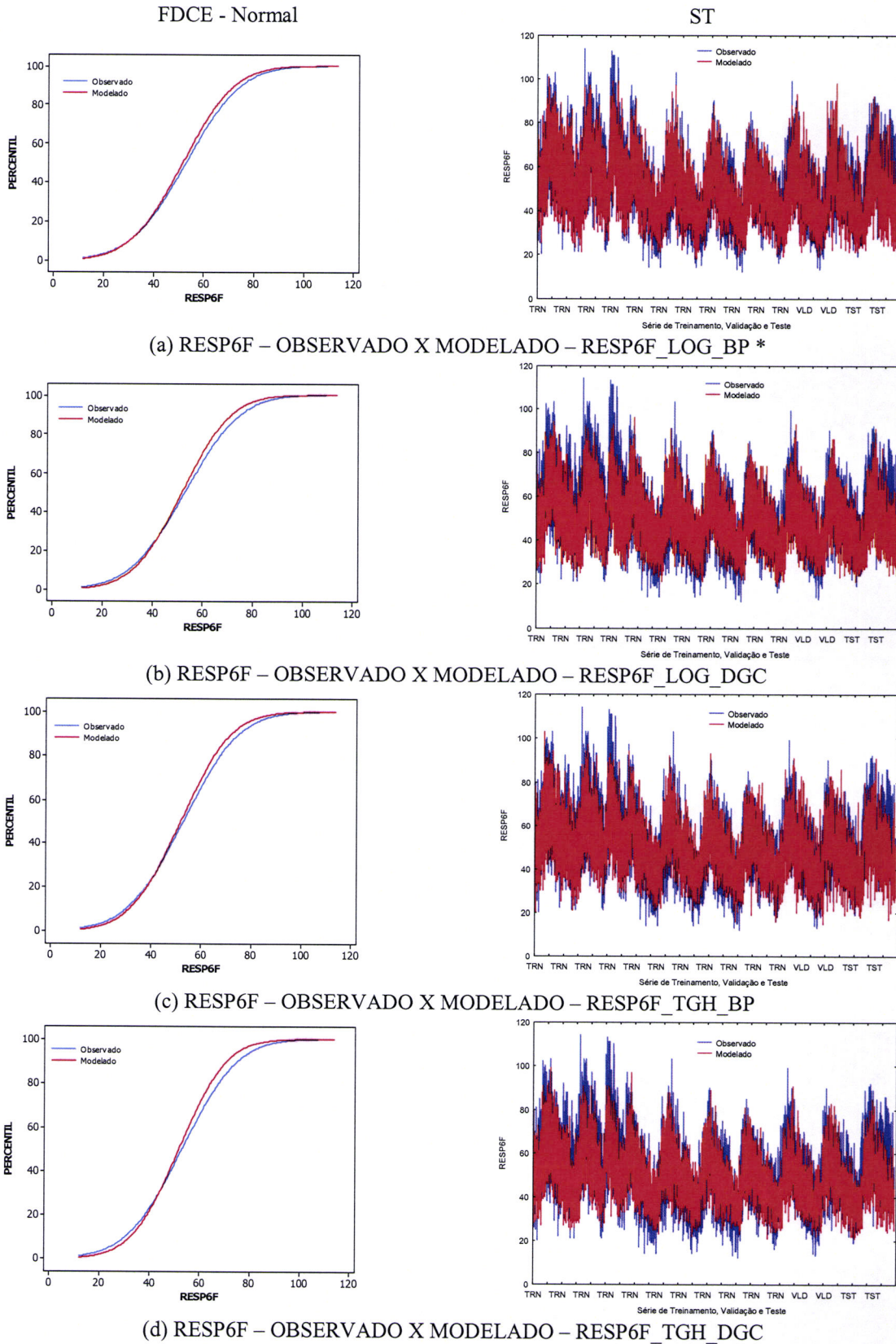
Figura 5.30 – Dendrograma de árvore através do método de Ward e distâncias euclidianas para todo o período de estudo: RESP6

5.7.2. Prognósticos Gerados pela RN – FDCE (Observado X Modelado)

A figura 5.31 (a - d) mostra a Função de Distribuição Cumulativa Empírica (FDCE) e a série temporal entre os valores observados e modelados através das redes neurais para RESP6F.

Na figura 5.31a, verificou-se que o melhor desempenho da rede neuronal foi através do algoritmo de treinamento Backpropagation (BP) e função de activação logística (LOG) para RESP6F_LOG_BP. Observa-se através das curvas de ajustes que os valores prognosticados pela rede detectam bem a variabilidade diária e sazonal das doenças respiratórias, porém, tendem a subestimar os valores observados durante a fase de treinamento, validação e teste. Deve-se ressaltar que a performance da rede melhora sensivelmente o reconhecimento do padrão de morbidade respiratória com a inclusão de todas as faixas etárias do sexo feminino durante o aprendizado.

Na Tabela 5.37, verifica-se os parâmetros da rede e estatística descritiva entre os valores observados e modelados durante o aprendizado realizado pela rede. A rede RESP6F_LOG_BP apresentou o melhor desempenho com arquitectura [62-115-1], taxa de aprendizagem (0,9), momento (1) e peso de $\pm 1,3$. Os parâmetros utilizados para um melhor ajuste da rede demonstraram um comportamento semelhante ao observado em termos da média e desvio padrão, enquanto que foi verificado um erro absoluto médio de 7,3.



*Melhor rede treinada.

Figura 5.31 – Comparação entre os valores observados de RESP6F e os valores modelados pela RNA através da Função de Distribuição Cumulativa Empírica (FDCE) e Série Temporal (ST).

Tabela 5.37 – Parâmetros da RNA e Estatística para RESP6F

RESP6F LOG BP *					
Parâmetros		Estatística	Observado	Modelado	Erro Absoluto
Arquitectura	[62-115-1]	Média	53,2	51,9	7,3
W	$\pm 1,3$	Desv.Pad	17,5	16,5	6,4
TA	0,9	Min	12	18,1	0,001
Momento	1	Max	114	100,6	47,9
Iterações	500	R		0,84	
Retreinamentos	3	R ²		0,71	
RESP6F LOG DGC					
Parâmetros		Estatística	Observado	Modelado	Erro Absoluto
Arquitectura	[62-115-1]	Média	53,2	52	8,8
W	$\pm 0,8$	Desv.Pad	17,5	15,5	6,9
TA		Min	12	18,8	0,01
Momento		Max	114	96,2	50,4
Iterações	500	R		0,78	
Retreinamentos	3	R ²		0,61	
RESP6F TGH BP					
Parâmetros		Estatística	Observado	Modelado	Erro Absoluto
Arquitectura	[62-83-1]	Média	53,2	52	8,2
W	$\pm 0,8$	Desv.Pad	17,5	15,9	6,9
TA	0,1	Min	12	15,2	0,007
Momento	1	Max	114	102,6	56,9
Iterações	500	R		0,80	
Retreinamentos	3	R ²		0,64	
RESP6F TGH DGC					
Parâmetros		Estatística	Observado	Modelado	Erro Absoluto
Arquitectura	[62-83-1]	Média	53,2	51,8	8,5
W	$\pm 0,6$	Desv.Pad	17,5	15	6,7
TA		Min	12	20,8	0,0009
Momento		Max	114	98,6	43,9
Iterações	500	R		0,79	
Retreinamentos	3	R ²		0,63	

