

UNIVERSIDADE FEDERAL DO RIO GRANDE DO SUL
ESCOLA DE ENGENHARIA
ENG. DE CONTROLE E AUTOMAÇÃO

LUIZ FREDERICO VILLALOBOS

**MODELO HÍBRIDO NEURONAL -
FENOMENOLÓGICO DE UM REATOR
INDUSTRIAL DE
HIDRODESSULFURIZAÇÃO**

Porto Alegre
2018

LUIZ FREDERICO VILLALOBOS

**MODELO HÍBRIDO NEURONAL -
FENOMENOLÓGICO DE UM REATOR
INDUSTRIAL DE
HIDRODESSULFURIZAÇÃO**

Trabalho de Conclusão de Curso (TCC-CCA)
apresentado à COMGRAD-CCA da Universidade
Federal do Rio Grande do Sul como parte dos re-
quisitos para a obtenção do título de *Bacharel em
Eng. de Controle e Automação* .

ORIENTADOR: Prof. Dr. Pedro Rafael Bolognese Fernandes

Porto Alegre
2018

LUIZ FREDERICO VILLALOBOS

**MODELO HÍBRIDO NEURONAL -
FENOMENOLÓGICO DE UM REATOR
INDUSTRIAL DE
HIDRODESSULFURIZAÇÃO**

Este Trabalho de Conclusão de Curso foi julgado adequado para a obtenção dos créditos da Disciplina de TCC do curso *Eng. de Controle e Automação* e aprovado em sua forma final pelo Orientador e pela Banca Examinadora.

Orientador: _____
Prof. Dr. Pedro Rafael Bolognese Fernandes, UFRGS
Doutor pela TU Dortmund – Dortmund, Alemanha

Banca Examinadora:

Prof. Dr. Marcelo Götz, UFRGS
Doutor pela Universität Paderborn – Paderborn, Alemanha

Prof. Dr. Renato Ventura Bayan Henriques, UFRGS
Doutor pela UFMG – Belo Horizonte, Brasil

Prof. Dr. Marcelo Farenzena, UFRGS
Doutor pela UFRGS – Porto Alegre, Brasil

Prof. Dr. Mario Sobczyk
Coordenador de curso
Eng. de Controle e Automação

Porto Alegre, dezembro de 2018.

DEDICATÓRIA

Dedico este trabalho aos meus colegas de curso, por todo carinho, ajuda e suporte que me deram ao longo de toda a minha caminhada na universidade.

AGRADECIMENTOS

Agradeço a Universidade Federal do Rio Grande do Sul pela formação que me foi dada e as diversas oportunidades que me geraram.

A minha família e amigos pelo suporte e paciência enquanto eu desenvolvia esse projeto.

A minha gata Florence, por ter estado do meu lado, ou atrás do meu monitor, durante todo o tempo que eu estive digitando ou programando.

RESUMO

Modelos de processo podem ser desenvolvidos dentro de um amplo espectro desde o totalmente fenomenológico, ou "caixa-branca", até o puramente baseado em dados, ou "caixa-preta", cada qual com suas vantagens e desvantagens em aplicações. Dentro desta perspectiva, este trabalho traz a proposta de uma estrutura de modelo híbrido neuronal-fenomenológico no qual um modelo fenomenológico e uma rede neuronal convencional, esta na forma de modelo do erro, são combinados de forma a gerar um único modelo superior em ambos em termos de capacidade de predição. O modelo fenomenológico tem parâmetros ajustados para uma ponto nominal/inicial de operação e o modelo do erro é treinado para se adaptar a efeitos não descritos nas equações do modelo, a variações nos parâmetros, etc. Através do estudo de caso real de um processo de hidrotreatamento, no qual se remove impurezas (compostos derivados do enxofre) do óleo diesel, os resultados da abordagem e suas vantagens são apresentados ao longo de uma janela de tempo de operação. É mostrado que o modelo híbrido é capaz de se adaptar mais rápido que uma rede neuronal comum, podendo ser implementada com menos dados e atingindo a mesma capacidade de predição que uma rede comum.

Palavras-chave: Redes Neurais, Modelagem de Processos, Modelo Híbrido, Reator de Hidrodessulfurização.

ABSTRACT

This work aims to show the development of a hybrid neural network, where a phenomenological model is merged with a neural network in order to produce a single model. The advantages of each one of the models are also presented, and their respective predictions compared. It is shown, that the hybrid neural network is capable of learning faster than the regular neural network, and as such, can be implemented with fewer data and with no loss of performance. This is done for a hydrotreatment plant, where diesel impurities, sulfur in particular, are removed. The results from this approach and its advantages are also shown.

Keywords: Neural Networks, Process Modeling, Hybrid Model.

SUMÁRIO

LISTA DE ILUSTRAÇÕES	9
LISTA DE TABELAS	11
LISTA DE ABREVIATURAS	12
LISTA DE SÍMBOLOS	13
1 INTRODUÇÃO	15
2 REVISÃO BIBLIOGRÁFICA	16
2.1 Redes Neurais	16
2.1.1 Função de Ativação	16
2.1.2 Treinamento da Rede	17
2.2 Redes Neurais Híbridas	18
2.3 Hidrodessulfurização	19
3 METODOLOGIA	21
3.1 Descrição do Processo	21
3.1.1 Fase Gasosa	22
3.1.2 Fase Líquida	22
3.1.3 Fase sólida	23
3.1.4 Cinética de HDS	23
3.1.5 Cinética de HDA	23
3.1.6 Cinética de HDN	23
3.1.7 Balanço de Energia	24
3.1.8 Modelagem da zona de <i>quench</i>	24
3.1.9 Estimação de parâmetros e simulação	24
3.2 Rede Neuronal Convencional	25
3.2.1 Keras	25
3.2.2 Desenvolvimento da Rede Neuronal	25
3.2.3 Dados de Planta	26
3.2.4 Treinamento	27
3.3 Rede Neuronal Híbrida	28
4 RESULTADOS	29
4.1 Ensaio do Horizonte Deslizante	29
4.2 Ensaio da Janela Móvel	32
4.3 Horizonte Deslizante e Janela Móvel - Remoção do Erro Fenomenológico	35

4.4	Ensaio da Janela Móvel - Rede Híbrida Alternativa	39
5	CONCLUSÃO	45
	BIBLIOGRAFIA	46
	APÊNDICE A TABELAS DE COEFICIENTE DE PEARSON	47

LISTA DE ILUSTRAÇÕES

1	Exemplo de uma rede neuronal.	17
2	Exemplo de reatores HDT comuns.	19
3	Diagrama de como o Reator HDT é conectado.	21
4	Comparação das concentrações normalizadas da saída da planta "S_out" e o modelo "S_out_modelo".	25
5	Resultado da média do primeiro ensaio de horizonte deslizante.	30
6	Resultado do desvio padrão do primeiro ensaio de horizonte deslizante.	30
7	Erro do modelo fenomenológico para os dados testados.	32
8	Exemplo do Ensaio da Janela Móvel.	33
9	Resultados do primeiro ensaio de janela móvel de 5% com avanço incremental.	33
10	Resultados do primeiro ensaio de janela móvel de 10% com avanço incremental.	34
11	Resultados do primeiro ensaio de janela móvel de 15% com avanço incremental.	34
12	Erro do modelo fenomenológico para a seção inicial dos dados.	35
13	Média absoluta do ensaio do horizonte deslizante da seção inicial dos dados.	35
14	Média absoluta do ensaio do horizonte deslizante da seção inicial dos dados.	36
15	Média absoluta do ensaio de janela móvel de 5%.	36
16	Média absoluta do ensaio de janela móvel de 10%.	37
17	Média absoluta do ensaio de janela móvel de 15%.	37
18	Desvio padrão do ensaio de janela móvel de 5%.	37
19	Desvio padrão do ensaio de janela móvel de 10%.	38
20	Desvio padrão do ensaio de janela móvel de 15%.	38
21	Resultados da média absoluta do horizonte deslizante para o novo modelo híbrido.	39
22	Resultados do desvio padrão do horizonte deslizante.	40
23	Resultados do ensaio de janela móvel de 5% para o modelo híbrido alternativo.	40
24	Resultados do ensaio de janela móvel de 10% para o modelo híbrido alternativo.	40
25	Resultados do ensaio de janela móvel de 15% para o modelo híbrido alternativo.	41
26	Desvio Padrão do ensaio de janela móvel de 5% para o novo modelo híbrido.	41

27	Desvio Padrão do ensaio de janela móvel de 10% para o novo modelo híbrido.	41
28	Desvio Padrão do ensaio de janela móvel de 15% para o novo modelo híbrido.	42
29	Predição do modelo híbrido e fenomenológico frente aos dados da planta.	42
30	Gráfico de dispersão dos modelos híbrido e fenomenológico.	44

LISTA DE TABELAS

1	Funções de Ativação e suas funções matemáticas.	17
2	Tabela com os erros dos dados de validação do ensaio.	31
3	Coefficientes de Pearson para as diferentes posições da janela e os diferentes modelos no ensaio de 15%.	43
4	Coefficientes de Pearson para as diferentes posições da janela e os diferentes modelos no ensaio de 5%.	48
5	Coefficientes de Pearson para as diferentes posições da janela e os diferentes modelos no ensaio de 10%.	49
6	Coefficientes de Pearson para as diferentes posições do horizonte deslizante.	50

