

# MONITORAMENTO DE PARTÍCULAS EM CALDEIRA DE RECUPERAÇÃO KRAFT POR *MACHINE LEARNING*

**Autores:** Elisa C. Werneck do Carmo<sup>1</sup>, Andréa P. Parente<sup>1</sup>, Maurício B. de Souza Jr.<sup>1\*</sup>, Flavio Waltz<sup>2</sup>

<sup>1</sup>EQ, UFRJ; <sup>2</sup>Radix Engenharia & Software S.A.

## RESUMO

A caldeira de recuperação é um equipamento-chave nas plantas de produção de papel e celulose pelo processo Kraft. A deposição de material particulado formado na caldeira representa um importante problema operacional, devendo ser monitorada e controlada para prevenção de paradas operacionais não programadas. O processo de formação de partículas nesse equipamento é complexo e modelos fenomenológicos são de difícil desenvolvimento. Uma alternativa promissora, que se insere no contexto da **Indústria 4.0**, é a aplicação de técnicas de aprendizado de máquinas, ou *Machine Learning*, para modelagem de tal fenômeno. O objetivo do presente trabalho foi propor uma metodologia para o emprego de redes neurais artificiais no monitoramento da formação de material particulado em uma caldeira de recuperação. Foram utilizadas redes do tipo mapeamento auto-organizável para o desenvolvimento de modelos classificatórios a partir de dados históricos de um ano de operação em uma planta industrial. Foi obtido um mapa auto-organizável, que permitiu a identificação de regiões de maior probabilidade de operação anormal. Esse método pode ser usado para identificar condições com maior potencial de risco operacional.

**PALAVRAS-CHAVES:** monitoramento; redes neurais; caldeira de recuperação; processo Kraft; *machine learning*.

## INTRODUÇÃO

A necessidade das indústrias de aumentar a sua produtividade, dinamizar a sua produção, adaptando-as às exigências dos clientes e ao aumento de sua competitividade, está levando à quarta revolução, chamada de **Indústria 4.0**. Tal revolução industrial se baseia no conceito de *smart factories*, com elevada automação, eficiência e integração.

A **Indústria 4.0** engloba tecnologias dotadas da capacidade de interagir com o ambiente em seu entorno e de resolver problemas, buscando soluções e tomando decisões de forma autônoma. Diversos conceitos são essenciais para compreender a complexidade e os potenciais dessa nova revolução industrial, destacando-se o de *Machine Learning* ou aprendizagem de máquinas.

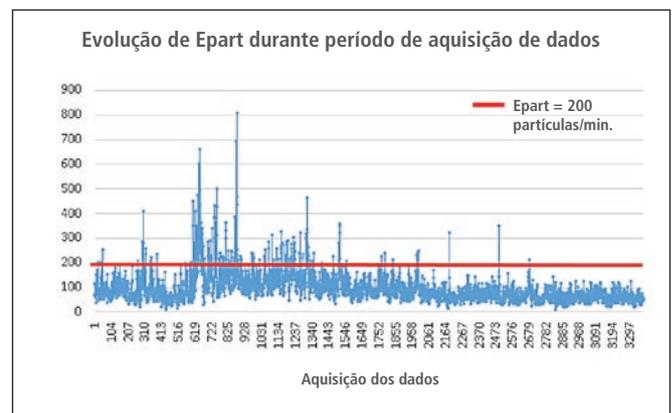
As técnicas de *Machine Learning* são técnicas de inteligência artificial, baseadas no aprendizado a partir de exemplos, que são usadas no reconhecimento de padrões. Dessa forma, máquinas são capazes

de aprender, de se adaptar e de agir de forma autônoma a partir dos dados de processo aos quais elas têm acesso (NILSSON et al., 2005). O presente estudo propõe o uso de *Machine Learning* na busca de soluções inovadoras para um dos desafios enfrentados pela indústria de papel e celulose: o monitoramento da formação de partículas em uma caldeira de recuperação.

Durante a queima de licor negro nas caldeiras de recuperação ocorre a formação de material particulado, que pode ser arrastado pelos gases de exaustão e depositado nas superfícies de troca térmica. Esta deposição provoca perda de eficiência de troca térmica e até mesmo bloqueio do escoamento dos gases, tornando necessária a realização de paradas não programadas. Assim, a quantidade de partículas formadas é um parâmetro que auxilia na avaliação da operação da caldeira e o seu monitoramento tem o potencial de reduzir a frequência de paradas para manutenção.

A Figura 1 ilustra, com os dados empregados no presente estudo, a formação de material particulado, variável *Epart*, ao longo de um ano, em uma caldeira industrial. Detalhes sobre a obtenção desses dados serão apresentados adiante. Considerando que valores de *Epart* acima de 200 partículas/min correspondem a condições operacionais anormais (COSTA et al., 2005), percebe-se que este limite, representado pela linha vermelha horizontal, é ultrapassado em vários momentos da operação da caldeira.