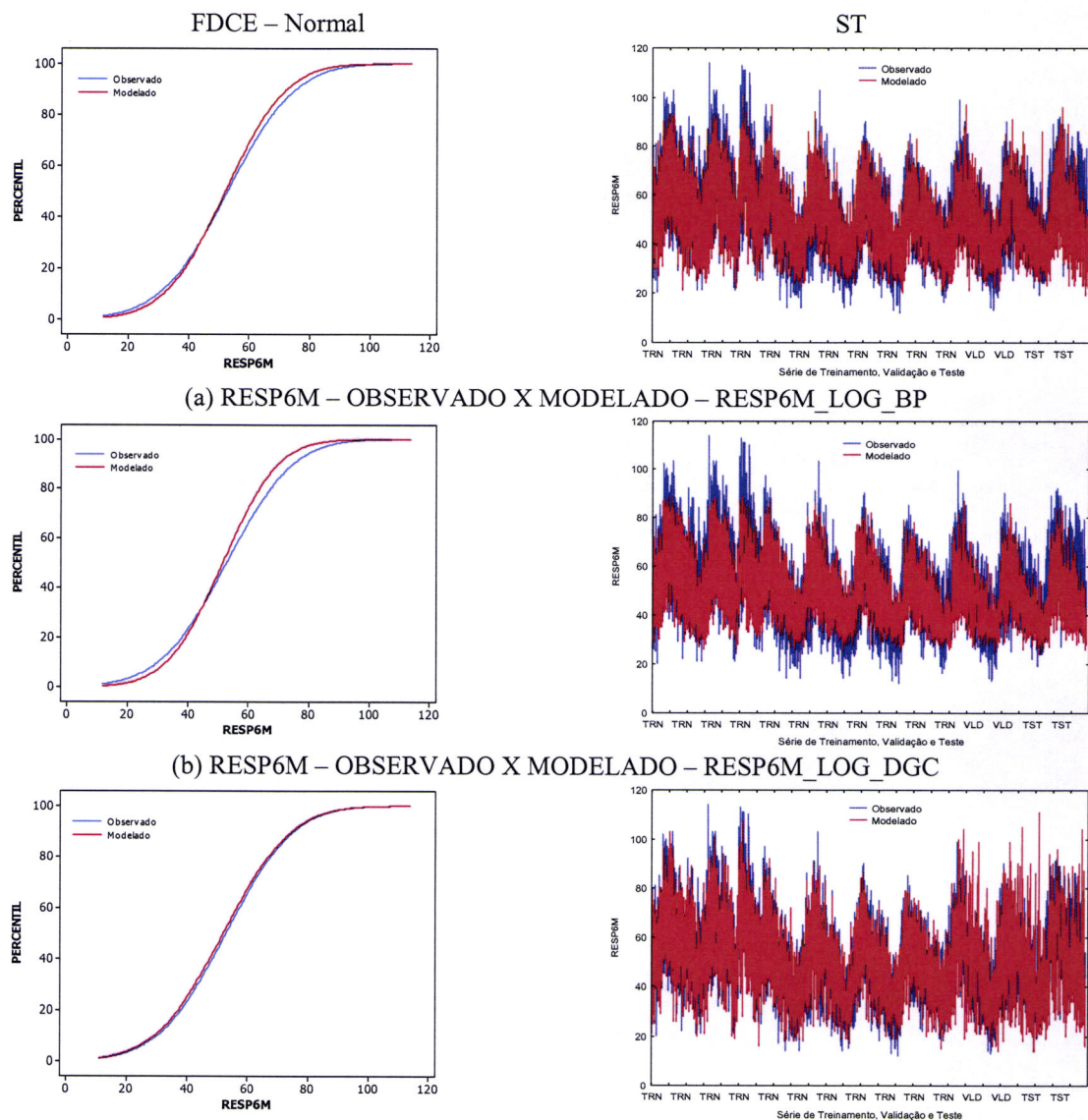
*Melhor rede treinada.

A figura 5.32 (a - d) mostra a Função de Distribuição Cumulativa Empírica (FDCE) e a série temporal entre os valores observados e modelados através das redes neuronais para RESP6M.

Na figura 5.32c, verificou-se que o melhor desempenho da rede neuronal foi através do algoritmo de treinamento Backpropagation (BP) e função de activação tangente hiperbólica (TGH) para RESP6M_TGH_BP. Observa-se através das curvas de ajustes que os valores prognosticados pela rede detectam bem a variabilidade diária e sazonal das doenças respiratórias, porém, tendem a subestimar os valores observados durante a fase de treinamento e sobrestimam na fase de validação e teste. Deve-se ressaltar que a performance da rede melhora sensivelmente o reconhecimento do padrão das admissões hospitalares

por afecções respiratórias, com a inclusão de todas as faixas etárias do sexo masculino durante o aprendizado.

Na Tabela 5.38 verifica-se os parâmetros da rede e estatística descritiva entre os valores observados e modelados durante o aprendizado realizado pela rede. A rede RESP6M_TGH_BP apresentou o melhor desempenho com arquitetura [62-100-1], taxa de aprendizagem (0,1), momento (1,0) e peso de $\pm 0,8$. Os parâmetros utilizados para um melhor ajuste da rede demonstraram um comportamento semelhante ao observado em termos da média e desvio padrão, enquanto que foi verificado um erro absoluto médio de 8,1.



*Melhor rede treinada.

Figura 5.32 – Comparação entre os valores observados de RESP6M e os valores modelados pela RNA através da Função de Distribuição Cumulativa Empírica (FDCE) e Série Temporal (ST).

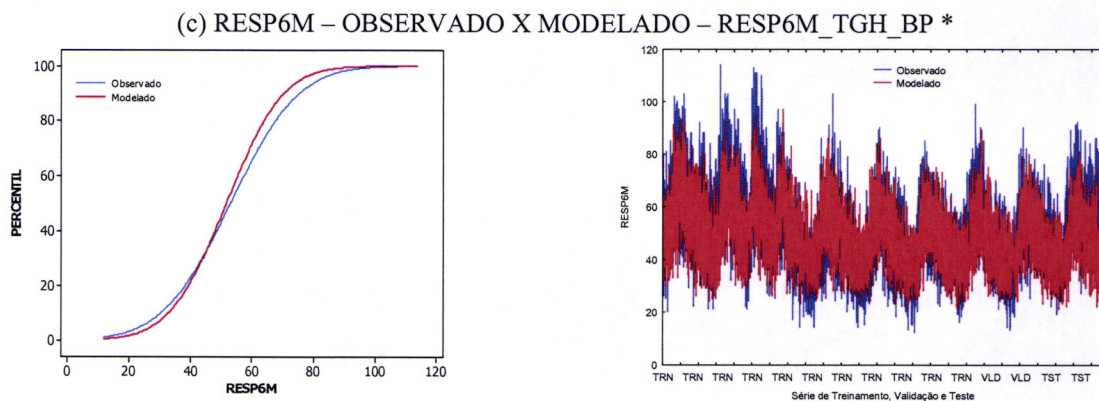


Figura 5.32 – cont.

Tabela 5.38 – Parâmetros da RNA e Estatística para RESP6M

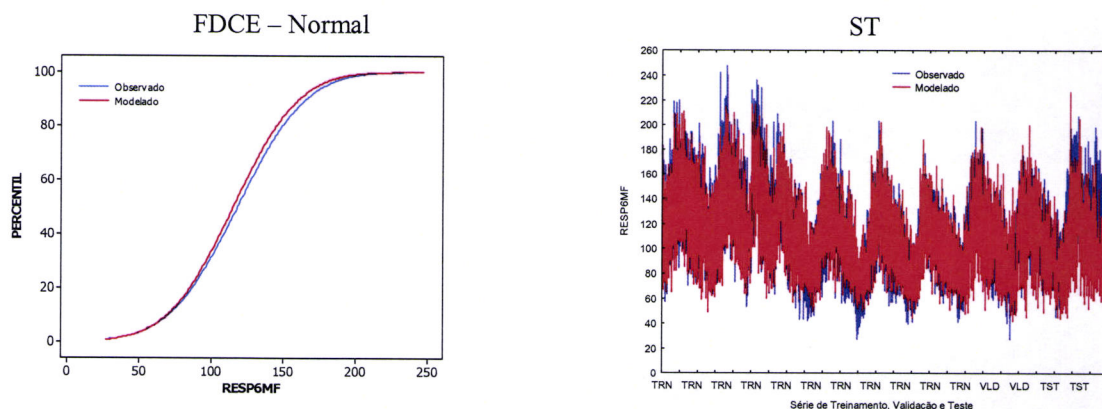
RESP6M LOG BP					
Parâmetros		Estatística	Observado	Modelado	Erro Absoluto
Arquitectura	[62-21-1]	Média	53,2	52,3	8,2
W	± 2,5	Desv.Pad	17,5	15,7	6,7
TA	6	Min	12	19,1	0,004
Momento	1,3	Max	114	100,8	52,2
Iterações	500	R	0,80		
Retreinamentos	3	R ²	0,65		
RESP6M LOG DGC					
Parâmetros		Estatística	Observado	Modelado	Erro Absoluto
Arquitectura	[62-21-1]	Média	53,2	51,7	8,3
W	± 1,5	Desv.Pad	17,5	14,5	6,5
TA		Min	12	24,5	0,002
Momento		Max	114	89,3	39,6
Iterações	500	R	0,80		
Retreinamentos	3	R ²	0,64		
RESP6M TGH BP *					
Parâmetros		Estatística	Observado	Modelado	Erro Absoluto
Arquitectura	[62-100-1]	Média	53,2	52,2	8,1
W	± 0,8	Desv.Pad	17,5	17,7	8
TA	0,1	Min	12	13,7	0,001
Momento	1	Max	114	110,5	64,5
Iterações	500	R	0,79		
Retreinamentos	3	R ²	0,63		
RESP6M TGH DGC					
Parâmetros		Estatística	Observado	Modelado	Erro Absoluto
Arquitectura	[62-100-1]	Média	53,2	51,8	8,8
W	± 0,5	Desv.Pad	17,5	14,6	6,9
TA		Min	12	21,2	0,001
Momento		Max	114	97,2	44,2
Iterações	500	R	0,77		
Retreinamentos	3	R ²	0,60		

*Melhor rede treinada.

A figura 5.33 (a - d) mostra a Função de Distribuição Cumulativa Empírica (FDCE) e a série temporal entre os valores observados e modelados através das redes neurais para RESP6MF.