LISTA DE ABREVIATURAS

<i>tanh</i>	Tangente Hiperbólica
ReLU	Unidade de Retificação Linear
HDT	Hidrotratamento
HDS	Hidrodessulfurização
HDN	Hidrodesnitrificação
HDA	Hidrodesaromatização

LISTA DE SÍMBOLOS

G	Função de ativação
x_i	Entradas de um neurônio
y_j	Saída de um neurônio
w_{ij}	Pesos de um neurônio
b_{ij}	Viés de um neurônio
X_i	Vetor de dados de um neurônio
X^{Nor}	Vetor de dados normalizado
$Y(X)$	Saída global do modelo híbrido
$F(X)$	Saída do modelo fenomenológico
$E(X)$	Saída da rede neuronal do modelo híbrido
u_G	Velocidade superficial da fase gasosa
T	Temperatura no leito catalítico,
R	Constante universal dos gases
p_i^G	pressões parciais
$k_i^L a_L$	coeficiente efetivo de transferência de massa gás-líquido
H_i	Constante de Henry
z	Comprimento do leito catalítico
u_L	Velocidade da fase líquida
C_i^L	Concentrações molar dos componentes na superfície do catalisador
C_i^S	Concentrações molar dos componentes na fase líquida
$k_i^S a_S$	Coeficiente efetivo de transferência de massa líquido-sólido
η	Fator de eficiência do catalisador
ρ_B	Densidade de bulk do catalisados no leito
ν	Coeficiente estequiométrico
K_{H_2S}	Constante de equilíbrio da cinética do tipo Langmuir-Hinshelwood
k_{HDS}	Constante de cinética de HDS

k_{HDAf}	Constante de cinética de reação HDA direta
k_{HDAr}	Constante de cinética de reação HDA indireta
$k_{HDN_{Nb}}$	Constante cinética da reação de nitrogênio básico
k_{HDN}	Constante cinética da reação de nitrogênio não básico
ρ_G	Massa específica da fase gasosa
ρ_L	Massa específica da fase líquida
c_p^L	Capacidade calorífica do diesel
ΔH_j	Entalpia molar
q	Vazão mássica de fluido de <i>quench</i>
l_s	Vazão mássica da fase líquida de entrada do <i>quench</i>
g_s	Vazão mássica da fase gasosa de entrada do <i>quench</i>
l_e	Vazão mássica da fase líquida de saída do <i>quench</i>
g_e	Vazão mássica da fase gasosa de saída do <i>quench</i>
ρ	Fração molar
T_s	Temperatura da mistura na entrada do <i>quench</i>
T_e	Temperatura da mistura na saída do <i>quench</i>
T_q	Temperatura da corrente de <i>quench</i>

1 INTRODUÇÃO

Reatores de hidrotratamento (HDT) possuem um papel fundamental no processo de fabricação do diesel (AQUINO GONÇALVES, 2017a). Estes são responsáveis por remover impurezas como o enxofre, metais, nitrogênio, entre outras, necessárias para garantir a qualidade do produto final. Além disto, para se obter baixos teores de enxofre, é necessária uma alta eficiência por parte dos reatores de hidrotratamento. Por estas razões, torna-se importante encontrar um modelo capaz de corretamente prever o comportamento do processo, visando o controle de otimização.

Além disso, nos últimos 30 anos, tem se intensificado o controle da emissão de poluentes no Brasil e no mundo (AQUINO GONÇALVES, 2017a). Por exemplo, a resolução ANP 42 de 2009 determinou a transição do diesel S1800 para o diesel S500¹, que se tornou o padrão nacional em 2013, e o S50 para o S10, cuja transição deve ser feita até 2020. A tendência indica que a concentração de enxofre no diesel deve diminuir com o passar do tempo.

Para gerar o modelo de reatores de hidrotratamento, desenvolve-se neste trabalho um modelo híbrido, onde se junta um modelo fenomenológico desenvolvido por Gonçalves (2017) com um modelo de redes neuronais. Isso é feito com o intuito de juntar os benefícios de ambas as abordagens e mitigar suas desvantagens.

O uso de modelos híbridos oferece diversas vantagens frente ao uso de redes comuns. Modelos convencionais funcionam como "caixa preta", não sendo possível obter um conhecimento preciso das relações e mecanismos internos, perdendo o significado físico dessas operações. A rede híbrida apresentada mantém o significado através do modelo fenomenológico, modelando como caixa preta apenas parte do sistema. Além disso, alguns tipos de redes híbridas são capazes de treinar de forma mais eficiente do que redes comuns, necessitando de menos dados, o que as tornam extremamente interessantes para dar início a um projeto, ou quando não há muitos dados disponíveis.

Para desenvolver o modelo, utilizam-se dados reais medidos em um reator industrial. Estes dados são empregados tanto para comparar os resultados do modelo obtido quanto para o treinamento da rede neuronal.

¹indica a concentração de enxofre em *ppm*, sendo S1800 equivalente a 1800 *ppm*, por exemplo.

2 REVISÃO BIBLIOGRÁFICA

Originalmente feitas para simular as sinapses de um cérebro humano, como descrito por Haykin (2009), redes neuronais artificiais tem provado serem extremamente úteis para apreender e representar modelos. Nas próximas seções, serão revisados alguns conceitos importantes referentes a utilização de redes neuronais.

2.1 Redes Neuronais

Redes neuronais são compostas por um elemento básico chamado neurônio, estes organizados em camadas. Em sua forma mais básica, um neurônio é apenas uma função matemática, que utiliza diferentes entradas multiplicadas por pesos para gerar uma saída. Normalmente essa saída é também aplicada sobre uma função de ativação, que é apenas uma função matemática aplicada sobre a saída do neurônio. As características das funções de ativação serão explicadas mais adiante.

A Equação (1) apresenta a forma como cada neurônio se comporta, o neurônio j recebe a entrada i da rede ou a saída de outros neurônios x_{ij} , multiplica por um peso w_{ij} , soma com o viés b_{ij} e então soma todos os resultados. Estes são então aplicados a função de ativação G que por sua vez gera a saída y_j .

$$y_j = G\left(\sum_{i=0}^m (w_{ij}x_{ij} + b_{ij})\right) \quad (1)$$

Um conjunto de neurônios artificiais é chamado de camada. Esse conjunto compartilha uma mesma função de ativação. Várias camadas são unidas de forma a gerar uma rede. Uma rede usual é composta de três camadas, chamadas de camada de entrada, camada oculta e camada de saída. Cybenkot (1989) provou que essa rede seria matematicamente capaz de representar qualquer função matemática utilizando a camada oculta com função de ativação sigmoide. Na Figura 1, observa-se a forma como os neurônios e as camadas são estruturados dentro de uma rede.

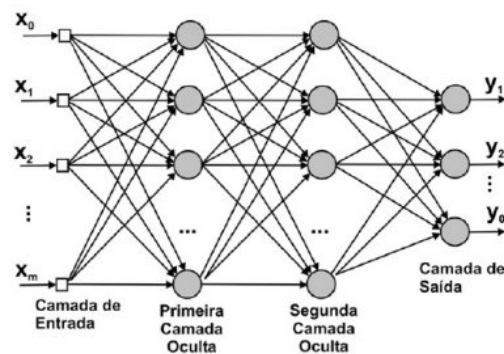
Na imagem, observar-se a forma como as conexões são realizadas. Cada saída de um neurônio transforma-se na entrada de cada neurônio de uma camada seguinte.

2.1.1 Função de Ativação

A função de ativação, representada pela função G na Equação 1, é uma função aplicada à saída de um neurônio em uma camada. Essa função é compartilhada pela camada inteira.

Essas funções possuem diversas utilidades. Primeiramente, elas adicionam não linearidades na rede. Isso possibilita que a rede aprenda uma maior gama de modelos do que

Figura 1: Exemplo de uma rede neuronal.



Fonte: Retirado de Fature (2017).

uma rede puramente linear.

Além disso, é importante ressaltar que as funções de ativação têm um efeito importante no treinamento. Algumas funções serão mais eficientes do que outras, sendo capazes de serem treinadas mais rapidamente. Um exemplo deste fenômeno é a função retificadora (ReLU), que vem apresentando resultados de treinamento melhores em relação a outras em diversas situações, e por isso vem ganhando popularidade (GLOROT; BORDES; BENGIO, 2011).

3 funções de ativação diferentes se destacam, a tangente hiperbólica, a sigmoide e a ReLU. A função sigmoide devido a sua propriedade de ser capaz de aproximar qualquer função. A função tangente hiperbólica (*tanh*) pela mesma razão que a sigmoide, ambas possuem o mesmo perfil, mas enquanto a saída da função sigmoide é limitada ao intervalo $[0,1]$, a da *tanh* é $[-1,1]$. Já a função ReLU é importante devido sua popularidade e uso em projetos de redes neurais, sendo atualmente a função de ativação mais popular em projetos de redes neurais (RAMACHANDRAN; ZOPH; LE, 2017). Cada uma das funções utilizadas e suas expressões matemáticas se encontram na Tabela 1:

Tabela 1: Funções de Ativação e suas funções matemáticas.

