

# “ABORDAGEM NEURO-DIFUSA PARA MODELIZAÇÃO DE UM SISTEMA DE BRANQUEAMENTO DE PASTA DE PAPEL”

R. P. Paiva<sup>1\*</sup>

A. Dourado<sup>2\*</sup>

B. Duarte<sup>3</sup>

## RESUMO

Actualmente, os requisitos relativos à qualidade de papel produzido pela indústria são bastante elevados. Deste modo, o processo de branqueamento da pasta é preponderante para a qualidade final do produto. Assim, de modo a que os requisitos do mercado sejam satisfeitos é fundamental haver um conhecimento profundo do sistema de branqueamento. Neste artigo propõe-se uma abordagem neuro-difusa capaz de prever a brancura da pasta em folha produzida pela Companhia de Celulose do Caima, S.A. Esta estratégia é desenvolvida em duas fases: na primeira utiliza-se uma técnica de agrupamento difuso de que resulta a obtenção de um conjunto de regras difusas, enquanto que na segunda, os centros e as larguras das funções de pertença são sintonizados através de uma rede neuronal difusa treinada pelo algoritmo de retropropagação do erro. Esta metodologia apresenta a vantagem de permitir bons resultados com sistemas fortemente não lineares, para além de descrever o sistema através de um conjunto de regras linguísticas, mais próximas da linguagem humana natural e, assim, mais intuitivas para os operadores. Os resultados obtidos com esta metodologia podem-se considerar aceitáveis pois que o erro da previsão é cerca de 0.5% do valor médio da variável prevista. No entanto verifica-se a necessidade de integrar no modelo a influência de atrasos variáveis para cada um dos *inputs*.

**Palavras-chave:** branqueamento de pasta de papel, agrupamento difuso, modelização neuro-difusa.

## 1. INTRODUÇÃO

---

<sup>1</sup> Assistente Estagiário no Departamento de Engenharia Informática da Universidade de Coimbra (DEIUC); Licenciado em Engenharia Informática; e-mail: ruipedro@dei.uc.pt

<sup>2</sup> Professor Associado do DEIUC; Doutorado e Agregado em Engenharia Electrotécnica; e-mail: dourado@dei.uc.pt

\* CISUC – Centro de Informática e Sistemas da Universidade de Coimbra, Departamento de Engenharia Informática, PÓLO II da Universidade de Coimbra, Pinhal de Marrocos, P 3030, Coimbra, Portugal,  
Tel. 039-790000, Fax: 039-701266

<sup>3</sup> Assessor Industrial; Doutorado em Engenharia Química; e-mail: nop61212@mail.telepac.pt, Companhia de Celulose do Caima, S A, P 2250 Constância, Portugal, Tel. 049-730000, Fax: 049-96284

O branqueamento de pasta de papel é um processo fortemente não linear, envolvendo fenómenos físicos de importância considerável, contudo insuficientemente compreendidos. Além do mais, a sequência de branqueamento é influenciada por um número elevado de variáveis cujo peso relativo não é conhecido com precisão. Em virtude da ausência de um modelo baseado nos princípios físico-químicos do processo e do elevado grau de dificuldade na sua obtenção o recurso a uma abordagem baseada em técnicas que permitam descrever o processo com base no conhecimento aproximado do processo e dados de operação recolhidos, nomeadamente algumas técnicas da chamada Inteligência Artificial, é uma alternativa de validade inequívoca. Do leque de escolhas possíveis, as estruturas difusas são de particular interesse, uma vez que constituem aproximadores universais não lineares e permitem a descrição de sistemas através de um conjunto de regras linguísticas próximas da linguagem natural. Além do referido, permitem a integração de conhecimento *a priori*, fornecido por operadores e engenheiros de processo. Assim, neste artigo propõe-se um método neuro-difuso aplicado à modelização de um sistema de branqueamento de pasta de papel, compreendendo duas fases. Na primeira fase é utilizada uma técnica de agrupamento difuso (agrupamento subtractivo) no espaço de entradas-saídas com o objectivo de se obter um conjunto de regras linguísticas descritivas do processo. Na segunda fase, os centros e larguras das funções de pertença anteriormente obtidas são ajustados através de uma rede neuronal difusa, treinada pelo algoritmo de retropropagação do erro. Uma importante característica do sistema é o seu atraso de transporte variável. Este fenómeno dificulta a aprendizagem, uma vez que a estrutura do sistema neuro-difuso não permite capturar o seu efeito. As consequências deste fenómeno são apresentadas e discutidas adiante.