Na figura 5.33c, verificou-se que o melhor desempenho da rede neuronal foi através do algoritmo de treinamento Backpropagation (BP) e função de activação tangente hiperbólica (TGH) para RESP6MF_TGH_BP. Observa-se através das curvas de ajustes que os valores prognosticados pela rede detectam bem a variabilidade diária e sazonal das doenças respiratórias, porém, tendem a subestimar os valores observados durante a fase de treinamento e sobrestimam na fase de validação e teste. Deve-se ressaltar que a performance da rede melhora sensivelmente o reconhecimento do padrão da morbidade para todas as faixas etárias com a inclusão de ambos os sexos durante o aprendizado.

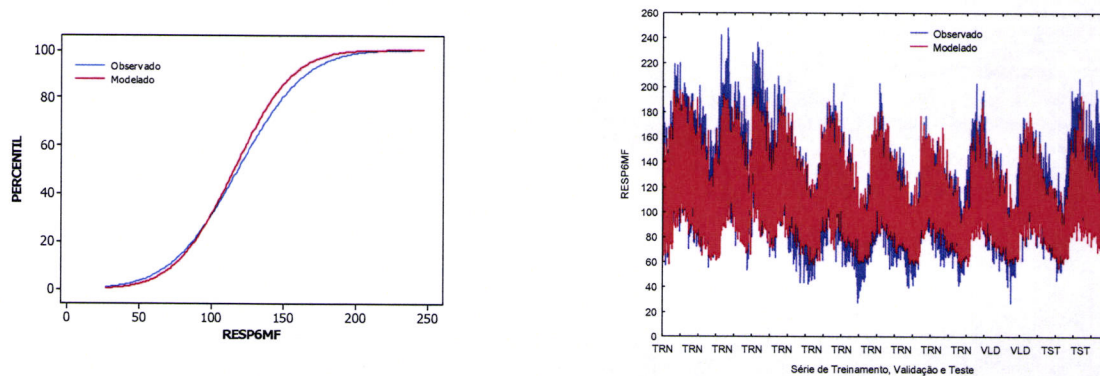
Na Tabela 5.39, verifica-se os parâmetros da rede e estatística descritiva entre os valores observados e modelados durante o aprendizado realizado pela rede. A rede RESP6MF_TGH_BP apresentou o melhor desempenho com arquitetura [62-126-1], taxa de aprendizagem (0,1), Momento (1,0) e peso de $\pm 0,3$. Os parâmetros utilizados para um melhor ajuste da rede demonstraram um comportamento semelhante ao observado em termos da média e desvio padrão, enquanto que foi verificado um erro absoluto médio de 15,7.



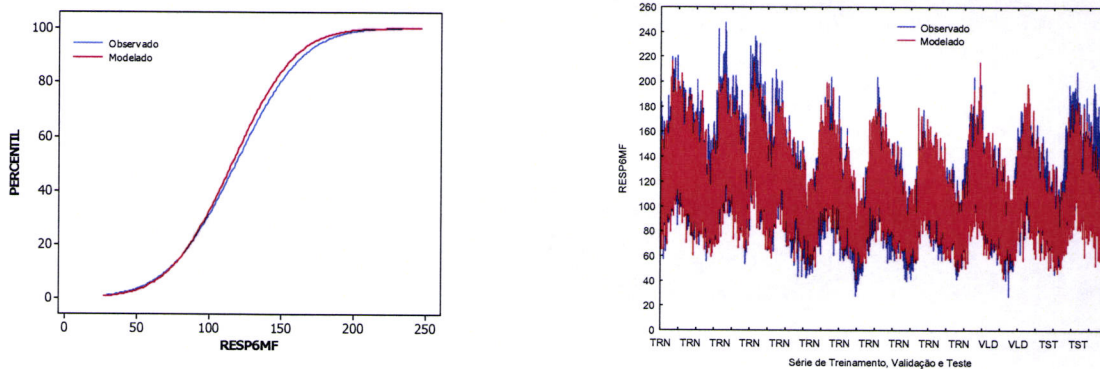
(a) RESP6MF – OBSERVADO X MODELADO – RESP6MF_LOG_BP

*Melhor rede treinada.

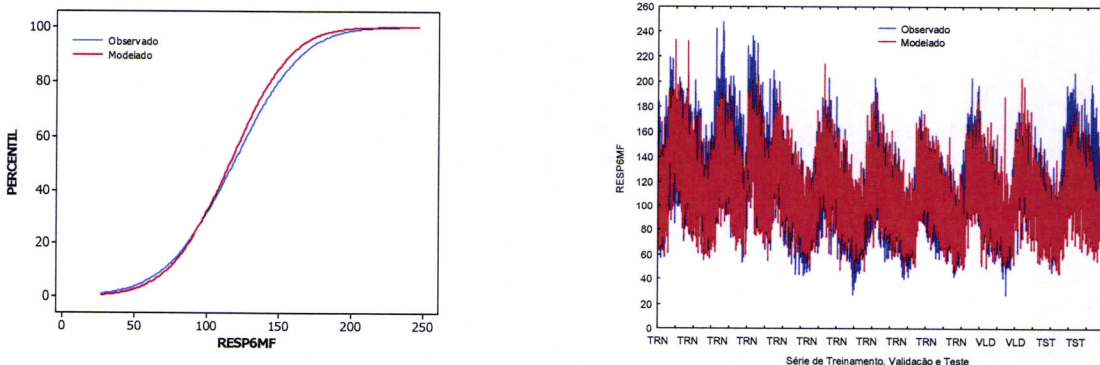
Figura 5.33 – Comparação entre os valores observados de RESP6MF e os valores modelados pela RNA através da Função de Distribuição Cumulativa Empírica (FDCE) e Série Temporal (ST).



(b) RESP6MF – OBSERVADO X MODELADO – RESP6MF_LOG_DGC



(c) RESP6MF – OBSERVADO X MODELADO – RESP6MF_TGH_BP *



(d) RESP6MF – OBSERVADO X MODELADO – RESP6MF_TGH_DGC

*Melhor rede treinada.

Figura 5.33 – cont.

Tabela 5.39 – Parâmetros da RNA e Estatística para RESP6MF

RESP6MF LOG BP					
Parâmetros		Estatística	Observado	Modelado	Erro Absoluto
Arquitetura	[62-31-1]	Média	118,6	115,3	15,4
W	± 3	Desv.Pad	37,1	35,5	13,8
TA	0,1	Min	27	42,4	0,001
Momento	1,1	Max	247	227	96,1
Iterações	500	R	0,84		
Retreinamentos	3	R ²	0,71		
RESP6MF LOG DGC					
Parâmetros		Estatística	Observado	Modelado	Erro Absoluto
Arquitetura	[62-31-1]	Média	118,6	115,9	16,5
W	± 1,3	Desv.Pad	37,1	32,4	13,1
TA		Min	27	55,1	0,004
Momento		Max	247	197,3	87,4
Iterações	500	R	0,83		

Retreinamentos	3	R^2	0,68		
RESP6MF TGH BP *					
Parâmetros		Estatística	Observado	Modelado	Erro Absoluto
Arquitectura	[62-126-1]	Média	118,6	116	15,7
W	$\pm 0,3$	Desv.Pad	37,1	34,6	12,6
TA	0,1	Mín	27	40,9	0,01
Momento	1	Max	247	216,1	87,9
Iterações	500	R	0,85		
Retreinamentos	3	R^2	0,72		
RESP6MF TGH DGC					
Parâmetros		Estatística	Observado	Modelado	Erro Absoluto
Arquitectura	[62-126-1]	Média	118,6	115,8	17,5
W	$\pm 0,8$	Desv.Pad	37,1	33,1	13,7
TA		Mín	27	44,3	0,01
Momento		Max	247	233,4	88,6
Iterações	500	R	0,80		
Retreinamentos	3	R^2	0,65		

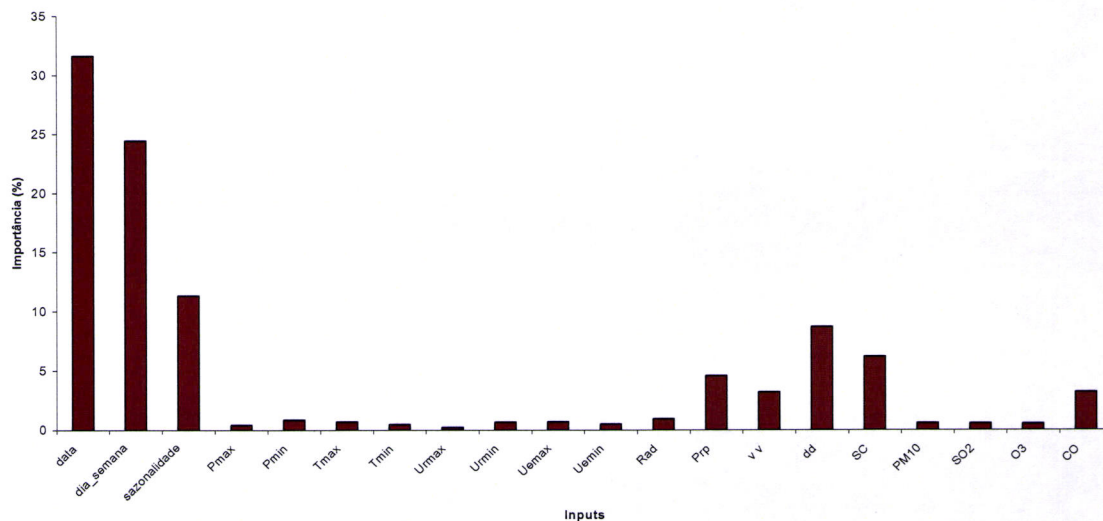
* Melhor rede treinada.