Função de Ativação	Equação	Intervalo de Saída
ReLU	$G(x) = \begin{cases} 0 & \text{se } x < 0 \\ x & \text{se } x \geq 0 \end{cases}$	$[0, x^+]$
<i>Tanh</i>	$G(x) = \frac{1-e^{-2x}}{1+e^{-2x}}$	$[-1, 1]$
Sigmoide	$G(x) = \frac{1}{1+e^{-x}}$	$[0, 1]$

Fonte: Retirado de Glorot, Bordes e Bengio (2011)

2.1.2 Treinamento da Rede

Um processo extremamente importante para um rede neuronal é o seu treinamento. É durante esse processo que ela se transforma de uma rede genérica para se tornar o modelo de algo específico. Durante o treinamento supervisionado (em que a saída da rede é diretamente comparada com dados reais), determinam-se os pesos w_{ij} de cada neurônio que otimizam a saída do modelo em relação aos dados reais. Essa otimização é definida pelo parâmetro de custo (também chamado de perda) que é definido durante a criação da rede.

Um das formas mais comuns da função de custo é o erro quadrático médio entre a saída da rede e o valor esperado, de maneira que o algoritmo procurará minimizar esse valor através de um otimizador não linear. Existem diversos algoritmos de otimização de redes. De maneira geral, a grande maioria utiliza o princípio de *backpropagation*, onde se analisa o custo, e este valor é "propagado" pela rede através do algoritmo de forma a reduzi-lo.

O algoritmo de otimização utilizado é um aspecto importante do treinamento. Recentemente, o algoritmo *Adam* tem recebido bastante destaque em pesquisas devido a sua precisão e velocidade de treinamento (KINGMA; BA, 2015). O método funciona calculando o gradiente do erro, dessa maneira encontrando-se a direção de otimização e variando os pesos dos neurônios. A vantagem do método de *Adam* frente a outros algoritmos é que ele possui taxas de aprendizado variáveis, de maneira que ele consegue ser mais rápido que métodos equivalentes quando longe e do ponto ótimo e mais preciso quando próximo.

O processo de aplicar o algoritmo de otimização sobre a rede é repetido diversas vezes. Em cada repetição, todos os dados de treino são inseridos e os pesos variados, no fim dos dados, o modelo é validado. Esse processo é executado até o modelo convergir ao seu ponto de ótimo. O nome dado a cada uma dessas iterações é *epoch*.

2.2 Redes Neurais Híbridas

Redes híbridas são construídas combinando o modelo fenomenológico de um sistema com as redes neuronais tradicionais. Elas apresentam diversas vantagens em relação redes comuns, a presença da parte analítica implicam que elas não são simplesmente uma caixa preta, e carregam informações sobre características físicas do modelo.

Na literatura encontram-se três principais formas de se unir os dois modelos. Na primeira forma, descrita por Chaichana et al. (2012), utilizam-se as redes neuronais para modelar algum aspecto do modelo analítico. Normalmente a parte modelada pelas redes neuronais será a mais não-linear.

A segunda forma, desenvolvida por Qi et al. (1999), utiliza redes neuronais para modelar parâmetros internos do modelo analítico. Dessa maneira, os valores que normalmente seriam modelados como constantes em suas equações variam de acordo com a entrada, gerando um resposta mais precisa para o modelo.

A terceira forma, estruturada por Psychogios e Ungar (1992), baseia-se na ideia de utilizar as redes neurais para corrigir o modelo analítico. Nesta implementação, calcula-se o erro entre o modelo analítico e os valores reais da planta. Então utiliza-se o erro como sendo o parâmetro de saída das redes neurais, sendo a mesma treinada para aprender tais padrões. Dessa maneira, ela torna-se uma fator de correção do modelo analítico. Essa forma de modelagem é a forma utilizada nesse trabalho.

Essa modelagem apresenta diversas vantagens frente a uma modelagem com apenas uma rede neuronal comum. Como foi dito anteriormente, ela permite que o modelo ainda tenha significado físico. Psychogios e Ungar (1992) mostram que o resultado do modelo fenomenológico é significativamente melhorado pela inserção da rede neuronal, mesmo analisando comparando modelos que utilizam Filtro de Kalman. Além disso, ele demonstra que a rede neuronal híbrida é mais rápida para treinar e necessita de menos informação que um modelo construído apenas com uma rede neuronal tradicional.

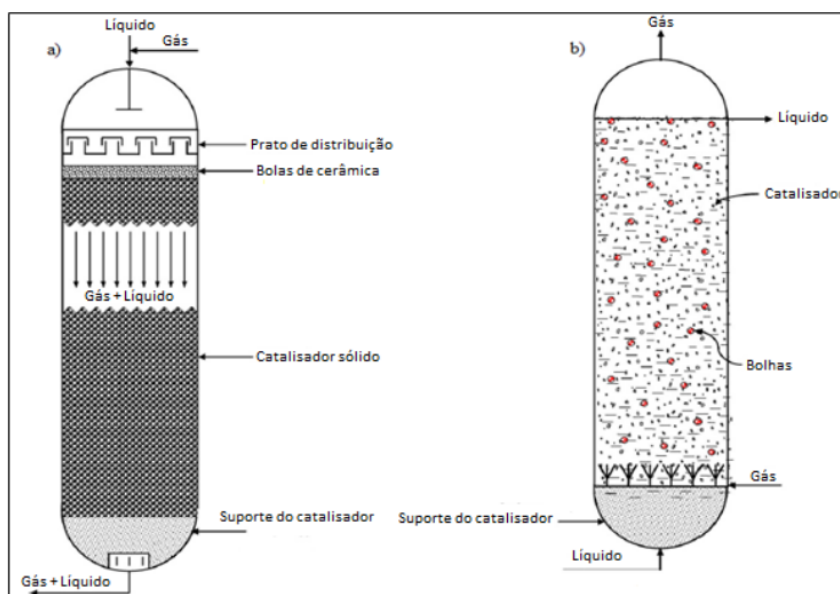
2.3 Hidrodessulfurização

O processo de Hidrodessulfurização (HDS) é um processo de hidrotratamento (HDT) catalítico onde se insere hidrogênio com o intuito de remover enxofre dos subprodutos da produção de petróleo, transformando os compostos organossulfurados em sulfeto de hidrogênio (AQUINO GONÇALVES, 2017a). Dentro desse trabalho foi analisada a hidrodessulfurização do diesel.

Diversos estudos vem sendo realizados abordando este tópico com o intuito de produzir o melhor modelo e simula-lo. Dietz (2014) desenvolveu um modelo a partir da cinética das reações de hidrotratamento. Gonçalves (2017) expandiu o modelo, desenvolveu um método de simulação dos modelos e estimação de parâmetros, gerando assim um modelo fenomenológico que representa o processo. Ainda assim, o erro do modelo fenomenológico continuou relativamente alto, de maneira que encontrar novas abordagens de modelar esse processo continua sendo importante.

A planta testada nesse trabalho é do tipo *trickle-bed*, onde a fase líquida escorre pelo leito catalítico (fixo ao reator) e escoo pela base. Pode-se observar na Figura 2 uma representação de dois reatores HDT. Na Figura 2a é apresentado um reator *trickle-bed*, enquanto na Figura 2b um do tipo *slurry*.

Figura 2: Exemplo de reatores HDT comuns.



Fonte: Retirado de Gonçalves (2017).

Segundo Gonçalves (2017), as ligas metálicas mais utilizadas para a criação de catalisadores para reações de HDS são de de cobalto-molibdênio (CoMo) ou níquel-molibdênio (NiMo), devido ao seu baixo custo, seletividade e resistência ao envenenamento.

No processo de HDT comumente se utiliza duas camadas de reatores HDT em série, entre cada um dos reatores, há uma etapa chamada *quench*. Nesta etapa, resfria-se a saída do reator, de forma a diminuir a adsorção do oposito, aumentando a remoção de enxofre, em especial para os compostos mais pesados. Além de controlar a temperatura, também se insere hidrogênio, de forma a aumentar a razão hidrogênio/sulfeto.

É interessante notar que, conforme descrito por Aquino Gonçalves (2017), o reator HDT do tipo *trickle-bed* possui diversas vantagens, entre elas:

- fluxo de líquido próximo ao empistonado, maior taxa de conversão;
- baixa perda do catalisador;
- não possui partes móveis;
- boa faixa de operação de temperatura e pressão;
- baixo custo operacional;
- velocidade do líquido é variável dependendo de; parâmetros internos
- pode-se utilizar diferentes regimes de fase gasosa-líquida, o que torna a operação flexível;
- baixa queda de pressão;
- não há risco de inundação.