O objetivo geral do presente estudo é, portanto, propor uma me-



**Figura 1.** Evolução de Epart durante o período de aquisição de dados

**Autor correspondente:** UFRJ, Escola de Química, CT, Av. Athos da Silveira Ramos, 149, Bloco E - Sala E-201 - Ilha do Fundão, Rio de Janeiro-RJ, Brasil CEP 21941-909, Caixa Postal 68542, e-mail: mbsj@eq.ufrj.br

metodologia com base em técnicas de *Machine Learning*, mais especificamente redes neurais, para prever a formação de material particulado em uma caldeira industrial de recuperação Kraft a partir de dados operacionais. Foram adotados aqui mapas auto-organizáveis – redes SOM ou *Self Organizing Maps* – a fim de fazer a análise de *cluster* ou agrupamentos dos dados.

Assim, o mapa deverá inferir, tendo por base dados operacionais da caldeira, se a formação de partículas está elevada (operação anormal) ou não (operação normal). Dessa forma, mesmo quando os sensores empregados na contagem de partículas estiverem em falha ou em manutenção, o operador terá um apoio operacional para avaliar a condição do processo.

Este artigo está organizado como segue. A motivação e os objetivos são apresentados nesta Introdução. A seguir, são descritos brevemente os fundamentos da caldeira de recuperação Kraft, com ênfase na formação de material particulado, e os dados empregados no estudo. Na sequência, é descrita a metodologia de diagnóstico da condição operacional baseada em redes SOM. Por fim, são apresentados os resultados e as conclusões. O Apêndice 1 expõe as equações empregadas no treinamento das redes.

## Caldeira de Recuperação Kraft

### 1. O Processo

O processo Kraft é o mais empregado na produção de polpa celulósica (COSTA, 2004). Nele, a caldeira de recuperação é responsável pela geração de vapor e energia elétrica para a planta a partir da queima do licor negro, assim como pela recuperação de reagentes não consumidos na digestão, o que permite a redução de impactos ambientais.

O licor negro, subproduto do digestor de deslignificação, é uma solução altamente alcalina. Ele é obtido após o processamento da polpa marrom, que separa o licor negro da polpa de celulose. Sua composição é complexa: além de conter a lignina e os açúcares extraídos da madeira, material orgânico combustível, apresenta os sais inorgânicos, que resultam da reação de deslignificação (BOUCARD, 2014).

As propriedades e a composição do licor negro variam em cada planta de acordo com as condições operacionais estratégicas de cada empresa. Essas diferenças são atribuídas, principalmente, aos seguintes fatores: condições operacionais do digestor, espécie da madeira utilizada, razão entre as quantidades de reagentes químicos alimentados ao digestor e a quantidade de madeira, propriedades do licor branco, rendimento da polpa e eficiência da lavagem do material descarregado do digestor.

A recuperação dos sais inorgânicos contidos no licor negro é crítica para reduzir os custos e o impacto ambiental do processo (GALDIANO, 2006). Para esta finalidade, foram desenvolvidas caldeiras de recuperação. Com isso, o descarte dos subprodutos é reduzido e há um ganho importante em termos ambientais. Uma caldeira de recuperação Kraft típica se divide basicamente em duas seções. Na

primeira seção, denominada fornalha, ocorre a queima do licor negro. A segunda seção é uma região de transferência de calor por convecção, na qual fica localizado um sistema de troca térmica para aproveitamento de calor, e compreende os seguintes componentes: superaquecedores, feixe convectivo, tubulões e economizadores. O circuito de água de uma caldeira de recuperação é muito semelhante ao das caldeiras aquatubulares utilizadas em geração de vapor em termoeletricas e na indústria de petróleo e gás. Assim, o detalhamento da segunda seção não será feito aqui, podendo ser encontrado em COSTA (2004).

A fornalha é uma seção muito específica da indústria de papel e celulose. É nela que se remove a umidade remanescente do licor negro e que se faz a combustão dos compostos orgânicos e a redução dos sais de sódio de forma a recuperar o sulfeto de sódio em forma fundida.

Pode-se dividir a fornalha em 3 zonas para melhor compreensão dos processos físico-químicos que nela ocorrem. O licor negro, contendo sólidos em suspensão, é alimentado na caldeira, na 1ª zona, por meio de uma série de bocais de aspersão na forma de gotículas com tamanhos controlados. A 1ª zona da fornalha é uma região de redução, na qual ocorrem, predominantemente, as etapas de secagem, pirólise e queima do carbono fixo do licor negro. A zona 2, localizada abaixo da primeira, é uma região de redução, compreendendo entradas de ar primário e ar secundário e saídas dos sais inorgânicos recuperados na forma de uma massa fundida. A zona 3, localizada acima da 1ª zona, é uma região de oxidação, com injeção de ar terciário.

É inevitável que exista o arraste de pequenas partículas sólidas pelos gases de exaustão, além da formação de sais na parte superior da fornalha, que seguem o fluxo de gases de exaustão. Este arraste pode causar diversos problemas, como deposição na superfície externa dos tubos de água, chegando a causar bloqueio da passagem dos gases e redução da eficiência da troca térmica, além do custo de remoção dessas partículas no separador eletrostático na saída da caldeira. O progresso das incrustações é acompanhado, principalmente, pelo monitoramento das seguintes variáveis operacionais: queda de pressão, temperatura do gás de exaustão ao longo da caldeira e temperatura do vapor produzido (COSTA, 2004).