O artigo presente está organizado em cinco secções. A Secção 2 descreve o algoritmo de modelização difusa. Na Secção 3, a planta de branqueamento é descrita resumidamente. A aplicação das técnicas neuro-difusas aos dados da planta de branqueamento é conduzida na Secção 4. Finalmente, na Secção 5 apresentam-se conclusões, nomeadamente as vantagens e limitações da técnica descrita, os principais problemas encontrados, assim como direcções para trabalho futuro.

## 2. MODELIZAÇÃO DIFUSA

Considere-se um sistema com várias entradas e várias saídas (sistema MIMO). Seja  $p \hat{I} N$  o seu número de entradas e  $q \hat{I} N$  o seu número de saídas.

$$Y(k) = f[\theta(k), \xi(k)], \quad Y(k) \in \mathfrak{R}^q, \quad \xi(k) \in \mathfrak{R}^q \quad (1)$$

onde  $f$  é uma função linear ou não linear,  $x(k)$  representa uma perturbação estocástica e  $q(k)$  representa as entradas e saídas passadas. Assim:

$$\theta(k) = [y_1(k-1), \dots, y_1(k-n_1), \dots, y_q(k-1), \dots, y_q(k-n_q), \dots, u_1(k-d_1), \dots, u_1(k-m_1-d_1+1), \dots, u_p(k-d_p), \dots, u_p(k-m_p-d_p+1)]^T \quad (2)$$

Os parâmetros  $n_1, \dots, n_q, m_1, \dots, m_p, d_1, \dots, d_p$  relacionam-se com a ordem e com o atraso puro discreto do sistema.

A equação (1) permite representar um processo através de um modelo não-linear auto-regressivo envolvendo uma variável exógena: modelo NARX. O seu objectivo final é obter uma aproximação da função  $f$ . Neste artigo, este problema é resolvido através de uma técnica de modelização difusa. Deste modo, o modelo é do tipo difuso auto-regressivo envolvendo uma variável exógena – modelo FARX.

## 2.1. ESTRUTURA DO MODELO FARX

A função  $f$  é aproximada por um sistema de inferência de Mamdani. O processo é representado por um conjunto de  $R$  regras do tipo (3):

$$R_i: \text{If } y_1(k) \text{ is } A_{1i} \text{ and } \dots \text{ and } y_q(k-n_q) \text{ is } A_{qn_i} \text{ and } \dots \text{ and } u_1(k-d_1) \text{ is } B_{1i} \text{ and } \dots \text{ and } u_p(k-m_p-d_p+1) \text{ is } B_{pm_i} \quad (3)$$

$$\text{then } y_1(k+1) \text{ is } C_{1i} \text{ and } \dots \text{ and } y_q(k-n_q+1) \text{ is } C_{qi}$$

onde  $A_{jki}$ ,  $B_{jki}$  and  $C_{ji}$  denotam os termos linguísticos de cada variável de saída e entrada, definidas pelas suas funções de pertença:  $\mu_{A_{jki}}, \mu_{B_{jki}}, \mu_{C_{ji}}, i = 1, 2, \dots, R$ .