E entre suas desvantagens:

- baixa eficácia, quando comparado com outros modelos;
- há uma restrição no tamanho máximo da partícula;
- passível de problemas de má distribuição;
- limitações na operação com líquidos viscosos;
- limitado a reações relativamente rápidas;
- o processo para recuperar calor perdido é complexo.

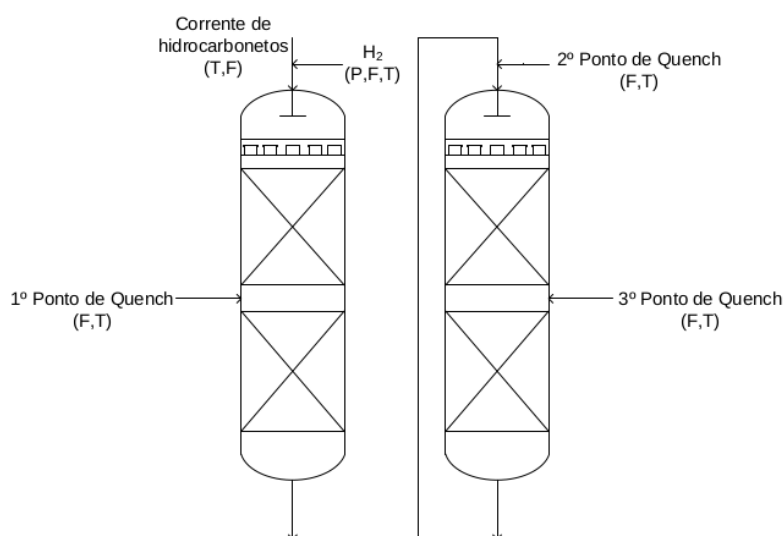
3 METODOLOGIA

O desenvolvimento desse trabalho se deu em 4 etapas: a configuração do modelo fenomenológico, que já havia sido implementado (AQUINO GONÇALVES, 2017a), a otimização de uma rede neuronal convencional para modelar o sistema, o desenvolvimento da rede híbrida, unindo os dois modelos e a comparação destes modelos com os dados reais.

3.1 Descrição do Processo

A planta possui duas zonas distintas, sendo estes o próprio reator de hidrotreatamento (HDT) e a zona de *quench*, que pode estar presente dentro do reator ou entre reatores. No modelo desenvolvido, estes componentes são interligados em série, de forma que a saída de uma seção é a entrada da próxima, e assim sucessivamente. Na planta, há um conjunto de 2 reatores HDT (cada um com 2 leitos catalíticos e uma zona de *quench* interno) interligados por outra zona de *quench*, formando o sistema completo, conforme a Figura 3.

Figura 3: Diagrama de como o Reator HDT é conectado.



Fonte: Retirado de Golçalves (2017).

É importante ressaltar que além do processo de hidrodessulfurização (HDS), cujo o resultado é o foco da análise, o modelo também considera simultaneamente o processo

de hidrodesnitrificação (HDN) e hidrodesaromatização (HDA), devido a competição que estes processos tem no processo de HDS. É assumido no modelo que a cinética dos compostos pode ser representado de maneira agregada.

O modelo é construído com base no balanço de massa e energia entre as fases gasosa, líquida e sólida. O sistema é modelado ao longo do comprimento do reator, de maneira que o resultado da concentração de saída de organossulfurados é o seu valor ao final do comprimento. Várias simplificações e considerações foram feitas para criar o modelo, estas são:

- a cinética para conversão de misturas de organossulfurados é formulada em função de compostos de referência;
- não existem variações radiais;
- sistema é composto por três fases;
- a velocidade de gás e líquido é constante ao longo do leito;
- o óleo não vaporiza;
- o reator esta em estado estacionário;
- as reações ocorrem exclusivamente na superfície do catalisador ;
- não há transferência de massa entre as partículas;
- gases ideais;
- as características variantes no tempo do modelo foram desprezadas.

3.1.1 Fase Gasosa

O balanço de massa da fase gasosa é dado por:

$$\frac{u_G}{RT} \frac{dp_i^G}{dz} + k_i^L a_L \left(\frac{p_i^G}{H_i} - C_i^L \right) = 0 \quad (2)$$

sendo u_G a velocidade da fase gasosa, $i = H_2, H_2S, NH_3$. T é a temperatura no leito catalítico, z o comprimento do leito catalítico, R é a constante universal dos gases, p_i^G são as pressões parciais, $k_i^L a_L$ é o coeficiente efetivo de transferência de massa gás-líquido. H_i é a constante de Henry no componente i , utilizada para aproximar as concentrações de equilíbrio na fase líquida pela lei de Henry.

3.1.2 Fase Líquida

O balanço de massa da fase líquida, para os mesmos componentes H_2, H_2S, NH_3 , é dado por:

$$u_L \frac{dC_i^L}{dz} - k_i^L a_L \left(\frac{p_i^G}{H_i} - C_i^L \right) + k_i^S a_S (C_i^L - C_i^S) = 0 \quad (3)$$

sendo u_L a velocidade da fase líquida, C_i^L e C_i^S são concentrações molares dos componentes i na fase líquida e na superfície do catalisador e $k_i^S a_S$ é o coeficiente efetivo de transferência de massa líquido-sólido.

É assumido que os compostos aromáticos, organossulfurados e organonitrogenados não são voláteis, então o balanço de massa para estes é:

$$u_L \frac{dC_i^L}{dz} + k_i^S a_S (C_i^L - C_i^S) = 0 \quad (4)$$

3.1.3 Fase sólida

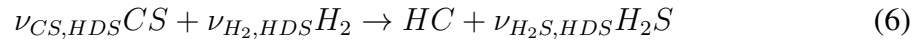
As reações químicas consomem ou produzem os componentes transferidos entre a fase líquida e o catalisador seguindo o seguinte balanço:

$$k_i^S a_S (C_i^L - C_i^S) = \sum_j -\nu_{ij} \rho_B \eta r_{j,r} \quad (5)$$

sendo i novamente igual à H_2, H_2S, NH_3 , porém também organossulfurados, organonitrogenados, aromáticos; j indica o tipo de reação química HDS, HDN e HDA. r_i é a taxa intrínseca de reação j , η é o fator de eficácia do catalisador, ρ_B é a densidade do catalisador no leito e ν_{ij} são os coeficientes estequiométricos.

3.1.4 Cinética de HDS

A partir da utilização de reações de forma agregada, pode-se definir a cinética de HDS - além das subsequentes cinéticas - como possuindo a seguinte equação:



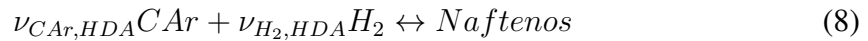
sendo que $\nu_{i,j}$ é o coeficiente estequiométrico. A taxa de reação é dado por:

$$r_{HDS} = k_{HDS} \frac{(C_{C-S}^S)(C_{H_2}^S)^{0,45}}{(1 + K_{H_2S} C_{H_2S}^S)^2} \quad (7)$$

A cinética representada é do tipo Langmuir-Hinshelwood com adsorção de sulfeto de hidrogênio. K_{H_2S} é a constante de equilíbrio e k_{HDS} é a constante cinética de HDS.

3.1.5 Cinética de HDA

A cinética de HDA considera a seguinte reação reversível:



e ocorre com a seguinte taxa de reação:

$$r_{HDA} = k_{HDAf} p_{H_2}^G C_{Ar}^S - k_{HDAr} (1 - C_{Ar}^S) \quad (9)$$

onde k_{HDAf} é a constante cinética de reação direta e k_{HDAr} a constante cinética de reação indireta.

3.1.6 Cinética de HDN

Há uma divisão entre compostos básicos e não básicos no equacionamento dos compostos HDN:



suas taxas de reação são:

$$r_{HDN_{Nb}} = k_{HDN_{Nb}} (C_{N_{Nb}}^S)^{1,5} \quad (11)$$

$$r_{HDN} = k_{HDN_{Nb}} (C_{N_{Nb}}^S)^{1,5} - k_{HDN} (C_N^S)^{1,5} \quad (12)$$

onde $k_{HDN_{Nb}}$ é a constante cinética da reação de nitrogênio básico e k_{HDN} a constante para o caso não básico.

3.1.7 Balanço de Energia

O balanço de energia é dado pela seguinte equação:

$$\frac{\partial T}{\partial z} = \frac{1}{u_G \rho_G c_p^G + u_L \rho_L c_p^L} \sum_j -\rho_B \eta r_{j,r} (-\Delta H_j) \quad (13)$$

onde ρ_G e ρ_L são as massas específicas para a fase gasosa e líquida respectivamente, c_p^L é a capacidade calorífica do diesel e ΔH_j é a entalpia molar da reação j .

3.1.8 Modelagem da zona de *quench*

A zona de *quench* é modelada como um misturador entre a corrente de *quench* e a saída do leito. As seguintes equações de balanço representam o misturador:

$$q + l_s + g_s = l_e + g_e \quad (14)$$

q é a vazão mássica do fluido de *quench*, l_s é a vazão mássica da fase líquida e g_s é a vazão mássica da fase gasosa, ambos da seção de reator prévia a zona de *quench*. l_e e g_e representam as mesmas variáveis, mas de entrada do leito catalítico após o *quench*.