A quantidade de partículas que sai da fornalha pode ser medida por instrumentos localizados na região de entrada dos superaquecedores. Em algumas instalações são utilizadas câmeras e processamento de imagens para contar o número de partículas que deixam a fornalha por unidade de tempo. Os dados usados no presente estudo são referentes a uma caldeira dotada de tal sistema de medição, em que existem dois sensores localizados no lado direito e no lado esquerdo da fornalha pouco abaixo dos superaquecedores.

Esses sensores apresentam baixa confiabilidade, dadas as condições rigorosas na caldeira. Assim, é necessária uma abordagem de inferência para o caso de falhas ou manutenção nos sensores. Sensores virtuais podem ser usados para inferir o número de par-

tículas a partir das condições de operação da caldeira. As RNA se constituem em uma excelente alternativa para o desenvolvimento desses sensores (WALTZ et al., 2017).

## 2. Os Dados Operacionais

Foram usados os dados operacionais tratados por Costa (2004) de uma caldeira de recuperação Kraft de uma fábrica de papel e celulose brasileira. Os dados correspondem a 12 meses de operação, incluem 12 variáveis de processos e 3.381 pontos operacionais (ou conjuntos de dados). O nome da empresa será omitido por razões de confidencialidade.

os dados foram obtidos no histórico do sistema de supervisão e aquisição de dados da indústria. Todas as variáveis empregadas na análise estão apresentadas na Tabela 1.

O número de partículas formadas, fornecido pela variável Epart, foi exibido na Figura 1. A medição é feita por câmeras e posterior processamento de imagem. A quantidade de partículas, de acordo com os dados disponíveis, variou entre 9,62 e 806,12 (valores mínimo e máximo, respectivamente). O valor médio de Epart é igual a 96,99 e o desvio padrão a 61,01.

A análise dos dados mostrou que a operação apresenta grandes variações do valor de Epart, assim como das demais variáveis, e que os valores de Epart não são distribuídos uniformemente. A distribuição não uniforme é esperada, já que os dados foram obtidos durante operação rotineira da caldeira com objetivo de acompanhar o processo e não de realizar uma varredura experimental de todos os valores possíveis de Epart. A menor quantidade de pontos com Epart elevado (acima de 200) também é esperada e evidencia que a caldeira funciona, na maior parte do tempo, dentro das condições desejadas em relação à formação de material particulado.

A Tabela 2 fornece os valores máximos e mínimos, a média e o desvio padrão de cada variável. As variáveis que apresentaram maiores valores de desvio padrão são a pressão de alimentação do ar secundário, Par2, e a quantidade de partículas, Epart. Ademais, foi possível observar que mudanças significativas ocorreram em relação às variáveis Var1 e Var2 a partir, aproximadamente, do conjunto de dados de número 884 e em relação à variável Par2, que sofreu um aumento a partir dos dados de número 1380, caracterizando uma mudança na forma de operação da caldeira.

**Tabela 1.** Variáveis de processo. Adaptado de COSTA, 2004

| TAG   | Descrição   | Local de medição   | Unidades            |
|-------|---|--|---------------------|
| Vvlu  | Vazão volumétrica de licor negro alimentado à caldeira. | Alimentação de licor negro.  | m <sup>3</sup> /h   |
| x     | Concentração de sólidos na alimentação do licor negro.  | Alimentação de licor negro.  | %                   |
| Tlu   | Temperatura de alimentação do licor negro.              | Alimentação de licor negro.  | °C                  |
| Var1  | Vazão de ar primário.                                   | Alimentação de ar primário.  | t/h                 |
| Var2  | Vazão de ar secundário.                                 | Alimentação de ar secundário.  | t/h                 |
| Var3  | Vazão de ar terciário.                                  | Alimentação de ar terciário.   | t/h                 |
| Par1  | Pressão de alimentação do ar primário.                  | Alimentação de ar primário.  | mmca                |
| Par2  | Pressão de alimentação do ar secundário.                | Alimentação de ar secundário.  | mmca                |
| Tar1  | Temperatura de alimentação do ar primário.              | Alimentação de ar primário.  | °C                  |
| Tar2  | Temperatura de alimentação do ar secundário.            | Sistema de alimentação de ar secundária.   | °C                  |
| PI    | Pressão de alimentação do licor negro.                  | Sistema de alimentação de licor negro.   | Kgf/cm <sup>2</sup> |
| Epart | Número médio de partículas por minuto.                  | Média entre as quantidades medidas na entrada da região dos superaquecedores no sentido do fluxo dos gases do lado direito e esquerdo. | partículas/min      |