## 2.2. IDENTIFICAÇÃO

Os parâmetros  $n_1, \dots, n_q, m_1, \dots, m_p, d_1, \dots, d_p$  têm que ser escolhidos adequadamente com base em conhecimento *a priori* ou por comparação de estruturas diferentes, de acordo com um dado critério objectivo. Assumindo que esta questão está ultrapassada, o objectivo é, então obter um conjunto de regras do tipo (3) e ajustar os parâmetros das funções de pertença, utilizando amostras de dados de operação do sistema:

$$X = [\theta(1) \ \dots \ \theta(N-1)]^T, \Psi = [Y(1) \ \dots \ Y(N-1)]^T \quad (4)$$

em que  $N$  representa o número de observações utilizadas na identificação.

### 2.2.1. AGRUPAMENTO DIFUSO

De modo a obter-se um conjunto de  $R$  regras, evitando-se os problemas inerentes às partições do tipo *grelha*, e. g., explosão da base de regras, é aplicada uma técnica de agrupamento difuso: *agrupamento substractivo* (subtractive clustering) [1]. Esta técnica, tal como outras técnicas de agrupamento difuso, é utilizada por permitir particionar o espaço de entradas-saídas de forma flexível.

O algoritmo de agrupamento substractivo consiste, essencialmente, numa forma modificada do Método da Montanha. Assim, seja o conjunto  $Z$  formado pela concatenação dos conjuntos  $X$  e  $Y$  (4). Assume-se que todos os pontos em consideração são normalizados em cada dimensão de modo a estarem limitados por um hiper-cubo. Cada ponto é visto como um potencial centro de grupo estando a ele associado um determinado potencial:

$$P_i = \sum_{j=1}^N e^{-\alpha \|z_i - z_j\|^2} \quad (5)$$

sendo  $\alpha = 4/r_a^2$  e  $r_a > 0$  uma constante definindo o raio da vizinhança de cada centro. Assim, o potencial associado a cada ponto depende da sua distância a todos os outros, o que leva a que uma vizinhança densa origine um potencial elevado para o seu centro.

Depois de calculado o potencial de cada ponto selecciona-se aquele com o maior potencial como o primeiro centro do grupo. Defina-se  $z_1^*$  como o centro do primeiro grupo e  $P_1^*$  como o seu potencial. O potencial de cada ponto  $z_i$  é, então, actualizado de acordo com a Eq. (6):

$$P_i \leftarrow P_i - P_1^* e^{-\beta \|z_i - z_1^*\|^2} \quad (6)$$

onde  $\beta = 4/r_b^2$  e  $r_b > 0$  é uma constante representando o raio da vizinhança com reduções sensíveis no seu potencial. Tipicamente,  $r_b = 1.5r_a$ . Assim, os pontos mais próximos do primeiro centro vão ter o seu potencial mais reduzido. Deste modo, a probabilidade de algum desses pontos ser escolhido como o próximo centro é menor. Este procedimento de escolha de centros e redução de potencial é efectuado iterativamente até que se satisfaça o critério de paragem seguinte:

Se  $P_k^* > e^{up} P_1^*$   
**Aceitar  $z_k^*$**  como centro de grupo e continuar

Caso contrário,  
 Se  $P_k^* < e^{down} P_1^*$   
**Rejeitar  $z_k^*$**  e terminar.

Caso contrário  
 Seja  $d_{min}$  = menor distância entre  $z_k^*$  e todos os centros já encontrados  
 Se  $d_{min}/r_a + P_k^*/P_1^* \geq 1$   
**Aceitar  $z_k^*$**  como centro de grupo e continuar

Caso contrário  
**Rejeitar  $z_k^*$**  e atribuir ao seu **potencial** o valor **0.0**.  
**Seleccionar** o ponto com o **potencial mais elevado** como o novo  $z_k^*$ .  
 Voltar a **testar**.