Há também o balanço de massa por fase:

$$q\nu + g_s = g_e \quad (15)$$

$$q(1 - \nu) + l_s = l_e \quad (16)$$

nesse balanço, ρ representa a fração molar do fluido de *quench*. O balanço de energia da zona de mistura é:

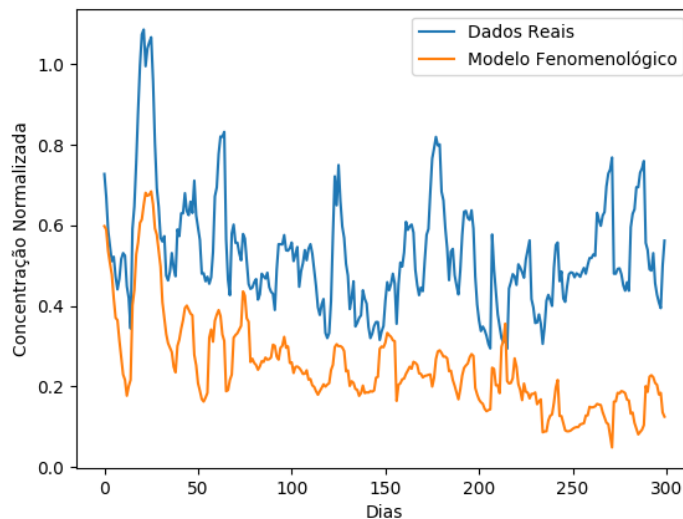
$$\int_{T_s}^{T_e} l_s c_p^L dT + \int_{T_s}^{T_e} g_s c_p^G dT + \int_{T_s}^{T_e} q c_p^q dT = 0 \quad (17)$$

onde T_e é a temperatura da mistura na saída do *quench*, T_s na entrada e T_q é a da corrente de *quench*.

3.1.9 Estimação de parâmetros e simulação

Utilizando os dados reais, 8 parâmetros ($k_{HDS}, k_{H_2S}, k_{HDAf}, k_{HDAr}, k_{HDN_{Nb}}, k_{HDN}, \rho_{ij}, \eta$) foram estimados por Gonçalves (2017). Também foram feitas simulações utilizando as ODEs e equações descritas utilizando a plataforma *JModelica*, que se comunica com *Python* (AQUINO GONÇALVES, 2017b). Pode-se observar o resultado do modelo fenomenológico na Figura 4.

Figura 4: Comparação das concentrações normalizadas da saída da planta "S_out" e o modelo "S_out_modelo".



Fonte: Autor.

Infelizmente a planta possui não-linearidades e aspectos de difícil inclusão dentro do modelo fenomenológico. Observa-se na Figura 4 que o modelo chega a ser 30 vezes menor que os dados de planta em alguns momentos. Mais detalhes sobre a modelagem, equacionamento e simulação desse processo podem ser encontrados em Gonçalves (2017) e Dietz (2014).

3.2 Rede Neuronal Convencional

Antes do desenvolvimento da rede neuronal híbrida, foi desenvolvida uma rede neuronal comum (sem a utilização do modelo fenomenológico) para se ter uma "rede controle", com estrutura similar a rede híbrida, para servir de comparação de resultados.

3.2.1 Keras

Foi escolhido para fazer o desenvolvimento a plataforma *Keras* (CHOLLET et al., 2015). Esse é um *framework* para a linguagem de programação *Python* e automatiza diversos aspectos do desenvolvimento de redes neurais.

Essa plataforma foi escolhida devido a fácil integração com o modelo fenomenológico, que havia sido desenvolvido em *Python*. Além disso, ela apresenta um *framework* simples, porém poderoso, que facilita a experimentação e a variação de modelos.

3.2.2 Desenvolvimento da Rede Neuronal

O modelo inicial testado foi o modelo proposto por Cybenkot (1989). Este possui uma única camada oculta e utiliza em suas camadas a função de ativação sigmoide. Devido à similaridade da sigmoide com a função tangente hiperbólica, esta também foi testada. Outra função de ativação testada foi a função de retificação linear (ReLU) (GLOROT; BORDES; BENGIO, 2011). A partir disso, testou-se a rede com diversas camadas diferentes, além de número de neurônios e estas três funções de ativação, de maneira a se convergir a um melhor resultado de validação. Esse resultado é estimado à partir do erro

médio quadrático entre os dados reais e os de validação, que representam 20% dos dados de treinamento selecionados aleatoriamente.

A rede selecionada foi de 4 camadas, todas com 8 neurônios (exceto a camada de saída, com um neurônio), sendo que a camada visível e a segunda camada oculta têm função de ativação ReLU, a primeira camada oculta tem função tangente hiperbólica e a camada de saída possui função de ativação sigmoide. Foi selecionado a função sigmoide para a última camada devido ao fato de ela ser limitada no intervalo $[0, 1]$. Como a saída do modelo tem natureza de concentração, e portanto, um valor positivo, decidiu-se usar uma função que limita sua saída nesse intervalo. Observa-se que a saída do sistema é normalizada, então os valores de concentração não serão limitados ao intervalo da função de ativação. Isso será abordado em maiores detalhes na Seção 3.2.4.

Além dessa estrutura, cada neurônio possui internamente um valor de viés. Esse valor é somado a multiplicação da entrada da rede com seu peso, e é variado juntamente com os pesos durante a etapa de treinamento.

Um problema comum de ocorrer com redes neuronais é chamado de *overfitting*. O *overfitting* ocorre quando uma rede chega ao seu limite de otimização, porém continua-se tentando treina-la. Nesse momento, ao invés de melhorar a resposta da rede, acaba-se reduzindo sua precisão. Isso ocorre, pois durante o *overfitting*, a rede tenta aprender a posição exata dos dados, memorizando também qualquer variação aleatório presente neles, e inserindo ela em futuras predições. Dois métodos foram empregados para se minimizar o efeito do *overfitting*, o primeiro é chamado de *dropout* e o segundo se chama *checkpoint*, que será explicado na Seção 3.2.4.

Dropout é um método que limita a quantidade de neurônios da rede sendo treinados (SRIVASTAVA et al., 2015). Em cada *epoch* de treinamento, alguns neurônios são aleatoriamente selecionados para serem congelados, de forma que seu gradiente não são calculados e seu pesos se mantêm constantes. Isso tem a função de ajudar a evitar que alguns neurônios dominem o processo de predição, pois caso algum comece a se destacar, ele inevitavelmente será congelado em alguns *epochs* e isso permitirá que outros neurônios tenham seus pesos atualizados. Experimentalmente, foram realizados diversos ensaios com a rede e foi decidido que o melhor resultado é congelando um neurônio em cada camada por *epoch*.

3.2.3 Dados de Planta

Os dados utilizados para treino são provenientes de ensaios de um reator HDT industrial. Há um total de 1008 pontos, abrangendo 4 anos de operação. Cada ponto representa dados médios diários da planta. É importante ressaltar que apesar das medições representarem um dia do processo, elas não são feitas diariamente de maneira que a distribuição de dados relativo ao tempo não é uniforme. Para o treino de ambos os modelos de redes neurais, foram utilizadas 10 variáveis de entrada - as mesmas utilizadas para calcular o modelo fenomenológico - listadas abaixo com suas respectivas descrições:

- FoR1 - Vazão de diesel na entrada do reator $[cm^3 \cdot s^{-1}]$
- FqN - Vazão de hidrogênio na entrada do reator $[cm^3 \cdot s^{-1}]$
- FqN1 - Vazão de hidrogênio na primeira zona de *quench* do reator $[cm^3 \cdot s^{-1}]$
- FqN2 - Vazão de hidrogênio na segunda zona de *quench* do reator $[cm^3 \cdot s^{-1}]$
- FqN3 - Vazão de hidrogênio na terceira zona de *quench* do reator $[cm^3 \cdot s^{-1}]$

- PinR1 - Pressão na entrada do primeiro reator [kPa]
- PinR2 - Pressão na entrada do primeiro reator [kPa]
- T_H2 - Temperatura do Hidrogênio [K]
- TinR1 - Temperatura de entrada do reator 1 [K]
- S_in - Concentração de organossulfurados na entrada [ppm]

Como variável de saída, foi utilizado "S_out", sendo este a concentração de saída de organossulfurados em partes por milhão.

3.2.4 Treinamento

Uma etapa importante para o treinamento de redes neuronais é a de normalizar as variáveis de entrada e saída para o intervalo entre $[0, 1]$. Isto é feito por duas razões, a primeira é para que todas as variáveis tenham o mesmo peso no cálculo interno do gradiente durante o treinamento, independente de seu valor absoluto. A segunda razão é que deixar as variáveis normalizadas torna o processo de treinamento mais rápido (J. JIN; LI; L. JIN, 2014). As variáveis são normalizadas conforme a Equação 18, sendo que X_j é um vetor com as variáveis de entrada ou saída da rede neuronal, de índice j , X_j^{Nor} é o resultado normalizado, X_j^{max} é o máximo valor que essa variável possui entre os dados analisados e X_j^{min} é o mínimo valor.

$$X_j^{Nor} = \frac{X_j - X_j^{min}}{X_j^{max} - X_j^{min}} \quad (18)$$

Ao se utilizar os dados normalizados para treinar a rede, seu resultado será, naturalmente, normalizado também. Por essa razão, ao se realizar uma operação de predição pela rede neuronal, é necessário realizar a operação reversa da Equação 18 sobre o valor de saída da rede.