Tabela 5.39 – cont.

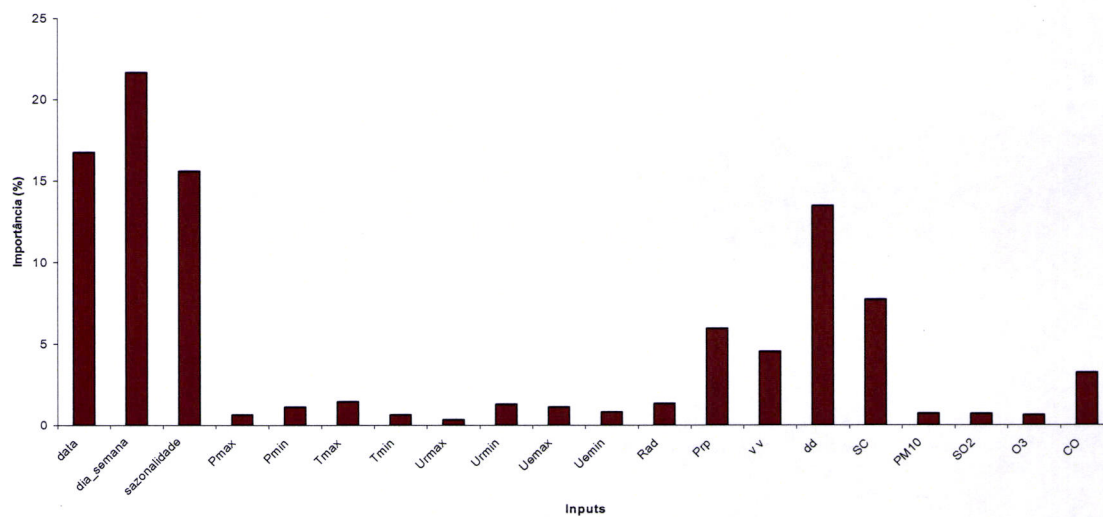
5.7.3. Análise de Sensibilidade da RN – Importância dos “inputs”

A figura 5.34 (a - c) demonstra a análise de sensibilidade dos “inputs”, para os melhores desempenho da rede para a morbilidade de doenças respiratórias envolvendo todas as idades, segundo a faixa etária e sexo.

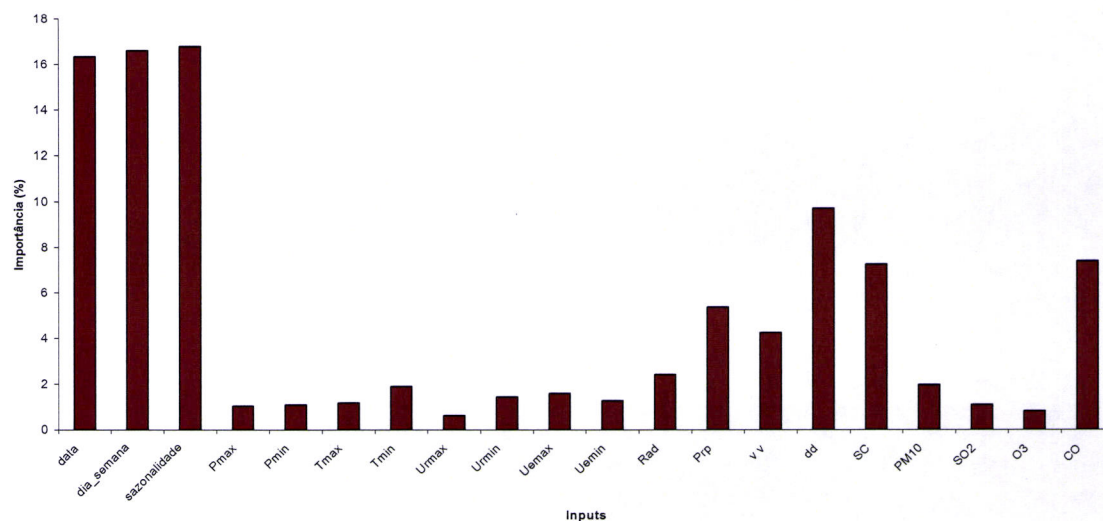
De um modo geral, observou-se que as variáveis de periodicidade denotaram uma importância de 55%, principalmente, a variabilidade semanal (dias_semana – 20%), enquanto que 40% está relacionada com as variáveis meteorológicas, principalmente a direcção do vento (dd – 14 %) e condição sinóptica (SC – 10%). As variáveis poluentes mostraram uma importância em torno de 5 %, principalmente, em relação ao monóxido de carbono (CO).



(a)



(b)



(c)

Figura 5.34: Análise de Sensibilidade: (a) RESP6F_LOG_BP – [62-115-1]; (b) RESP6M_TGH_BP – [62-100-1] e (c) RESP6MF_LOG_BP – [62-31-1]

5.7.4. Índice de Risco Meteorotrópico (IRM) – Matriz de Confusão

As tabelas 5.40, 5.41 e 5.42 mostram através da matriz de confusão o grau de acerto (valores em azul) do Índice de Risco Meteorotrópico (IRM) entre os valores preditos pelas melhores redes treinadas e os valores observados, durante a fase de aprendizado.

Os valores prognosticados pelas redes (RESP6F_LOG_BP – [62-115-1] e RESP6M_TGH_BP – [62-100-1]) apresentaram um grau de acerto em torno de 80% sobre o valor observado, enquanto que para RESP6MF_TGH_BP – [62-31-1] o grau de acerto foi em torno de 83%.

O IRM 2 ($q_{0.25} < \text{RESP6} \leq q_{0.50}$), faixa com maiores ocorrências de morbidades para RESP1, apresentou uma **Frequência Normal** em relação ao valor esperado pelas afecções respiratórias, enquanto que a classe de **Baixa Frequência** (IRM1) e **Frequência Acima do Normal** (IRM3) tenderam a sobrestimar e subestimar os valores observados, respectivamente. Para o IRM 4 não foi prognosticado nenhuma **Frequência de Alerta**. A informação da matriz de confusão corrobora com a explicação descrita nas análises de Funções de Distribuição Cumulativa Empírica (FDCE).

Tabela 5.40 – Matriz de Confusão para RESP6F_LOG_BP – [62-115-1].

MODELADO	OBSERVADO			
	IRM 1	IRM 2	IRM 3	IRM 4
	$\text{RESP6F} \leq q_{0.25}$	$q_{0.25} < \text{RESP6F} \leq q_{0.50}$	$q_{0.50} < \text{RESP6F} \leq q_{0.75}$	$\text{RESP6F} \geq q_{0.75}$
	≥ 40	[41....64]	[65....113]	≤ 114
≥ 40	967	140	15	0
[41....64]	210	1305	14	0
[65....113]	6	99	377	0
≤ 114	0	8	38	0

Tabela 5.41 – Matriz de Confusão para RESP6M_TGH_BP – [62-100-1].

MODELADO	OBSERVADO			
	IRM 1	IRM 2	IRM 3	IRM 4
	$RESP6M \leq q_{0.25}$	$q_{0.25} < RESP6M \leq q_{0.50}$	$q_{0.50} < RESP6M \leq q_{0.75}$	$RESP6M \geq q_{0.75}$
	≥ 40	[41....64]	[65....113]	≤ 114
≥ 40	913	166	25	0
[41....64]	199	1316	22	0
[65....113]	22	141	327	0
≤ 114	0	9	39	0

Tabela 5.42 – Matriz de Confusão Para RESP6MF_TGH_BP – [62-31-1].

MODELADO	OBSERVADO			
	IRM 1	IRM 2	IRM 3	IRM 4
	$RESP6MF \leq q_{0.25}$	$q_{0.25} < RESP6MF \leq q_{0.50}$	$q_{0.50} < RESP6MF \leq q_{0.75}$	$RESP6MF \geq q_{0.75}$
	≥ 91	[92....143]	[144....246]	≤ 247
≥ 91	960	106	17	0
[92....143]	139	1381	19	0
[144....246]	22	141	357	0
≤ 247	0	7	30	0

CAPÍTULO VI – Conclusões

Baseado nos resultados discutidos no capítulo anterior foram obtidas as seguintes conclusões:

1. Através dos métodos estatísticos multivariados (ACP e AC), pode-se verificar que a variabilidade das variáveis envolvidas durante as análises apresentaram uma explicação bastante plausível em relação as morbilidades respiratórias.