**Tabela 2:** Estatísticas das variáveis de entrada

|               | Vvlu  | x     | Tlu    | Var1   | Var2   | Var3  | Par1   | Par2   | Tar1   | TAr2   | PI   | Epart  |
|---------------|-------|-------|--------|--------|--------|-------|--------|--------|--------|--------|------|--------|
| Valor Mínimo  | 25,09 | 70,71 | 127,15 | 82,35  | 99,64  | 18,02 | 12,72  | 123,31 | 120,37 | 109,34 | 1,76 | 9,62   |
| Valor Máximo  | 74,73 | 87,39 | 143,46 | 156,1  | 143,61 | 93,07 | 247,78 | 510,71 | 165,23 | 205,94 | 4,31 | 806,12 |
| Média         | 63,81 | 81,46 | 140,23 | 122,09 | 129,14 | 65,82 | 112,91 | 324,84 | 153,36 | 141,24 | 3,91 | 96,99  |
| Desvio padrão | 3,96  | 2,24  | 1,07   | 14,49  | 8,13   | 7,37  | 20,53  | 103,08 | 4,61   | 13,73  | 0,23 | 61,01  |

## METODOLOGIA

## 1. Conceitos Preliminares: As Redes SOM

A natureza complexa da formação de material particulado na caldeira dificulta o desenvolvimento de modelos matemáticos convencionais a partir de métodos analíticos e fenomenológicos, tornando interessante, como alternativa, o uso de redes neuronais artificiais (RNA), ou, em inglês, *Artificial Neuronal Networks* (ANN).

Costa et al. (2005) modelaram a formação de material particulado de uma caldeira industrial a partir de redes neuronais multicamadas. Foram testadas redes preditivas e classificadoras. O presente trabalho propõe uma abordagem alternativa, pelo uso das redes SOM. Essas redes diferem das redes multicamadas, por terem uma aprendizagem não supervisionada, ou seja, não precisam de exemplos de saída do processo para realizar sua tarefa, mas são capazes de reconhecer características nos dados de entrada que permitem seu agrupamento e posterior classificação. Isso confere às redes SOM um enorme potencial na identificação de condições novas, podendo assistir o operador na automação do monitoramento.

Como as redes multicamadas são bastante conhecidas na literatura (DE SOUZA JR., 1993) e não serão empregadas neste trabalho, elas não serão revisadas aqui. Os fundamentos da rede adotada neste estudo são apresentados na sequência.

O princípio de funcionamento do mapeamento auto-organizável ou SOM (do inglês, *Self-Organized Map*) é baseado na organização do córtex humano, em que cada região é excitada por sensações diferentes e específicas (ROSA e VALLE, 2012). Dessa forma, em uma rede de mapeamento auto-organizável, entradas diferentes excitam neurônios diferentes de tal modo que neurônios que se encontram próximos uns aos outros são ativados por padrões (entradas) apresentando similaridades.

Ao contrário das redes multicamadas, a SOM é uma rede de camada única em forma de grade uni ou bidimensional, na qual os neurônios se posicionam lado a lado, formando vizinhanças. A camada única é, portanto, uma representação da saída de rede de mapeamento auto-organizável. O treinamento das redes SOM é apresentado no Apêndice 1.

## 2. Classificação do Número de Partículas

A abordagem proposta foi o uso de redes SOM para a classificação dos dados em função do número de partículas. O estudo do uso de redes para classificação se baseou na separação dos dados em 7, 5, 4, 3 e 2 classes. As classes foram estabelecidas conforme proposto por Costa (2004) e suas definições são apresentadas na Figura 2.

## 3. Análise de Cluster

Neste trabalho, utilizou-se a ferramenta *Automated Neural Networks* (SANN) do pacote *Data Mining* do software STATISTICA® versão 8.0. Foram avaliadas redes bidimensionais de mapeamento auto-organizável com diferentes topologias. Variou-se o número de neurônios de 4 a 180, assim como a sua disposição na rede, e foram testadas redes com as seguintes topologias: 2x2, 3x3, 1x100, 3x40,

| Faixa Epar | Conjuntos de dados | Classe e (conjuntos de dados na classe) |             |             |             |             |
|------------|--------------------|---|-------------|-------------|-------------|-------------|
|            |                    | A                                       | B           | C           | D           | E           |
| 0 - 50     | 481                | 1<br>(481)                              | 1<br>(481)  | 1<br>(481)  | 1<br>(2236) |             |
| 50 - 100   | 1755               | 2<br>(1755)                             | 2<br>(1755) | 2<br>(1755) |             | 1<br>(3184) |
| 100 - 150  | 710                | 3<br>(710)                              | 3<br>(710)  | 3<br>(710)  | 2<br>(948)  |             |
| 150 - 200  | 238                | 4<br>(238)                              | 4<br>(238)  |             |             |             |
| 200 - 250  | 106                | 5<br>(147)                              |             |             |             |             |
| 250 - 300  | 41                 |   |             | 4<br>(435)  |             |             |
| 300 - 350  | 25                 | 6<br>(34)                               | 5<br>(197)  |             | 3<br>(197)  | 2<br>(197)  |
| 350 - 400  | 9                  |   |             |             |             |             |
| >400       | 16                 | 7<br>(16)                               |             |             |             |             |

Figura 2. Definição das classes estabelecidas e estudadas na classificação do número de partículas

3x60, 4x10, 5x7, 5x8, 5x20, 6x6, 10x10 e 20x5 (em que o primeiro número indica o número de linhas, e o segundo, o número de colunas da matriz formada pelos neurônios, e o produto desses valores fornece o total de neurônios da rede).