Um detalhe importante ao se treinar uma rede neuronal é evitar utilizar todos os dados para o treinamento, mantendo sempre um percentual dos dados para validação dos resultados. Dessa maneira, pode-se testar o modelo com dados não utilizados para treinamento e calibração, evitando assim inserir um viés sobre o teste de validação do modelo. Por esse motivo, se reservou 20% dos dados para testes e validação do resultado, sendo que estes dados são escolhidos aleatoriamente antes de se treinar o modelo. Para realizar o treinamento, utiliza-se o algoritmo de otimização *Adam*, devido a sua precisão e velocidade (KINGMA; BA, 2015). Além disso, inicializa-se os pesos e vieses dos diferentes neurônios com valores aleatórios no intervalo $[0,1]$.

O resultado do erro médio quadrático entre esses dados, chamados de dados de validação, e a saída do modelo é a forma utilizada para definir o quão bom o modelo é. Um modelo será melhor do que outro quando seu erro for menor.

Através do processo de validação do modelo durante o treinamento, pode-se realizar uma técnica de redução de *overfitting* chamada de *checkpoint*. Nessa técnica, no final de cada *epoch*, calcula-se o erro do modelo sendo treinado e compara-se com o menor erro já encontrado pelo modelo. Caso o erro desse novo modelo seja menor, salva-se esse modelo, sobreescrevendo o último modelo salvo. No fim do treinamento, o modelo salvo será o modelo com menor erro. Isso evita o *overfitting*, pois quando o modelo começa a realizar o *overfit*, seu erro com relação a dados não utilizados no treinamento aumentam,

e então ele deixa de ser salvo. Essa técnica é similar a de *early-stopping*, mas oferece a vantagem de não necessitar configurar os parâmetros da parada, apesar de ser mais lenta (PRECHELT, 2012). Na técnica de *early-stopping*, o treinamento é parado ao se encontrar o menor valor.

3.3 Rede Neuronal Híbrida

Decidiu-se implementar o modelo híbrido como sendo um fator de correção sobre o modelo fenomenológico. Esta é a forma mais simples de montar o modelo híbrido, o que a torna ideal para uma primeira implementação. Dessa maneira, tem-se que a saída do modelo híbrido é:

$$Y(X) = F(X) + E(X^{Nor}) \quad (19)$$

Sendo $Y(X)$ o resultado do modelo híbrido, $F(X)$ o resultado do modelo fenomenológico, $E(X)$ o resultado da rede neuronal modelada e X o vetor de entradas do sistema, sendo que para o modelo híbrido, ele é normalizado conforme descrito na Seção 3.2.4. É importante ressaltar que modelar o erro do modelo fenomenológico ao invés do sistema completo diminui a quantidade de informação que a rede necessita aprender, sendo esta os aspectos não modelados.

A primeira etapa para modelar o sistema é a criação da saída para a rede, ou seja, um vetor com todos os erros do modelo fenomenológico. Isso é feito utilizando o modelo para calcular o valor de concentração de saída em cada ponto em que se há dados reais. Após isso ser feito, se subtrai o valor real pelo estimado em cada um desses pontos, gerando assim o vetor de erros.

Com esse vetor, é possível realizar o treinamento, de maneira análoga ao que foi realizado na Seção 3.2. Em um primeiro momento, manteve-se a estrutura da rede neuronal comum, com o intuito de comparar apenas as metodologias durante os ensaios, evitando assim que diferentes redes possam ser um fator de divergência nos resultados. Por essa razão, a rede possui novamente 4 camadas, de 8 neurônios cada, exceto naturalmente a última, com função de ativação ReLU na primeira e terceira, tangente hiperbólica na segunda e sigmoide na última. Para treinamento, foi novamente usado o método de *Adam*.

Naturalmente, o resultado da previsão desta rede neuronal não é mais o resultado final. Para realizar uma previsão do modelo para um certo vetor de entrada X_e , calcula-se tanto o resultado do modelo fenomenológico $F(X)$ quanto o da rede para esse vetor. Então aplica-se o inverso da Equação 18 sobre a saída do modelo híbrido para se encontrar o valor de saída não normalizado. A partir disso, aplica-se a Equação 19 para se obter o resultado final da previsão.

4 RESULTADOS

Foram realizados ensaios para comparar os diferentes tipos de redes neuronais. Inicialmente, testes de horizonte deslizante foram realizados para simular a passagem de tempo na planta. Também foram realizados testes com janelas móveis, também para simular a passagem de tempo, mas representando melhor o processo de previsão. Os ensaios foram realizados para os três modelos.

4.1 Ensaio do Horizonte Deslizante

O primeiro ensaio realizado foi o do horizonte deslizante. Nesse ensaio, separa-se todos os dados da planta em duas seções, sendo a primeira utilizada para fins de treinamento e validação do modelo, e a segunda para fins de teste e verificação.

Essas seções são variadas com passos definidos de 5 % do tamanho dos dados (o que representa um total de 50 dias de medições reais), aumentando a janela de treinamento e diminuindo a de testes. É importante notar que, no ensaio, a janela de treinamento começa com 15 % dos dados e termina com 85 %. Segue um exemplo da execução desse ensaio:

- treina-se a rede para os 15% dos dados da seção inicial de treino;
- testa-se a rede frente os 85% dos dados da seção inicial de teste;
- avança-se o horizonte em 5%, o processo de treino e teste é repetido, neste caso, com 20% e 80% dos dados, respectivamente;
- repete-se esse processo até o horizonte chegar aos 85%.

Para cada teste de horizonte realizado, foram analisados 4 fatores. A média absoluta e o desvio padrão do erro entre o modelo e os dados reservados para testes e média absoluta e o desvio padrão do erro entre o modelo e os dados de validação do treinamento, que representam 15 % dos dados de treinamento. Dessa maneira, pode-se analisar a magnitude e a distribuição do erro em um ensaio que simula previsão, além de validar o modelo para os dados utilizados.

Esse ensaio foi feito para ambos os modelos neuronais, assim como o modelo fenomenológico puro. Os gráficos para o erro médio absoluto e o desvio padrão sobre o erro de previsão do modelo dessa análise se encontram na Figura 5 e 6, juntamente com a Tabela 2, dos resultados de validação do ensaio:

Figura 5: Resultado da média do primeiro ensaio de horizonte deslizante.

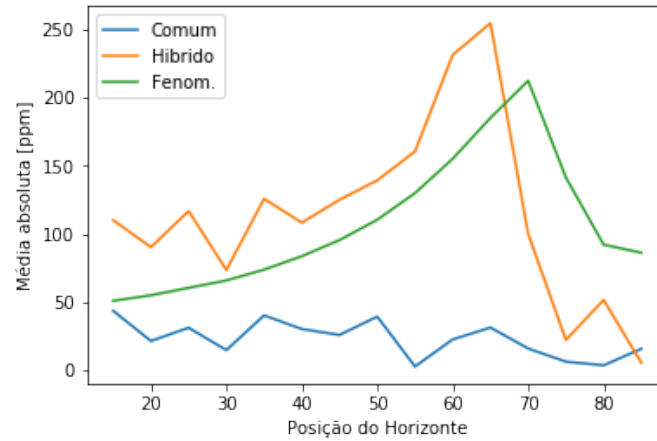
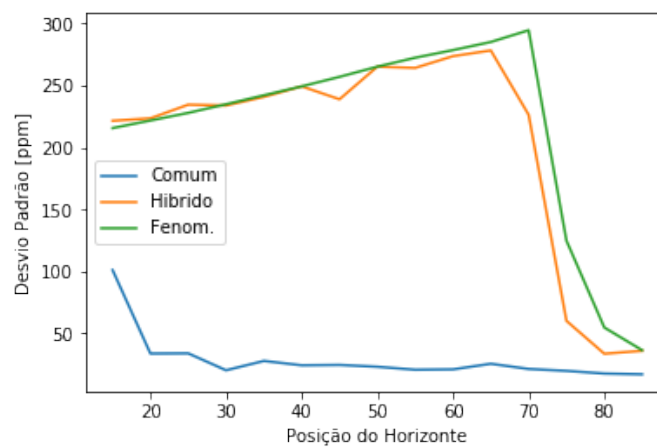


Figura 6: Resultado do desvio padrão do primeiro ensaio de horizonte deslizante.



Fonte: Autor.

Tabela 2: Tabela com os erros dos dados de validação do ensaio.