Fim Se

Fim Se

Fim Se

Aqui,  $e^{up}$  especifica um limiar para o potencial acima do qual o ponto é seguramente aceite como centro,  $e^{down}$  especifica o limiar oposto, em que o ponto é seguramente rejeitado. Na terceira situação, verifica-se se o ponto em questão se caracteriza por um compromisso aceitável entre ter um potencial suficientemente elevado e estar suficientemente afastado dos centros já determinados. Tipicamente,  $e^{up}=0.5$  e  $e^{down}=0.15$ .

Depois de se obter um conjunto de regras difusas, estas são incluídas no sistema de inferência. Antes, porém, é necessário obter os centros e larguras das funções de pertença. Os centros serão os mesmos que os centros dos grupos obtidos. Quanto às larguras, e uma vez que se usam funções gaussianas, os seus desvios padrões são determinados pela fórmula heurística:

$$\sigma_{ij} = r_a \cdot \frac{\max(z_{kj}) - \min(z_{kj})}{\sqrt{8}}, k = 1, \dots, N \quad (7)$$

## 2.2.2. REDE NEURONAL DIFUSA

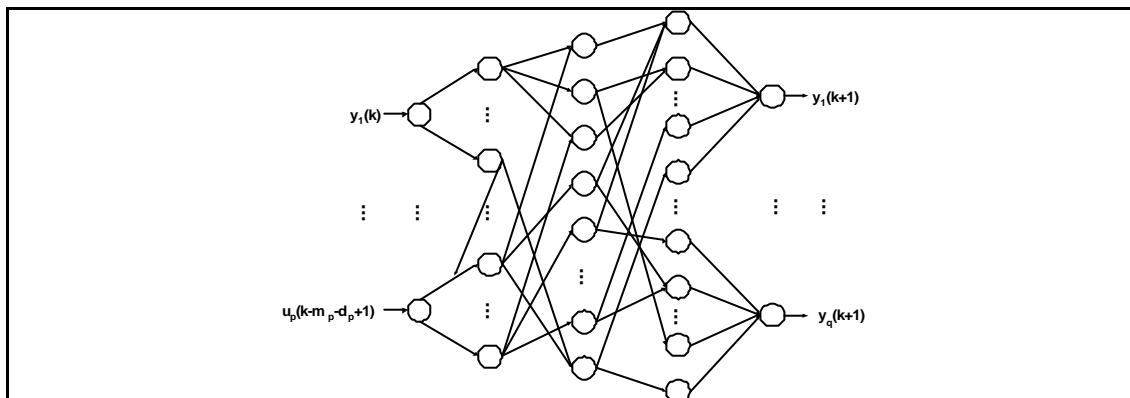


Figura 1: Estrutura da rede neuronal difusa

Após a determinação de um sistema de inferência difuso inicial com base no agrupamento difuso há que otimizar os seus parâmetros, i.e., os centros e larguras das funções de pertença. Essa tarefa foi levada a cabo por meio do treino de uma rede neuronal difusa (RND) com base no algoritmo de retropropagação do erro.

A estrutura da RND é apresentada na Figura 1. Essa rede é composta por cinco camadas. A primeira camada contém os nós de entrada, que representam as variáveis linguísticas de entrada. Esta camada simplesmente passa os valores das entradas para a camada seguinte. Os nós da camada 2 desempenham o papel de termos linguísticos de cada variável de entrada, representados por funções de pertença gaussianas. Esta camada é responsável pela *fuzzificação* dos valores reais de entrada. Na camada 3, cada nó representa uma regra do sistema de inferência difuso. Os antecedentes de cada regra são definidos através do estabelecimento adequado de ligações entre os nós da camada 2 e

os nós da camada 3. Nesta camada, cada regra é disparada com base numa das possíveis operações *AND* difusas. Uma vez que há algumas regras com o mesmo conseqüente, a camada 4 integra essas regras, utilizando uma qualquer forma do operador *OR* difuso. Os nós da camada 4 definem os termos linguísticos de cada variável de saída, representados por funções de pertença gaussianas, similares aos da camada 2. A última camada, camada 5, é a de saída. A sua função é realizar a *desfuzzificação*, i.e., converter os valores difusos nos valores reais do processo. Neste trabalho, método do centro da área é utilizado para essa conversão.