- ✓ A extracção da primeira componente elucida uma resposta das afecções respiratórias, devido ao aumento dos níveis das concentrações dos poluentes em oposição a velocidade do vento seguido de um decréscimo da temperatura do ar. Assim como a variabilidade explicada pela segunda componente, explica uma diminuição das doenças respiratórias devido a associação construtiva das variáveis meteorológicas (temperatura e humidade específica mínima do ar) em oposição a morbilidade. A terceira componente é responsável pela formação do ozono, devido a presença da radiação solar, que é coerente do ponto de vista meteorológico. É importante ressaltar uma significativa associação da morbilidade verificada na quarta componente, devido a acção do vento, neste caso, apesar do vento actuar como um agente dispersivo dos poluentes, provavelmente, transporta a poluição para um outro lugar menos poluído.
- ✓ As análises de "Cluster" apresentaram o mesmo padrão demonstrado pelas extracções das componentes, em relação as doenças do aparelho respiratório, para quase todas as idades e sexo. Contudo, não apresentou uma boa descrição na formação do grupo para adolescente.

2. O emprego das redes neuronais para as previsões das doenças respiratórias, em estudos do tipo ecológicos.

- ✓ Deve-se ressaltar que o algoritmo “BackPropagation” foi o que melhor representou os padrões das doenças respiratórias durante a fase de treino da rede.
- ✓ Através das funções de distribuição cumulativa empíricas, para as melhores redes treinadas, verificou-se que as curvas de ajustes dos valores prognosticados pela rede detectaram bem a variabilidade diária e sazonal das doenças respiratórias, porém, tendem a subestimar os valores observados durante a fase de treino e sobrestimam na fase testada, i. e, a performance da rede torna-se instável para os valores extremos.
- ✓ A inclusão de mais uma camada escondida poderia melhorar a performance do aprendizado durante a fase de treinamento, assim como uma série temporal mais extensa poderia dar mais robustez nas saídas da rede.
- ✓ Os parâmetros utilizados para a melhor calibração da rede apresentaram um comportamento semelhante ao observado em termos da média e desvio padrão.
- ✓ De um modo geral, observou-se que as variáveis de periodicidade denotaram uma maior sensibilidade durante a fase de aprendizado, em torno de 45%, principalmente, a variabilidade semanal (dias_semana – 18%), enquanto que 55% estava relacionada com as variáveis meteorológicas, principalmente a direcção do vento (dd – 14 %) e condição sinóptica (SC – 10%). As variáveis poluentes mostraram uma importância em torno de 5 %, sendo mais sensível ao monóxido de carbono (CO).
- ✓ Através da matriz de confusão verificou-se que os valores preditos pela rede mostraram um grau de acerto em torno de 70% para as morbidades respiratórias do sexo masculino e feminino, conforme a faixa etária. A rede quando incluía ambos os sexos de uma determinada idade, geralmente, aumentava o grau de acerto para 75%.
- ✓ As previsões da morbidade respiratórias quando envolvia todo grupo de faixa etária e sexo, RESP6 (F, M e MF); a melhor rede treinada mostrou

uma performance mais robusta e com um grau de acerto em torno de 82% quando comparado com os valores observados.

- ✓ O Índice de Risco Meteorotrópico – IRM 2 ($q_{0.25} < \text{RESP} \leq q_{0.50}$), apresentou as maiores ocorrências de morbidades respiratórias, observando uma **Frequência Normal** em relação ao valor esperado pelas afecções respiratórias, enquanto que a classe de **Baixa Frequência** (IRM1) e **Frequência Acima do Normal** (IRM3) tenderam a sobrestimar e subestimar os valores observados, respectivamente. A **Frequência de Alerta** (IRM 4) não foi prognosticada sobre os valores observados.

Apesar dos problemas referenciados acima, pode-se concluir que a metodologia desenvolvida neste trabalho, para a elaboração do Índice de Risco Meteorotrópico, apresentaram resultados satisfatórios nos prognósticos da variabilidade das doenças respiratórias, segundo a classificação etária e o sexo na Cidade de São Paulo.

↓ Considerações Finais

- É importante observar que a modelagem através das RNA's não requer o conhecimento do fenómeno físico em questão, considerando que este conhecimento esteja disponível (aprendizagem por exemplos). Estudos biometeorológico do tipo ecológicos têm a vantagem (custo-benefício) quando comparados com outros estudos epidemiológicos (estudos transversais, estudos de coorte e estudos casos-testemunhas).
- No caso das redes treinadas, observou-se uma boa resposta com as variáveis utilizadas neste estudo, porém, é importante destacar, que as morbidades respiratórias na Cidade de São Paulo apresentam uma relação directa com os dias da semana. É possível que as redes treinadas apresentem uma melhor resposta, inserindo informações concernentes a feriados e paralisações do serviço público hospitalar, que ocorreram durante o período analisado.

- A classificação sinóptica adoptada nesta pesquisa, poderia ser aperfeiçoada através de modelos regionais do tempo, com resoluções (ordem de 10km) que viabilizariam acrescentar “novos sistemas”, que actuam na região de estudo. A informação desses “novos sistemas” é possível que melhore a resposta da morbidade respiratória.
- Finalmente, com a rede treinada e avaliada, ela poderia ser integrada em um sistema operacional, onde utilizaria as informações de modelos regionais do tempo (por exemplo, RAMS e MM5), assim como, as informações das concentrações de poluentes gerados por modelos de dispersão atmosférica (por exemplo, HYSPLIT e HYPACT). Desta forma, o Índice de Risco Meteorológico (IRM) e o prognóstico das doenças respiratórias, para a Cidade de São Paulo, poderiam ser disponibilizados para o Sistema Único de Saúde (SUS).

SUGESTÕES PARA TRABALHOS FUTUROS

- Utilizar a metodologia desenvolvida neste estudo para:
 - Outras regiões;
 - Doenças do Aparelho Circulatório e;
 - Prognósticos dos poluentes.
- Verificar a resposta da rede com duas ou mais camadas escondidas.

7. Referências Bibliográficas

Abdul-Wahab, S.A. & S.M. Al-Alawi, 2002. Assessment and prediction of tropospheric ozone concentration levels using artificial neural networks. *Environmental Modelling & Software* **17**, pp. 219–228.

Abreu, M. L., 1993. Previsores Sinóticos - Meteorológicos para a Dispersão de Poluentes na Região da Grande São Paulo. São José dos Campos, INPE. INPE-3600-TDL/197. DRH-DME. FRH/MET, pp. 131.

Andrade, M. F. & Carvalho, L. M., 1994. Estudos dos Componentes Principais Associados a Parâmetros Meteorológicos e seus Impactos na Determinação de Fontes de Aerossóis Atmosféricos em São Paulo. Anais do VIII Congresso Brasileiro de Meteorologia e II Congresso Latino-Americano e Ibérico de Meteorologia, pp. 125-129.

Antó, J. M. & Sunyer, J., 1986. Environmental Health: A Point-Source Asthma Outbreak. *Lancet*, **1**: pp. 900-903.

Balling, R. C., 1984. Classification in Climatology. In: Gaile, G. L. & Willmott, C. J., eds. *Spatial Statistics and models*. Dordrecht, D Reidel Publishing Company, pp. 81-108.

Bascom, R.; Bromberg, P. A.; Costa, D.A.; Devlin, R.; Dockery, D. W.; Frampton, M. W.; Lambert, W.; Samet, J. M.; Speizer, F. E.; Utell, M., 1996b. Health Effects of Outdoor Pollution. *Am. J. Respir. Crit. Care Med.*, **153**, pp. 477-98.

BATES, D. V.; SIZTO, R., 1983. Relationship between Air Pollution Levels and Hospital Admissions in Southern Ontario, Canada. *J. Public Health*, **74**, págs.:117-122.

Benvenuto, F. and Marani, A., 2000. Neural networks for environmental problems: data quality control and air pollution nowcasting. *Global Nest: The International J*, gnest.org, 2, pp. 281-292

Bertsekas D., 1995. *Non Linear Programming*. Athena Scientific.

Bibi, H., Nutman, A., Shoseyov, D., Shalom, M., Peled, R., Kivity, S. and Nutman, J., 2002. Prediction of Emergency Department Visits for Respiratory Symptoms Using an Artificial Neural Network. *Chest*, **122**, pp. 1627-1632

Bishop, C. M., 1995. *Neural Networks for Pattern Recognition*. Oxford University Press, pp. 482.

Boelter, J. K. & Davidson, H. J., 1997. Ozone Geration by Indoor, Electrostatic Air Cleaners. *Aerossol Science and Technology*, **27**, N. 14, pp. 689-708.

Böhm, G. M.; Saldiva, P.H.N.; Pasqualucci, C.A.; Massad, E.; Martins, M.A.; Zin, W.A.; Cardoso, W.V.; Criado, P.M.P.; Komatsuzaki, M.; Sakay, R.S.; Nigri, E.M.; Lemos, M.; Capelozzi, V.D.; Crestana, C.; Silva, R., 1989. Biological Effects of Air Pollution in Sao Paulo and Cubatão. *Environ. Res.*, **49**, pp. 208-16.

Braga, A. L. F., 1998. *Quantificação dos Efeitos da Poluição do Ar sobre a Saúde da População Pediátrica da Cidade de São Paulo e Proposta de Monitorização*. Tese de doutorado, Faculdade de Medicina, USP-SP, pp. 225.