Foram adotados os valores recomendados pelos desenvolvedores do software para os demais parâmetros da rede, que se encontram resumidos na Tabela 3.

Tabela 3. Parâmetros utilizados para a geração de redes de mapeamento auto-organizável.

| Parâmetros                             | Valores definidos   |
|--|---|
| Número de redes geradas a cada análise | 1   |
| Amostra de treinamento                 | 80% do conjunto de dados  |
| Amostra de teste                       | 20% do conjunto de dados  |
| Amostra de validação                   | 0% do conjunto de dados   |
| Número de neurônios                    | De 4 a 180  |
| Entradas                               | Vvlu, x, Tlu, Var1, Var2, Var3, Par1, Par2, Tar1, Tar2, PI            |
| Ciclos (épocas) de treinamento         | 1.000   |
| Taxas de aprendizado                   | De 0,1 a 0,02   |
| Condições de parada de treinamento     | A partir de 10 ciclos, quando a melhora na performance é de $10^{-7}$ |
| Vizinhança                             | Inicia-se com 3 neurônios e finaliza com 0                            |
| Inicialização dos pesos                | Distribuição normal, valores aleatórios                               |

A rede gera um mapa topológico, escolhido retangular, em que cada neurônio corresponde a um elemento, e cada conjunto de dados é associado a um neurônio.

## RESULTADOS

Estudou-se o problema utilizando redes neuronais de mapas auto-organizáveis em duas etapas. A análise das redes geradas se baseou na capacidade de agrupar dados com características similares em regiões visualmente identificáveis.

Em uma primeira etapa, para a avaliação de cada rede, foi realizada uma análise baseada na divisão dos dados em duas classes: a classe 1 inclui os conjuntos de dados em que Epart é menor ou igual a 200 e a classe 2 reúne os conjuntos de dados em que Epart é maior que 200 (caso E na Figura 2). A análise consistiu em identificar regiões em que há uma predominância de dados de classe 2, ou seja, de operação anormal.

Para tal, foram identificados os neurônios vencedores para cada conjunto de dados. Para cada neurônio, foi calculado o percentual de casos de cada classe que nele foram agrupados. Assim, foi possível buscar regiões do mapa em que houve grande concentração de determinada classe. Por exemplo, na rede SOM 20x5, o neurônio 1 foi o vencedor 42 vezes, das quais 17 correspondem a dados de classe 1 e 25 da classe 2. A partir desses resultados, calculou-se a porcentagem de dados de classe 2 que foram agrupados no neurônio 1. Nesse caso, 60% dos dados agrupados no neurônio 1 são de classe 2 e, portanto, este neurônio foi considerado como sendo majoritariamente pertencente a uma região da "classe 2".

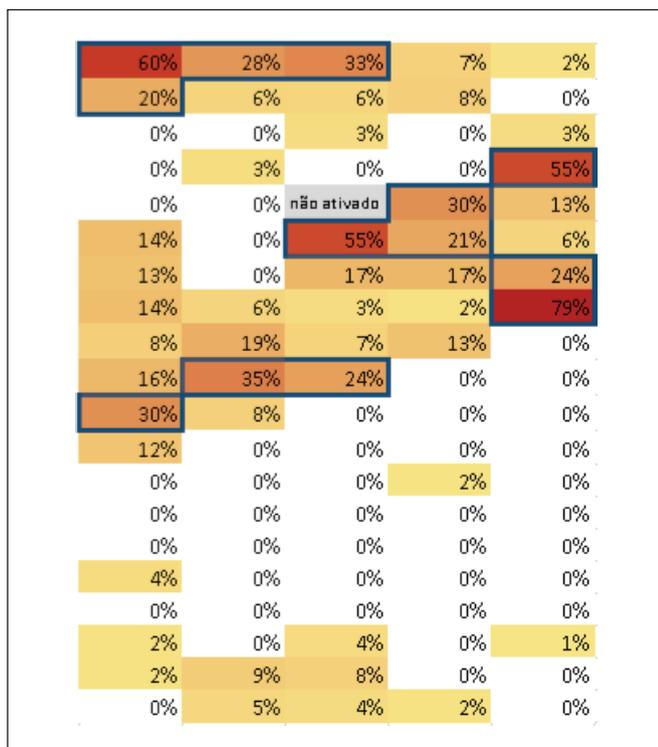


Figura 3. Mapa auto-organizável com 100 neurônios dispostos em uma matriz 20x5, indicando regiões de grande incidência de casos de classe 2

A partir das porcentagens calculadas para os dados de classe 2 em cada neurônio, foram construídos mapas evidenciando esses valores de modo a permitir a identificação visual de regiões com maior incidência de casos de classe 2. Uma escala de cor foi utilizada para facilitar a visualização, cores mais escuras correspondem a zonas com mais casos de classe 2.