Tal como foi referido anteriormente, o objectivo da rede neuronal difusa apresentada é o de otimizar os centros e larguras das funções de pertença gaussianas. Para tal, implementa-se um algoritmo de aprendizagem supervisionada, baseado nos dados adquiridos (4). Lin [2] apresenta uma descrição detalhada da aplicação da retropropagação às RNDs.

### 3. PLANTA DE BRANQUEAMENTO

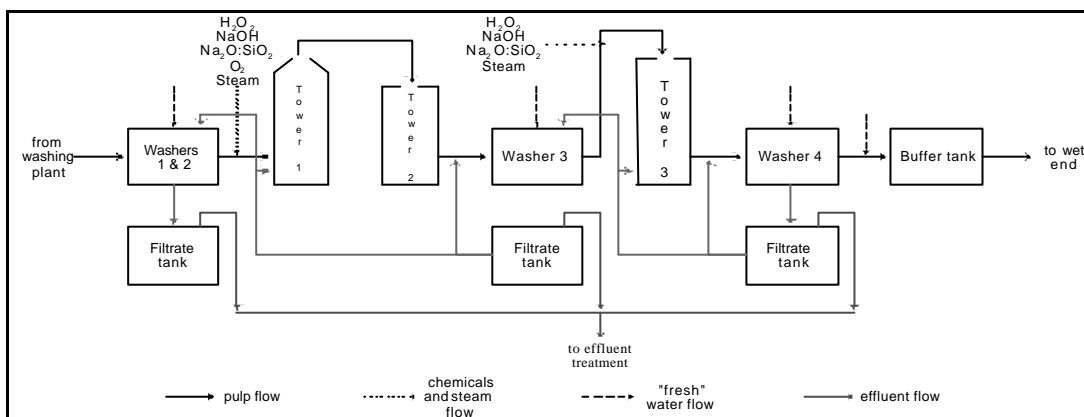


Figura 2: Esquema da secção de branqueamento da Companhia de Celulose do Caima, S.A.

A planta de branqueamento da Companhia de Celulose do Caima, S.A., objecto de modelação, é TCF (Totally Chlorine Free), sendo composta por dois estágios correspondentes a 3 torres. No primeiro estágio, dá-se a extracção (E) com Hidróxido de Sódio e a oxidação com Oxigénio (O) e com Peróxido de Hidrogénio (P). No segundo estágio ocorre apenas uma extracção e uma oxidação com Peróxido de Hidrogénio. Trata-se, pois, de uma sequência  $EOP_{NaOH} EP_{NaOH}$  (Figura 2).

A pasta proveniente da secção de Depuração com consistência de aproximadamente 4% é lavada em dois filtros rotativos em funcionamento paralelo com filtrado proveniente de um terceiro filtro situado a jusante e água limpa. A consistência da pasta à entrada dos filtros é de 1%, sendo de cerca de 12% à saída. O filtrado desta lavagem é parcialmente recirculado para os próprios lavadores, sendo o restante enviado à estação de tratamento de