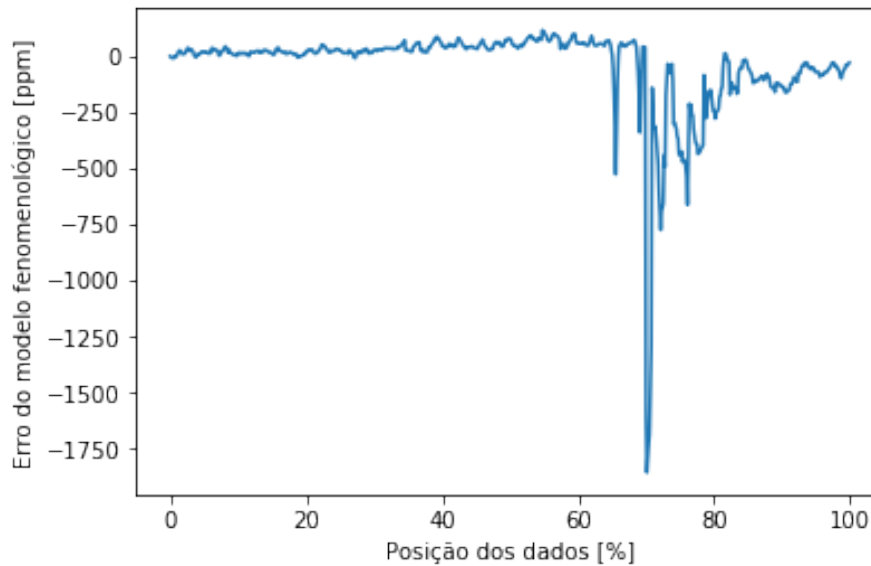
Posição do Horizonte	Erro Val. - Comum [ppm]	Erro Val. - Híbrido [ppm]
15%	28.97	5.31
20%	35.59	8.66
25%	21.84	5.68
30%	24.12	6.61
35%	20.14	6.17
40%	20.87	12.12
45%	18.36	7.92
50%	16.66	15.45
55%	14.86	6.15
60%	16.49	6.73
65%	13.88	7.64
70%	17.09	7.85
75%	13.87	13.40
80%	15.19	15.57
85%	15.05	20.44

Fonte: Autor

Algumas constatações sobre esse resultado:

- O modelo híbrido tem uma média pior que o modelo fenomenológico puro (no caso do horizonte) até os 65% de avanço.
- O desvio padrão do modelo híbrido e do fenomenológico é quase idêntico até a marca de 65 %, quando o híbrido decresce, juntamente com o fenomenológico.
- O modelo comum teve um erro menor que o híbrido e o fenomenológico em todas as médias do horizonte, mas perdeu para o híbrido em todas as comparações dos erros sobre os dados de validação, exceto nos dois horizontes finais.
- O erro entre o modelo fenomenológico e o processo tem uma tendência de crescimento constante até o valor de 60% dos dados. Após esse valor, ele muda drasticamente de comportamento, variando em módulo e em intensidade, conforme mostrado na Figura 7.

Figura 7: Erro do modelo fenomenológico para os dados testados.



Fonte: Autor.

Como foi constatado, observa-se nos resultados que o modelo híbrido possui um erro maior que o modelo fenomenológico até o horizonte de 65 %. Como o modelo híbrido também incorpora o modelo fenomenológico, isso significa que ele está prejudicando a predição. Apesar disso, o modelo comum apresentou um erro menor que o fenomenológico e o híbrido em todos os horizontes, demonstrando que ambas as redes são capazes de aprender nas condições estabelecidas (a rede híbrida deve aprender, inclusive, com mais facilidade que a rede comum). Isso é corroborado pelo fato de que a rede híbrida apresentou um erro menos do que a rede comum em quase todos os dados de validação, com exceção dos últimos dois horizontes.

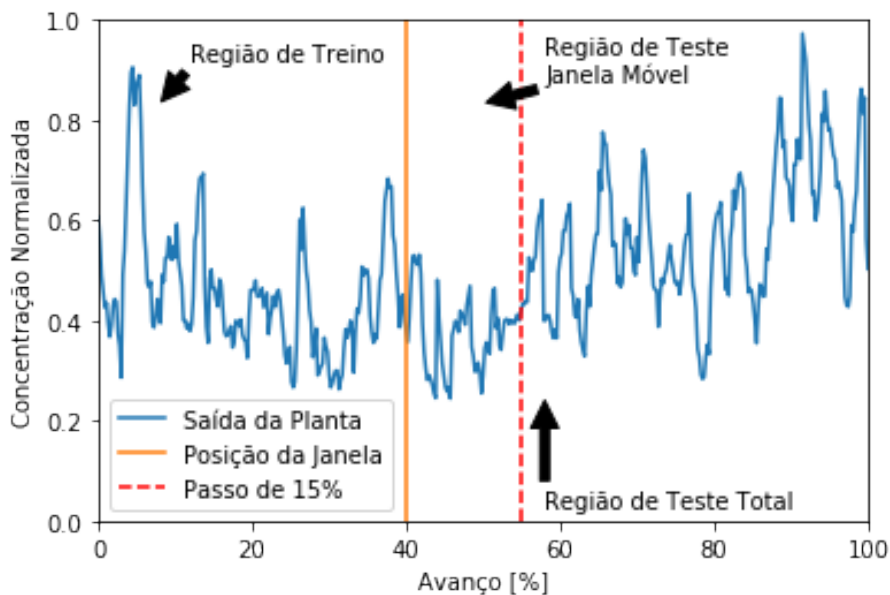
Tudo isso indica que o problema foi de fato a necessidade da rede híbrida de extrapolar o modelo treinado para o caso da variação grande do erro, pois a rede híbrida ganhou da comum nos dados de validação, situação onde ela não precisaria extrapolar os dados, em quase todos os casos. Já a rede comum não precisa realizar essa extrapolação, devido ao fato de que esse erro é introduzido pelo modelo fenomenológico, o que explica o seu desempenho relativo a rede híbrida na janela de teste.

4.2 Ensaio da Janela Móvel

Para se testar essa teoria, foram feitos testes de janela móvel. Nesse teste, ao invés de calcular apenas um erro para todo o intervalo da seção de testes do horizonte, tentou-se definir o erro logo após o fim da predição. Dessa maneira, calculou-se o erro dos 3 modelos para 5%, 10% e 15% de avanço (em relação a quantidade de dados total) de forma incremental. Ou seja, apenas uma fração dos dados de teste logo após a posição da janela (equivalente ao horizonte, nesse ensaio). A Figura 8 mostra um exemplo desse teste, sendo que a região entre a linha laranja e vermelha (nesse caso, em 15 %) é a região a ser testada. Nota-se que tudo a direita da janela seria o que normalmente seria testado no caso de horizonte deslizante.

Com este ensaio, consegue-se obter um mapa da magnitude do erro logo após o treinamento da rede, ao invés de um erro global. Isso é interessante para uma aplicação *online*,

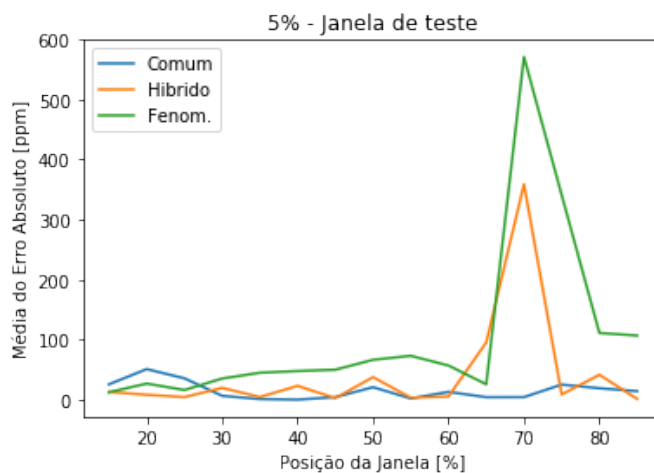
Figura 8: Exemplo do Ensaio da Janela Móvel.



Fonte: Autor

pois quando o modelo estiver operando nessa situação, a rede estaria sendo constantemente retreinada e todas as previsões teriam um horizonte definido. O resultados dessa análise se encontram na Figura 9, 10 e 11.

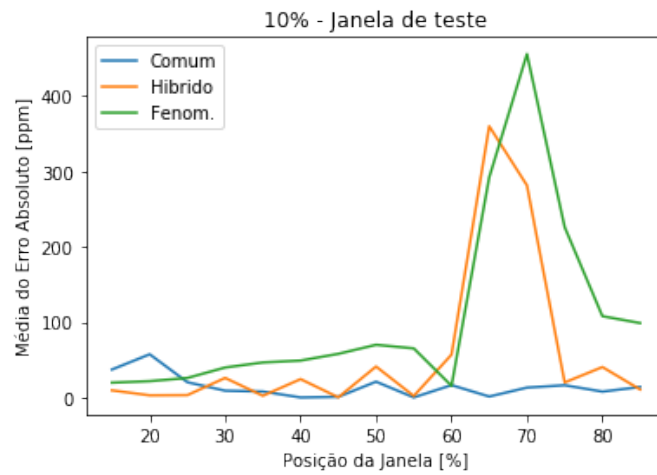
Figura 9: Resultados do primeiro ensaio de janela móvel de 5% com avanço incremental.