A Figura 3 apresenta o mapa criado para a rede 20x5. Nela, cada valor percentual corresponde a um neurônio e a contagem do número do neurônio é realizada a partir da primeira linha (na primeira linha estão localizados os neurônios de 1 a 5, na segunda, de 6 a 10, e assim por diante). Análises análogas foram feitas para as demais configurações da rede, mas não serão exibidas aqui por questão de espaço e porque os resultados para o mapeamento 20x5 foram considerados satisfatórios e ilustrativos.

Os neurônios com porcentagens nulas foram ativados somente por dados de classe 1. Existem 4 neurônios que agrupam majoritariamente dados de classe 2 (mais de 50%) e há pelo menos 3 regiões que agrupam porcentagens superiores a 20%. Os quatro neurônios majoritários da rede 20x5 capturam 27% dos dados da classe 2.

A Figura 4 fornece uma visualização qualitativa do mapa gerado pela rede, em que as diferentes regiões são indicadas com cores diferentes. As regiões em verde indicam elevada probabilidade de operação sem problemas de emissão de partículas. As regiões em vermelho, pelo contrário, indicam possibilidade de emissão acima de 200 partículas/min. Os quatro neurônios majoritários estão destacados com vermelho mais escuro. As regiões amarelas podem ser entendidas como faixas de transição.

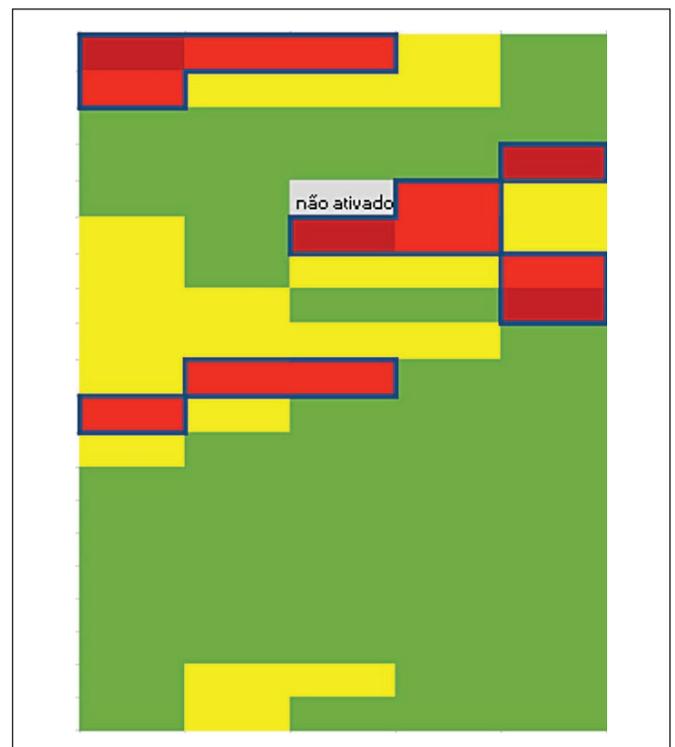


Figura 4: Ferramenta para detecção de condições anormais baseada na rede 20x5

A região em verde da Figura 4 concentra 72% de todos os conjuntos de dados, dos quais apenas 0,6% pertencem à classe 2. A região em vermelho agrupa somente 9% dos dados de operação, 61% dos quais pertencem à classe 2. Isso mostra que se um dado inédito for agrupado pela rede nesta região, haverá grande probabilidade de ele pertencer à classe 2. A região em amarelo agrupa 19% dos dados operacionais, dos quais 10% pertencem à classe 2. Percebe-se, então, que o mapa permite classificar os dados, porém existe uma faixa de indeterminação. Por um lado, uma parte dos dados de classe 2 não são indicados na região em vermelho (falsos negativos) e, por outro, alguns dados de classe 1 são classificados como operação anormal (falsos positivos).

Dessa forma, o modelo gerado a partir de uma rede de mapeamento auto-organizável permitiu a geração de uma ferramenta visual para auxiliar o operador na identificação de regiões de maior probabilidade de operação anormal.

Em uma segunda etapa da análise, a rede teve seu comportamento avaliado quando os dados foram separados em 3, 4, 5 e 7 classes. O procedimento utilizado para este estudo foi similar ao reportado acima, entretanto, como a complexidade da análise aumenta com o número de classes, em vez de serem criados mapas contendo os valores percentuais de uma determinada classe na posição de cada neurônio, optou-se por indicar qual a classe que ativou majoritariamente aquele neurônio. Foi estabelecido um valor percentual mínimo de 50% para que uma classe fosse considerada majoritária em um dado neurônio. Caso nenhuma classe fosse preponderante, o neurônio recebia o valor "0". Essa metodologia é chamada de "proporção mínima com vizinhança por Voronoi" (DE SOUZA JR., 2003).

Foram obtidos resultados como os ilustrados na Figura 5, para classificação em três classes (caso D na Figura 2).

Comparando o mapa da Figura 5 com o da Figura 4, percebe-se que a região de transição (em amarelo) é bastante semelhante à região do mapa de três classes, na qual predomina a classe 2. Portanto, o resultado para duas e três classes pode ser entendido como o mesmo resultado expresso de forma diferente. Entretanto, a análise com base em três classes apresenta uma desvantagem pela presença de zonas indefinidas, indicadas com o número zero na Figura 5. A presença de zonas indefinidas ocorre porque é mais difícil um neurônio concentrar mais de 50% de acertos numa dada classe para um número de classes maior.