efluentes. Após esta primeira lavagem a pasta é misturada com os agentes branqueadores adicionados ao primeiro estágio (Hidróxido de Sódio, Peróxido de Hidrogénio e Oxigénio). Simultaneamente, a sua temperatura é aumentada para cerca de 80 °C pela adição de vapor directo, após o que a mistura é bombeada para duas torres em série, onde as reacções de branqueamento (coloração dos cromóforos e corte das cadeias coradas) ocorrem em maior extensão. A primeira torre tem um tempo de retenção de 180 minutos, enquanto na segunda ele é de 60 minutos. Após isto, a pasta é lavada no Lavador 3 usando como corrente de lavagem o filtrado de um outro Lavador situado a jusante. O filtrado do Lavador 3 destina-se à lavagem nos primeiros dois filtros, situados antes do primeiro estágio de branqueamento. Esta recirculação visa recuperar água, mas sobretudo os agentes químicos branqueadores no estágio EP. Após esta lavagem a pasta é misturada com Peróxido de Hidrogénio e Hidróxido de Sódio, sendo igualmente a sua temperatura aumentada, pelo uso de vapor directo, para cerca de 80°C. A pasta é a seguir enviada à terceira torre, cujo tempo de residência é de 120 minutos. Da torre 3 a pasta é bombeada para o Lavador 4 e lavada com filtrado proveniente da máquina de formação. O filtrado desta lavagem é enviado ao Lavador 3, enquanto que a corrente de pasta é enviada à Máquina húmida e posteriormente à Secagem, onde a sua humidade é reduzida para cerca de 10%

#### 4. RESULTADOS DE SIMULAÇÃO

As técnicas apresentadas são aplicadas à modelização da variável Brancura na pasta em folha produzida pela C.C. Caima.

Algumas das variáveis medidas não são suficientemente excitadoras. Assim, a contribuição dessas variáveis para a qualidade final do modelo não é facilmente avaliada utilizando-se unicamente as amostras de dados recolhidas. Adicionalmente, e de acordo com informações dos peritos, as variáveis mais importantes são o caudal de peróxido de hidrogénio, o brilho à entrada e o pH. Este conhecimento baseado em experiência foi incorporado na modelização, sendo estas as variáveis utilizadas para descrever a dinâmica do sistema. De notar, que foram realizadas algumas experiências com todas as variáveis disponíveis, o que não melhorou os resultados finais obtidos.

O sistema de inferência difuso obteve-se a partir dos dados de entrada-saída do processo utilizando-se agrupamento difuso e sintonizando-se as funções de pertença como foi descrito na Secção 2.2.1. O intervalo de amostragem foi definido na fábrica como sendo de 1 hora; este intervalo parece suficiente, uma vez que o sistema apresenta uma dinâmica bastante lenta. Foram realizadas simulações com N=400 exemplos de treino. Ao parâmetro  $r_a$  foi atribuído o valor 1.7, o que originou 54 regras difusas. Na Figura 3 são apresentados os resultados do treino. Os resultados de validação do modelo são apresentados na Figura 4. Para a validação, obteve-se para o índice *root mean square error* o valor 0.4.

Do gráfico da Figura 4 transparece que o modelo obtido não apresenta capacidades de generalização satisfatórias. Tal facto poderá ser atribuído a ruído demasiado elevado nas medições da Brancura da pasta em folha e, principalmente, a inconsistências nos conjuntos de treino e de validação, resultantes da variação do atraso do sistema. O tempo total de residência da pasta varia entre 7 e 10 horas (dependendo do caudal de entrada). A técnica descrita não permite tratar este fenómeno adequadamente. Assim, está a ser considerada uma estratégia que consiga capturar o efeito da variabilidade do atraso do sistema.