Bryson, A. E. & Y. Ho, 1969. *Applied Optimal Control*. Blaisdell.

Bucher, K. & Haase, C., 1993. Meteorotropy and medical-forecasts. *Experientia*, **49(9)**, pp. 759-68.

Cannon, A. J. and Lord, E. R., 2000. Forecasting Summertime Surface-Level Ozone Concentrations in the Lower Fraser Valley of British Columbia: An Ensemble Neural Network Approach. *J. Air & Waste Manage. Assoc.*, 50, pp.322-339

Ccoyllo, Ó. R. S., 1997. Interação entre Poluentes Atmosféricos e a Circulação Local na Cidade de São Paulo. Tese de dissertação, Depto. de Ciências Atmosféricas, USP-SP, pp. 97.

Ccoyllo, Ó. R. S., 1997. Interação entre Poluentes Atmosféricos e a Circulação Local na Cidade de São Paulo. Tese de dissertação, Depto. de Ciências Atmosféricas, USP-SP, pp. 97.

CETESB, 1993. Relatório de Qualidade do Ar no Estado de São Paulo – SP.

CETESB, 2000. Relatório de Qualidade do Ar no Estado de São Paulo – SP.

CETESB, 2002. Relatório de Qualidade do Ar no Estado de São Paulo – SP.

Cigizoglu, H. K., 2003. Estimation, Forecasting and Extrapolation of River Flows by Artificial Neural Networks. *Hydrological Sciences Journal*, 48, pp.349-361

Conde, F. C., 2001. Uma análise de componentes principais de efeitos ambientais sobre a morbidade de doenças respiratórias em São Paulo. Dissertação de Mestrado. Depto. De Ciências Atmosféricas, USP-SP, pp. 122.

Cybenko, G., 1988. Continuous Valued Neural Network with two Hidden Layers are Sufficient. Technical Report, Department of Computer Science, Tufts University.

Cybenko, G., 1989. Approximation by Superposition of a Sigmoid Function. *Mathematics of Control, Signals and Systems*, 2, pp. 303-314.

Deryapa, N. R., 1986. Bioclimatological Aspects of population health: Meteorotropic diseases: World Climate Programme Applications, 2, WCAP- N° 2.

DHN, Cartas Diárias de Pressão à Superfície, 1993 a 2002. Divisão de Hidrografia e Navegação, Serviço de Previsão do Atlântico Sul – Ministério da Marinha.

Ellenius, J. and Groth, T., 2000 . Transferability of neural network-based decision support algorithms for early assessment of chest-pain patients. *Int. J. Med. Inform.* 60:1–20.

Fausset, L., 1994. *Fundamentals of Neural Networks: architectures, algorithms, and applications*, New York: Prentice Hall. ISBN 0-13-334186-0.

Fircket, D., 1931. Sur les Causes des Accidents Survenus Dans la Vallée de la Meuse, lors des Brouillards de Décembre 1930. *Bull. Acad. Roy. Med. Bel.*, 11, pp. 671-683.

FUNDAÇÃO SEADE, 1992. *Pesquisa das Condições de Vida na Região Metropolitana de São Paulo*. Fundação Seade, São Paulo.

Geman S., E. Bienenstock, e R. Doursat, 1992. Neural Network and the Bias-variance Dilemma. *Neural Computation*, 4, pp.1-58.

Gonçalves, F. L. T.; L. M. V. Carvalho; Conde, F. C.; M. R. D. O. Latorre, Saldiva, P. H. N., Braga, A. L. F., 2005. The effects of air pollution and meteorological parameters on respiratory morbidity during the summer in São Paulo city. *Environment International*, 31, pp. 343-349.

Grossberg S., 1976. Adaptive Pattern Recognition and Universal Recoding: Parallel Development and Coding of Neural Feature Detectors. *Biol. Cybern.*, 24, pp. 121-134.

Grossberg S., 1987. Competitive Learning: From Interactive Activation to Adaptive Resonance *Cognitive Science*, 11, pp. 23-63.

Hagan, M. & Menhaj, M., 1994. Training Feedforward Network with the Marquardt Algorithm. *IEEE Transactions on Neural Networks*, November, 5(6), pp. 989-993.

Hall, T., 1998. Precipitation Forecasting Using a Neural Network. *Weather and Forecasting*, 14, pp. 338–345.

Haykin, S.; 1994. *Networks – A comprehensive Foundation*. 2nd edition. Macmillan.

Hazucka, M. J. & BATES, D. V., 1985. Combined Effect of Ozone and Sulfur Dioxide on Human Pulmonary Function. *Nature*, 257, pp. 50-51.

Hebb, D. O., 1949. *The Organization of Behavior*. Wiley.

Hirsch, S.; Shapiro J. and Frank, P., 1997 Use of an artificial neural network in estimating prevalence and assessing underdiagnosis of asthma. *Neural Computing & Applications*, 5, pp. 124-128

Hopfield, J. J., 1982. Neural Networks and Physical Systems with Emergent Collective Properties. *Proc. Nat. Acad. Sci.*, 79, pp. 2554-2558.

Höppe, P., 1997. Aspects of Human Biometeorology in Past, Present and Future. *Int. J Biometeorol*, 40, pp. 19-23.

INSTITUTO BRASILEIRO DE GEOGRAFIA E ESTATÍSTICA - IBGE, 2000. Sinopse preliminar do censo demográfico 2000, 7.

Jauregui, E., 1986. Incidence of Respiratory Illness and Air Pollution Levels in Mexico. *Climate and Man's Health*. *Climate and Human Health: World Climate Programme Applications*, 1, WCAP- N° 2.

Jendritzky, G. & Sievers, U., 1989. Human Biometeorological Approaches with Respect to Urban Planning. In: Driscoll, D. & Box E. D., Eds. *Proceedings of the XIIth Congress of the International Society of Biometeorology*, West Lafayette, USA. The Hague, SPB Academic Publishing, pp. 25-39.

Jones, T. S. et al., 1982. Morbidity and Mortality Associated with the July 1980 Heat Wave in St. Louis and Kansas City. *Journal of the American Medical Association*, **247**, pp. 3327-3330.

Kaiser, H. F., 1960. The Varimax Criteria for Analytical Rotation in Factor Analysis. *Psychometric*, **23**, pp.:141-151.

Kalkstein, L. S. & Davis, R. E., 1989. Weather and the Human Mortality: An evolution of Demographic and Inter-regional Responses in the United States. *Annals of the Association of American Geographers*, **79**, pp. 44-64.

Kalkstein, L. S. & Smoyer, K. E., 1993. The Impact of Climate Change on Human Health: Some International Implications. *Experientia*, **49(11)**, pp. 969-979.

Kalkstein, L. S., 1991. A New Approach to Evaluate the Impact of Climate on Human Mortality. *Environ. Health Perspectives*, **96**, pp. 145-150.

Karnin, E. D., 1990. A Simple Procedure for Pruning Back-Propagation Trained Neural Networks. *IEEE Transactions on Neural Networks*, **1:(2)**, pp. 239-242.

Keiding, K.; Jensen, F.; Heidman, 1986. Absolute Modeling of Urban Aerosol Composition by Factor Analysis. *Analytica Química Acta*, **181**, pp. 79-85.

Koren, H. S.; Delvin, R. B.; Graham, D. E.; Mann, R.; Mcgee, M. P.; Horstman, D. H.; Kozumbo, W. J.; Becker, S.; House, D. E.; Mcdonnell, W. F.; Bromberg, P. A, 1989. Ozone-Induced Inflammation in the Lower Airways of Human Subjects- *Am. Rev. Respir. Dis.*, **139**, pp. 407-415.

Kóvacs, Z. L., 2002. *Redes Neurais Artificiais*. São Paulo, Editora Livraria da Física.

Lay, D., 1997. *Algebra Linear and its Applications*. Addison Wesley.

- Le Cun., Y., 1985. A learning Procedure for Assymmetric Threshold Network. In Proceedings of Cognitive 85, pp. 599-604.
- Leighton, R. M; Spark, E., 1997. Relationship between Synoptic Climatology and Pollution Events in Sydney. *Int. J Biometeorol.*, **41**, pp. 76-89.
- Lin, C. A., 1997. Efeitos da Poluição Atmosférica sobre a Morbidade Respiratória Aguda na População Infantil de São Paulo. Tese de doutorado, Faculdade de Medicina, USP-SP, pp. 115.
- Logan, W. P. D., 1953. Mortality in London Fog Incident. *Lancet*, **1**, pp. 336-338.
- Lu, W. Z.; Fan, H. Y.; Leung, A. Y. T. and Wong, J. C. K., 2002 Analysis of Pollutant Levels in Central Hong Kong Applying Neural Network Method with Particle Swarm Optimization. *Environmental Monitoring and Assessment*, **79**, pp 217-230
- Manfredini, L. M., 1988. Condições Meteorológicas para Previsão de Situações Críticas de Poluição na Grande São Paulo. Dissertação de Mestrado. IAG – USP.
- Martyn, D. 1992. *Climates of the World. Developments in Atmospheric*, **18**. Polish Scientific Publishers. Warsawa.
- Mausner & Bahn, 1999. *Introdução à epidemiologia (2ª Ed.)*. Lisboa: Fundação Calouste Gulbenkian.
- McCulloch W. S. & Pitts, W., 1943. A Logical Calculus of the Ideas Immanent in Nervous Activity. *Bulletin of Mathematical Biophysics*, **5**, pp. 115-133.
- McGregor, G. R; Walters, S.; Wordley, J., 1999. Daily hospital respiratory admissions and winter air mass types, Birmingham, UK. *Int J Biometeorol* (1999) **43**: pp. 21–30.