Para a análise em mais classes, há um aumento da complexidade do mapa com poucos ganhos em relação à melhoria dos resultados; por isso esses mapas não serão exibidos aqui.

## CONCLUSÕES

Foi obtida uma rede neuronal artificial de topologia 20x5 em um mapa auto-organizável (SOM) que permite a identificação de regiões de operação normal, regiões de transição e regiões de maior probabilidade de operação anormal, ou seja, com emissão de material particulado acima de 200 partículas/min em uma caldeira de recuperação Kraft. A região de operação normal concentra 72% de

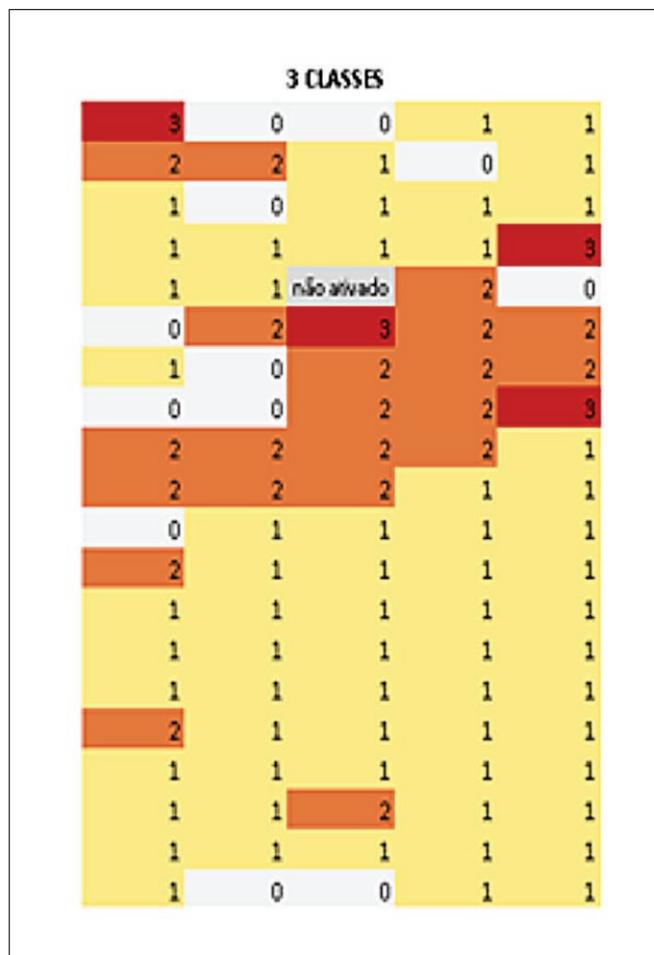


Figura 5. Comparação da rede de topologia 20x5 para análise em 3 classes

todos os conjuntos de dados, dos quais apenas 0,6% apresentam valores acima de 200 partículas/min. A região de maior probabilidade de operação anormal agrupa somente 9% dos dados de operação, 61% dos quais apresentam emissão de partículas superior ao limite operacional. A existência de dados de operação anormal localizados em regiões de operação normal gera falsos negativos, enquanto a situação inversa gera falsos positivos.

Embora a modelagem com base em redes SOM não tenha sido capaz de segregar completamente todos os pontos anormais de operação, o que limita seu uso para o controle do processo, este método pode ser usado para gerar recomendações referentes à segurança operacional, indicando que a operação pode estar comprometida.

Os resultados poderiam ser melhorados caso se dispusesse de uma distribuição uniforme dos dados nas classes, pois o número de exemplos de operação normal é muito maior do que os de falhas, o que prejudica o treinamento da rede neuronal. No entanto, a metodologia proposta pode ser empregada para bases de dados que podem ser aumentadas à medida que novos padrões de dados forem sendo armazenados. ■