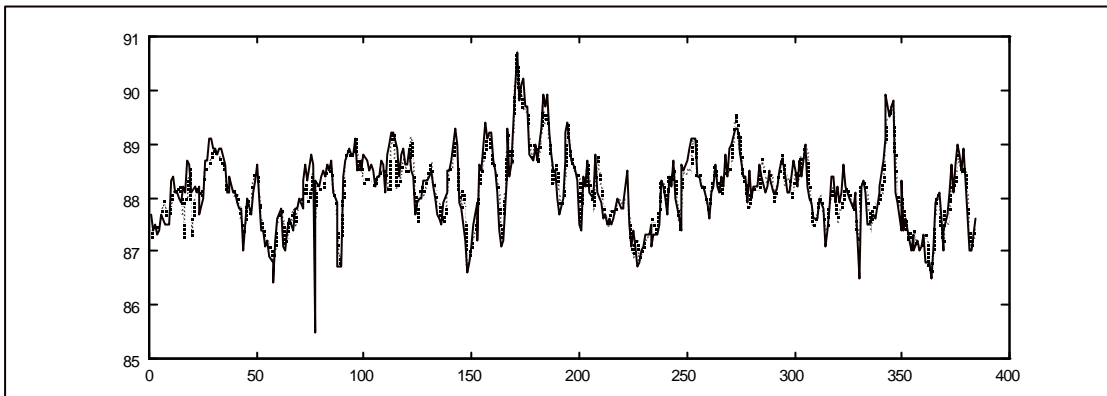


Figura 3. Treino: — Dados reais do processo - - - Saída da RND

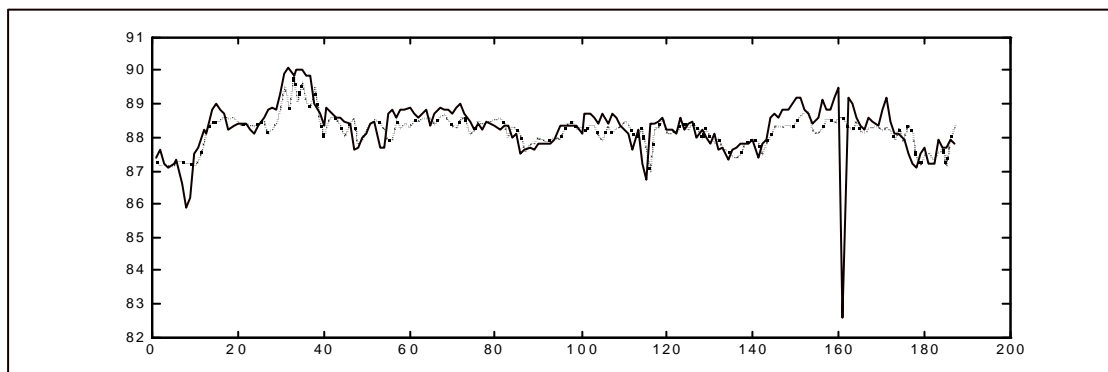


Figura 4. Validação: — Dados reais do processo - - - Saída da RND

## 5. CONCLUSÕES

Neste artigo foi apresentada uma estratégia de modelização neuro-difusa. O modelo é obtido em duas fases. Na primeira, é implementado o algoritmo de agrupamento subtractivo, de modo a obter-se um conjunto de regras difusas. Seguidamente, treina-se uma rede neuronal difusa com o fim de ajustar os parâmetros das funções de pertinência. Esta técnica foi aplicada à modelização da brancura de pasta de papel em folha. No entanto, foram encontrados alguns problemas que limitaram a precisão do modelo obtido. A origem desses problemas parece vir, fundamentalmente, do atraso variável do sistema, que não



é capturado pelo modelo. Presentemente está a ser desenvolvido um modelo tendo em conta o atraso variável. O atraso será estimado em cada instante de amostragem, de modo a que o brilho final possa ser previsto com maior precisão na base temporal adequada. Esta aproximação requer um esquema de aprendizagem em linha, que está também a ser investigado presentemente [3].

## 6. REFERÊNCIAS

- [1] - Chiu, S. 1994. "Fuzzy model identification based on cluster estimation", *Journal of Intelligent and Fuzzy Systems*, **Vol. 2, No. 3**.
- [2] - Lin, C.-T. 1995. "A neural fuzzy control scheme with structure and parameter learning", *Fuzzy Sets and Systems*, **No. 70**, 183-212.
- [3] - Lin, C.-T.; Lee, C. 1995. "Fuzzy adaptive learning control network with on-line neural learning", *Fuzzy Sets and Systems*, **No. 71**, 25-245.