Mendel, J. M., & McLaren, R. W., 1970. Adaptive, Learning, and Pattern Recognition Systems; Theory and Applications, chapter Reinforcement-learning control and pattern recognition systems. New York: Academic Press, pp. 287-318.

Michaelides, S. C., Pattichis, C. S. and Kleovoulou, G., 2001 Classification of rainfall variability by using artificial neural networks. *Int. J. Clim.*, 21, pp. 1401-1414

Minsky M. L., 1961. Steps Towards Artificial Intelligence. In *Proc. of the Institute of Radio Engineers* 49, pp. 8-32.

Minsky, M. L., & S. Papert, 1969. *Perceptrões: an Introduction to Computational Geometry*. MIT Press, Massachusetts.

Miraglia, S. G. Elk.; Saldiva, P. H. R. N.; Böhm, G. R., 2005. An Evaluation of Air Pollution Health Impacts and Costs in São Paulo, Brazil. *Environmental Management* 35, No. 5, pp. 667–676

Muller, R. A. & Jackson, A. L., 1985. Estimates of Climatic Air Quality Potential at Shreveport, Louisiana. *Journal of Climate and Applied Meteorology*, 24, pp. 293-301.

Munn, R. E., 1968. *Biometeorological Methods*. Acad. Press., pp. 336.

NATIONAL RESEARCH COUNCIL. U.S. BOARD ON ENVIRONMENTAL STUDIES COMMITTEE ON ADVANCES IN ASSESSING HUMAN EXPOSURE TO AIRBONE POLLUTANTS. HUMAN EXPOSURE ASSESSMENT FOR AIRBONE POLLUTANTS, 1996. *Advances and opportunities*. Washington, National Academy of Sciences, 1991 apud BASCON et al., pp.12.

Nazir, J.; Barlow, D. J.; Lawrence M. J.; Richardson, C. J.; Shrubbs, I.. 2002. Artificial Neural Network Prediction of Aerosol Deposition in Human Lungs. *Pharmaceutical Research*, 19, No. 8, pp. 1130-1136.

Nganga, J. K. & Ngugi, G. K., 1986. The Influence of Weather on Measles and Pneumonia in Nairobi. *Climate and Human Health: World Climate Programme Applications*, 1, WCAP- N° 2.

OBSERVAÇÕES DE SUPERFÍCIE EFETUADAS NA ESTAÇÃO METEOROLÓGICA DO IAG – USP. São Paulo, Boletim Anual (ISSN 1415-4374).

Panero, C; Pablo DE F.; Sánchez, J. M.; Tomás, C., 1997. Statistical Modeling and Prediction of Pollutants in the Urban Atmosphere of Salamanca, Spain. *Int. J. Biometeorol.*, **40**, pp. 223-33.

Parker D., 1985. *Learning Logic: Casting the Cortex of the Human Brain in Silicon* Technical Report, Center for Computational Research in Economics and Management Science, MIT.

Perry, A. H., 1983. Growth Points in Synoptic Climatology. *Progress in Physical Geography*, **7**, pp. 90-96.

Perry, R.; Young, R. J., 1977. *Handbook of the Air Pollution Analysis*. London, Chapman and Hall, pp. 506.

Persinger, M. A., 1980. *The Wheeler Matrix and Human Behavior*. New York, Praeger.

Pope, C. A. III & Dockery, D. W., 1992. Acute Health Effects of PM₁₀ Pollution on Symptomatic and Asymptomatic Children. *Am. Rev. Resp. Dis.*, **145**, pp. 1123-1128.

Pope, C. A. III, 1991. Respiratory Diseases Associated with PM₁₀ Pollution in Utah Salt Lake, and Cache Valley. *Arch Environ. Health*, **46**, pp. 90-7.

Pope, C. A., III; Schwartz, J., 1996. Time Series for the Analysis of Pulmonary Health Data. *Am. J. Respir. Crit. Care Med.*, **54**, pp. 5229-33.

Pryor, W. A.; Lightsey, J. W., 1981. Mechanism of Nitrogen Dioxide reactions: Initiation of Lipid Peroxidation and the Production of Nitrous Acid- *Science*, **214**, pp. 435-7.

Queirós, M.; Bonito-Vitor, A.; Costa-Pereira, A.; Maia, J. C., 1990. Childhood Asthma and Outdoor Air Pollution in Oporto Area. *Allergol. Et Immunopathol.*, **18**, pp. 291-5.

R. S. Sutton., R. S.; Barto, A. G.; Williams, R. J., 1991. Reinforcement Learning is Direct Adaptive Optimal Control. In *Proc. of the American Control Conference*, pp. 2143-2146.

Reed, R., 1993. Pruning Algorithms – *A survey*. *IEEE Transactions on Neural Networks*, **4(5)**:740-746.

Riedmiller M., 1994. Rprop – Description and Implementation Details. Technical Report University of Karlsruhe, 1994.

Rosenblatt F., 1958. *The Perceptron: A Probabilistic Model for Information Storage and Organization in the Brain*. *Psychol. Rev.*, **65**, pp. 386-408.

Rumelhart D. E. & McClelland J. L., 1986. *Parallel Distributed Processing*, volume 1: Foundations. The MIT Press.

Rumelhart D. E., Hinton, G. E., Williams, R. J., 1986. Learning Representations by Back-propagation Error. *Nature*, **323**, pp. 533-536.

Rusanov, V. I., 1986. Climate and Man's Health. *Climate and Human Health: World Climate Programme Applications*, **1**, WCAP- N° 2.

Sanchez, M., Pascual, D., Ramos, C., and Perez, I. , 1990. Forecasting Particulate Pollutant Concentration in a City from Meteorological Variables and Regional Weather Patterns. *Atmos. Env.*, **24A**: pp. 1509-19.

Schwartz, J., 1996. Air Pollution and Hospital Admissions for Respiratory Disease. *Epidemiol.*, 7, pp. 20 – 8,.

Schwartz, J., Marcus, A., 1990. Mortality and Air Pollution in London: A Times Series Analysis. *Am. J. Epidemiol.*, 131, pp. 185-94.

Schwartz, M. D., 1991. An Integrated Approach to Air Mass Classification in the North Central United States. *Professional Geographer*, 43(1), pp. 77-91.

Setzer, A.; Cavalcanti; I. F. A; Marcelino, B. C., 1980. Influências Meteorológicas na Poluição Atmosférica em São Paulo. São José dos Campos. INPE-1669-RPE/107, pp. 36.

Silverman, D. and Dracup, J. A., 1999. Artificial Neural Networks and Long-Range Precipitation Prediction in California. *J. App. Meteor.*, 39, pp. 57–66.

Spellman, G., 1999. An application of artificial neural networks to the prediction of surface ozone concentrations in the United Kingdom. *Appl. Geogr.* 19:123–136.

Sugahara, S., 1991. Flutuações Interanuais, Sazonais e Intrasazonais da Precipitação no Estado de São Paulo. Dissertação de Doutorado. Depto. De Ciências Atmosféricas, USP-SP.

Sunyer, J.; Aantó, J. M.; Murilo, C.; Saez, M., 1991. Effects of Urban Air Pollution on Emergency Room Admissions for Chronic Obstructive Pulmonary Disease. *Am. J. Epidemiol.*, 134, pp. 277-286.

Sycara K., 1988. Using Case-Based Reasoning for plan Adaptation and Repair. In *Case Based Reasoning Workshop, DARPA, Morgan Kaufmann*, pp.425-434.

Tikhonov A. N. & V. Y. Arsenin, 1977. *Solutions of Ill-Posed Problems*. V.H. Winston & Sons, John Wiley & Sons, Washington D.C. Translation editor Fritz John.

Tkacz, G., 2001. Neural network forecasting of Canadian GDP growth. *Int. J. Forecasting* 17:57–69.

Viotti, P.; Liuti, G. and Di Genova P., 2002. Atmospheric urban pollution: applications of an artificial neural network (ANN) to the city of Perugia. *Ecological Modelling*, 148, pp.27-48

Wax, C. L.; Borengasser, M. J. & Muller, R. A., 1978. Barataria Basin: Synoptic Weather Environmental Responses. Baton Rouge, Center for Wetland Resources Louisiana State University (Coastal zone Management Series).

Werbos P., 1974. Beyond Regression: New Tools For Prediction and Analysis in the Behavioral Sciences. PhD thesis, Harvard University.