## REFERÊNCIAS BIBLIOGRÁFICAS

- BOUCARD, H. **Contributions to the understanding of the hydrothermal processes: application to black liquor**. 2014. 309 f. Tese (Doutorado em Engenharia Química e de Processo) – École des Mines d’Albi-Carmaux, Universidade de Toulouse: Toulouse, França, 2014.
- CASTRO F. C. C.; CASTRO M. C. F. Mapas Auto-organizados de Kohonen – SOM. In: **Redes Neurais Artificiais: Curso de Engenharia Elétrica**. PUC-RS, Faculdade de Engenharia, Departamento de Engenharia Elétrica, Rio Grande do Sul, 2014. cap. 7, p. 1-19.
- COSTA, A. O. S. **Descrição matemática da caldeira de recuperação Kraft**. 2004, 150 f. Tese (Doutorado em Engenharia Química) – COPPE/UFRJ, 2004.
- COSTA, A. O. S.; SOUZA Jr., M. B.; BISCAIA Jr., E. C.; LIMA, E. L. Monitoring particulate material formation in a kraft furnace recovery boiler. **Tappi Journal**, Atlanta, Estados Unidos da América, v. 4, n. 4, p. E33-47, 2005.
- DE SOUZA Jr., M. B. **Redes neuronais multicamadas aplicadas à modelagem e controle de processos químicos**. 1993, 194 f. Tese (Doutorado em Engenharia Química) - COPPE/UFRJ, Rio de Janeiro-RJ, 1993.
- DE SOUZA Jr., M. B. **Redes neurais com Statística®**. In: Curso de treinamento para a Versão 6. Escola de Química da UFRJ. UFRJ, Rio de Janeiro, 2003, apostila, 39 p.
- GALDIANO, G. P. **Inventário do ciclo de vida do papel offset produzido no Brasil**. 2006. 280 f. Dissertação (Mestrado em Engenharia) – Escola Politécnica da USP, Departamento de Engenharia Química, São Paulo, 2006.
- NILSSON, N. J. **Introduction to machine learning**: An early draft of a proposed textbook. Robotics Laboratory, Department of Computer Service. Stanford University, Stanford, Califórnia, Estados Unidos da América, 2005, p. 188. Disponível em: <<http://robotics.stanford.edu/people/nilsson/mlbook.html>>. Acesso em: 10 jul. 2017.
- ROSA, R. A. F.; VALLE, M. E. Um estudo da aplicação de redes neurais auto-organizáveis para a identificação autônoma de fonemas portugueses. **Revista Ciências Exatas e Naturais**, v. 17, n. 2, p. 199-218, jul./dez. 2012.
- WALTZ, F.; DE SOUZA Jr. M. B.; MORADA, T.; FILHO L. Indústria 4.0 no aumento dos resultados da indústria de papel e celulose. **Webinar ABTCP: Indústria 4.0**. 2017. Disponível em: <<http://www.radixeng.com.br/docs/news/ApresentacaoABTCP.pdf>>. Acesso em: 20 jun. 2017.
- WASSERMAN, P. D., *Neural Computing: Theory and Practice*, Van Nostrand, Reinhold, 1989.

## APÊNDICE 1 – TREINAMENTO DA REDE SOM

O treinamento das redes SOM segue o treinamento de Kohonen (WASSERMAN, 1989) e compreende três etapas: competição, cooperação e adaptação sináptica (ROSA e VALLE, 2012). Uma etapa preliminar é necessária para a inicialização dos pesos e do raio da gaussiana. Wasserman (1989) sugere uma inicialização com valores aleatórios.

Após inicialização dos pesos, é começado o processo competitivo que tem por objetivo a definição do neurônio vencedor. Para tal, um padrão de entrada é alimentado à rede e é realizada a busca pelo neurônio que minimiza a distância em relação ao vetor de entrada. Assim, as distâncias são calculadas entre o vetor de entrada ( $\underline{x}$ ) e o vetor de pesos ( $\underline{w}_j$ ) correspondente a cada neurônio. Geralmente, utiliza-se a distância euclidiana. A distância ( $D_j$ ) do neurônio  $j$  em relação à entrada  $\underline{x}$  de dimensão ( $N \times 1$ ), sendo  $w_{ij}$  o peso entre a entrada  $i$  e o neurônio  $j$ , e  $n$  o número da iteração, é dada por:

$$D_j(n) = \|\underline{x} - \underline{w}_j\| = \sqrt{\sum_{i=1}^N (x_i(n) - w_{ij}(n))^2} \quad (1)$$

Para encontrar o neurônio vencedor, deseja-se encontrar a distância mínima, ou seja, o neurônio vencedor  $v(x)$  é dado pela Equação (2):

$$v(x) = \underset{j}{\operatorname{argmin}} \|\underline{x} - \underline{w}_j\| \quad (2)$$

Determinado o neurônio vencedor, dá-se o processo cooperativo. Uma função  $h_{j,v}$  define a vizinhança do neurônio vencedor, centrada nele próprio. Ela deve ser simétrica e atingir seu valor máximo no neurônio vencedor. Sua amplitude deve diminuir monotonicamente até ser nula conforme a distância lateral entre o vencedor e demais neurônios aumenta. Um exemplo de função que atende a essas exigências é a gaussiana. O centro da gaussiana é nulo ( $\mu=0$ ) para o neurônio vencedor. Para que o raio da vizinhança diminua com o tempo de treinamento, pode-se utilizar um desvio-padrão dinâmico, que diminui com o tempo. O decaimento pode ser linear ou exponencial.

Segue-se o processo adaptativo para o ajuste dos pesos dos neurônios na vizinhança topológica do neurônio vencedor (CASTRO e CASTRO, 2014):

$$\underline{w}_j(n+1) = \underline{w}_j(n) + \eta(n) \times h_{j,v(x)}(n) \times (\underline{x}(n) - \underline{w}_j(n)) \quad (3)$$

Uma vez que os pesos tenham sido ajustados, retorna-se ao início do processo iterativo com a apresentação de um novo padrão e as etapas de competição, cooperação e adaptação são repetidas até que não haja mais modificações significativas no mapa.

**AGRADECIMENTO:** Professor Maurício B. de Souza Jr. agradece ao CNPq (Processo No. 3106682016-6).