Widrow, B. and Hoff, M. E., 1960. Adaptive Switching Circuits. In 1960 IRE Western Electric Show and Convention Record, Part 4, pp. 96-104.

Widrow, B., 1962. Generation and information storage in networks of adaline neurons", in *Self-Organizing Systems*, M. Yovits and G. Giacobi and G. Goldstein (eds.), Spartan Books, Washington, D.C., , pp.435-461.

Wilks, D. S., 1995. *Statistical Methods in Atmospheric Sciences*. Department of Soil, Crop, and Atmospheric Sciences Cornell University, Ithaca-New York, pp. 467.

Yana, H.; Jiangb, Y; Zhenge, J.; Pengc, C.; Lid, Q., 2006 A multilayer perceptron-based medical decision support system for heart disease diagnosis. *Expert Systems with Applications*, 30, pp. 272–281

Yarnal, B., 1993. *Synoptic Climatology*, in *Environmental Analysis. A Primer Studies in Climatology Series*. London, Belhaven Press.

Zakrzewski, F. S., 1997. Principles of Environmental Toxicology. State University of New York at Buffalo. Second Edition. American Chemical Society, Washington – D.C.

ANEXO A

A resposta $y(t)$ no domínio do tempo de um sistema dinâmico pode ser aproximada por:

$$y(t) = f(C, Z(t)) \quad (\text{A-1})$$

Onde $f(.)$ é uma função de mapeamento aproximada entre os parâmetros de entrada e a saída, $Z(t)$ é um vector conhecido representando as entradas (excitação) no sistema e C é um vector de coeficientes constantes, no caso das RNA's o vector C representa todos os pesos sinápticos da mesma.

- **Método do Gradiente Descendente**

O método do Gradiente Descendente, também comumente chamado de "Steepest Descent", é um método tradicional para solução de problemas de optimização sem restrição.

Neste método parte-se de um valor inicialmente arbitrário para o vector C . A partir daí calcula-se o vector gradiente e caminha-se iterativamente em direcção contrária ao mesmo. Deste modo o deslocamento da solução vai à direcção do ponto de mínimo mais próximo. Matematicamente o método do Gradiente Descendente é dado por:

$$C_{j+1} = C_j - \eta G_j \quad (\text{A-2})$$

Onde η é a chamada taxa de aprendizagem, C_j e C_{j+1} são o vector solução na iteração actual e na próxima, respectivamente, e G_j corresponde ao vector gradiente calculado na iteração actual.

A Tabela A-1 ilustra este método sob a forma de algoritmo. O algoritmo deve ser interrompido quando for atingido um número máximo de iterações e/ou um valor aceitável pré-estabelecido para o erro quadrático médio for alcançado.

ETAPA	PROCEDIMENTO
1	$j = 0$ Use um método de pré-condicionamento ou arbitre um valor inicial para o vector de pesos C_0 e arbitre um valor para η .
2	$j = j + 1$ Calcule o vector gradiente G_j
3	Faça $C_{j+1} = C_j - \eta \cdot G_j$
4	Se o critério de parada não for atendido volte para etapa 2

Tabela A-1 – Algoritmo do Método do Gradiente Descendente.

Este método não apresenta garantia de convergência (KÓVACS, 2002). Pelo facto de adoptar a direcção de busca do mínimo como sendo a mesma do vector gradiente, o custo computacional deste método também se torna elevado.

- **Método do Gradiente Conjugado**

O método do gradiente conjugado pertence a uma classe de métodos de optimização de segunda ordem, que colectivamente são conhecidos como métodos da direcção conjugada. Este método existe sob a forma linear e não-linear.

✓ Forma linear

A seguir será apresentado um resumo do desenvolvimento em sua forma linear. Uma apresentação mais extensa pode ser vista em Haykin (1994).

O método em sua forma linear tem um desenvolvimento teórico baseado na minimização da seguinte função:

$$f(X) = \frac{1}{2} X^T B X^T - F^T X^T + d \quad (\text{A-3})$$

Onde X e F são vectores, B é uma matriz, e d é um escalar.

É possível demonstrar que se a matriz B é simétrica e positiva, então, a função (A-3) é minimizada por:

$$BX = F \quad (\text{A-4})$$

Um conjunto de vectores não nulos S_0, S_1, \dots, S_n é um conjugado de B se a seguinte condição for satisfeita:

$$S_i B S_j^T = 0 \quad (\text{A-5})$$

Para todo i e j tal que $i \neq j$

Uma importante propriedade dos vectores conjugados de B é que eles são linearmente independentes. Logo eles formam uma base do espaço vectorial de B .

Para um dado conjunto de vectores S , o método do gradiente conjugado para minimização da função (A-3) é definido por:

$$X_{j+1} = X_j + \eta_j S_j \quad (\text{A-6})$$

Onde X_0 é vector inicial arbitrário e η_j é um escalar definido por (A-7)

$$f(X_j + \eta_j S_j) = \min (f(X_j) + \eta_j S) \quad (\text{A-7})$$

O procedimento para a escolha de η_j , de forma a minimizar a função $f(X_j + \eta_j S_j)$ para um dado j , é conhecido como busca em linha (line search). Para o método em sua forma linear a taxa de aprendizagem η é dada por:

$$\eta_j = \frac{S_j^T B \varepsilon_j}{S_j^T B S_j} \quad (\text{A-8})$$

Onde ε_j é vector erro definido por:

$$\varepsilon_j = X_j - X^* \quad (\text{A-9})$$

Começando de um ponto arbitrário X_0 , o método garante que converge para um ponto de mínimo X^* da função (A-3) em um número finito de iterações.

O funcionamento deste método implica na disponibilidade de um conjunto de vectores S_0, S_1, \dots, S_n conjugados de B . Esses vectores podem ser obtidos pelas expressões:

$$R_j = F - B X_j \quad (\text{A-10})$$

$$S_j = R_j - \beta_j S_{j-1} \quad (\text{A-11})$$

Onde o vector S_j é obtido através de uma combinação linear de R_j e S_{j-1} , onde β_j é um escalar definido por:

$$\eta_j = \frac{S_j^T B R_j}{S_{j-1}^T B S_{j-1}} \quad (\text{A-11})$$

O método em sua forma linear presta-se para a solução de sistemas lineares de equações (equação (A-4)) pois nestes casos conhecemos a matriz B e o vector independente F . O seu emprego para o cálculo dos pesos nas RNA's é possível através do método na forma não-linear, descrita a seguir.

✓ Forma não-linear

O método do gradiente conjugado não-linear apresenta o mesmo desenvolvimento que o linear. Entretanto como não conhecemos a matriz B e o vector F , o vector resíduo é dado por:

$$R_j = -G_j \quad (\text{A-12})$$

O cálculo de β_j deve ser feito a partir da fórmula de Polak-Ribière:

$$\beta_j = \max \left[\frac{R_j (R_j - R_{j-1})}{R_j^T R_j}, 0 \right] \quad (\text{A-13})$$

O método de busca em linha, adoptado no método não linear, deve minimizar o erro quadrático médio. Basta actualizar o vector:

$$X_j = X_j + \eta S_j \quad (\text{A-14})$$

Variando-se o valor de η até que se encontre o erro médio quadrático.

Haykin (1994), descreve um mecanismo de busca em linha que se divide em duas etapas:

1. Faça η variar até que se descubra um intervalo no qual o erro quadrático médio seja o mínimo.

2. Quebre o intervalo em subintervalos de comprimentos progressivamente menores até que o mínimo seja encontrado.

Desta forma, o método do gradiente conjugado não linear aplicado ao cálculo do vector C , associado às RNA's pode ser escrito sob a forma de algoritmo na Tabela A-2.

O algoritmo deve ser interrompido quando for atingido um número máximo de iterações e/ou quando um valor máximo aceitável pré-estabelecido para o erro médio quadrático for alcançado.

O método do gradiente conjugado faz uso de um vector que dá a direcção de busca. Isso faz com que o ponto de mínimo seja encontrado muito mais rápido que no método do gradiente descendente, além da convergência ser garantida.

ETAPA	PROCEDIMENTO
1	$j = 0$ Use um método de pré-condicionamento ou arbitre um valor inicial para o vector de pesos C_0
2	Calcule o vector gradiente: G_0
3	Faça $S_0 = R_0 = -G_0$
4	Use um método de busca em linha para o cálculo de η
5	Atualize o vector C : $C_{j+1} = C_j + \eta \cdot S_j$
6	Calcule o novo vector gradiente G_{j+1}
7	Faça: $R_{j+1} = -G_{j+1}$
8	Use a fórmula de Polak-Ribière para calcular o escalar β : $\beta_j = \max \left\{ \frac{R_j \cdot (R_j - R_{j-1})}{R_j^T \cdot R_j}, 0 \right\}$
9	Atualize o vector direcção S : $S_{j+1} = R_{j+1} + \beta_j \cdot S_j$
10	$j = j + 1$ Se o critério de parada não for atendido volte para a etapa 4.

Tabela A-2 – Algoritmo do Método do Gradiente Conjugado não linear.