Fonte: Autor.

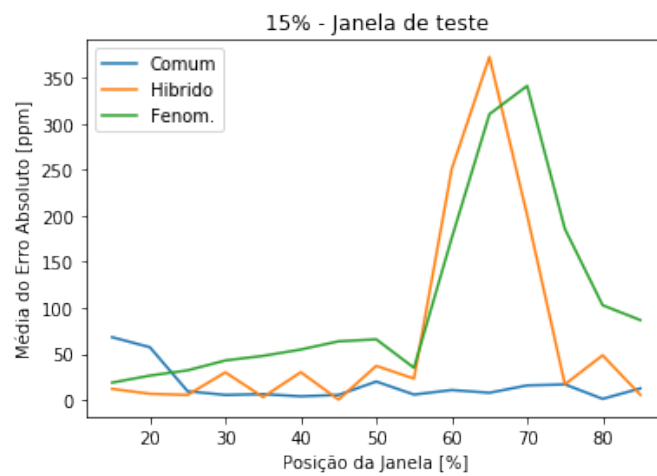
Nesse ensaio é possível observar que o modelo híbrido apresenta um erro menor que o fenomenológico e o comum desde o início dos ensaios; porém, seu erro varia abruptamente junto com o modelo fenomenológico.

Figura 10: Resultados do primeiro ensaio de janela móvel de 10% com avanço incremental.



Fonte: Autor.

Figura 11: Resultados do primeiro ensaio de janela móvel de 15% com avanço incremental.

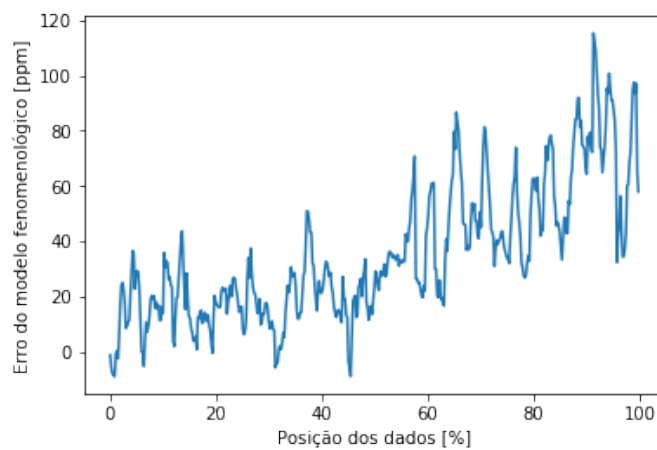


Fonte: Autor.

4.3 Horizonte Deslizante e Janela Móvel - Remoção do Erro Fenomenológico

Decidiu-se, em função dessa grande variação de erro repetir os ensaios do horizonte deslizante, mas sem a parte final dos dados, para obter uma melhor relação entre o modelo híbrido e o comum na condição onde o modelo fenomenológico esta operando corretamente. Nessa novo ensaio, utilizam-se apenas os 471 dias iniciais, de forma a evitar o pico de variação do erro do modelo fenomenológico. Observa-se o erro do modelo fenomenológico para esse intervalo na Figura 12:

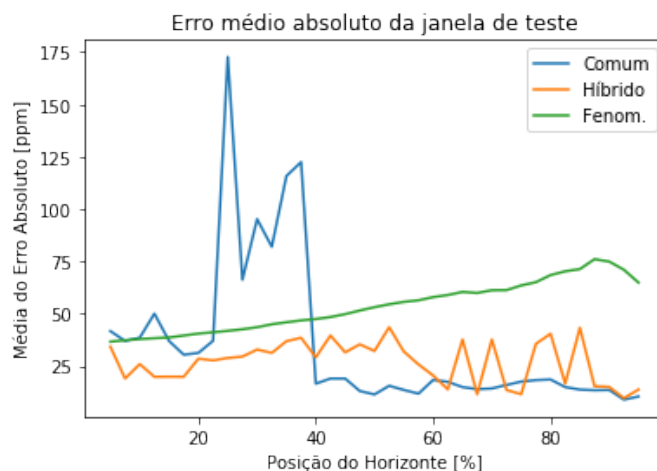
Figura 12: Erro do modelo fenomenológico para a seção inicial dos dados.



Fonte: Autor.

Para visualizar melhor o treinamento das redes, repetiu-se novamente o ensaio com avanço menor, de 2,5% e também começando na posição de 5% e terminando na de 95%, comparando o erro médio total na janela de teste inteira para esse grupo de dados. O resultado esta mostrado nas Figuras 13 e 14.

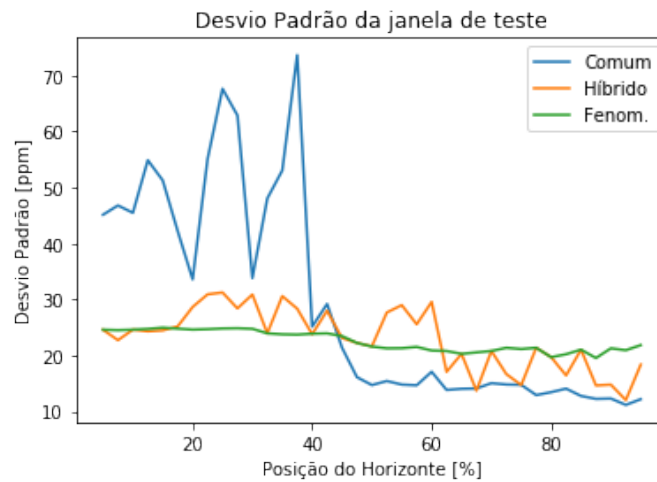
Figura 13: Média absoluta do ensaio do horizonte deslizante da seção inicial dos dados.



Fonte: Autor.

Nesse ensaio observa-se uma grande redução do erro em relação ao apresentado na Figura 13. Decidiu-se realizar também o mesmo ensaio de janela móvel, comparando o

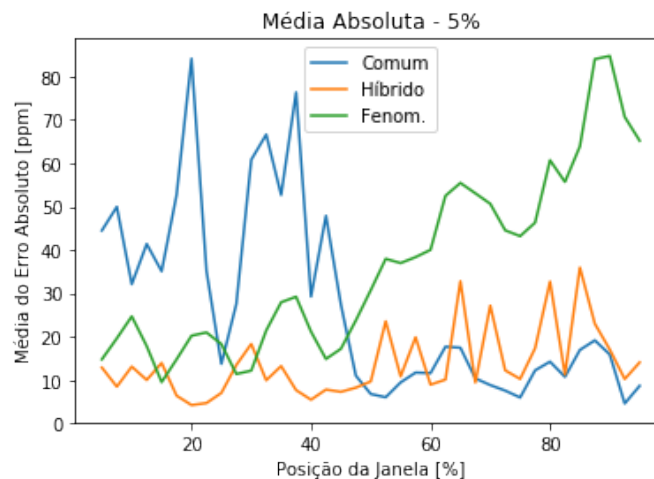
Figura 14: Média absoluta do ensaio do horizonte deslizante da seção inicial dos dados.



Fonte: Autor.

erro com avanços de 5%, 10% e 15% (equivalente a 30, 60 e 90 dias, respectivamente), também começando na posição de 5% dos dados e terminando na 95% e com avanço de 2,5%. Um detalhe importante é que como a quantidade de dados é menor nesse ensaio do que o realizado anteriormente, esse avanço representa uma quantidade menor de pontos absoluto. Os resultados da média absoluta desse ensaio se encontram nas Figuras 15, 16 e 17. Também pode se observar o resultado do desvio padrão dessa análise nas Figuras 18, 19 e 20.

Figura 15: Média absoluta do ensaio de janela móvel de 5%.

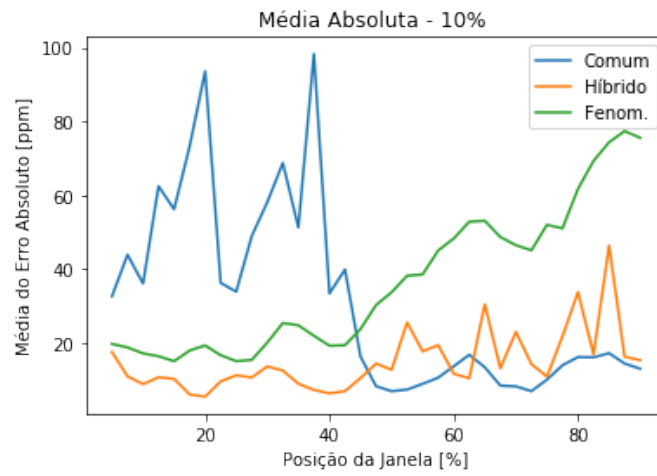


Fonte: Autor.

Esse resultado mostra duas características muito interessantes. Primeiramente, como esperado, o modelo híbrido possui o melhor resultado no início da análise. Isso se deve ao fato que ele precisa aprender menos características que o modelo comum, de forma que ele consegue aprender mais rápido. Além disso, por ele ser construído em cima do modelo fenomenológico, ele parte já de um erro reduzido comparado ao comum.

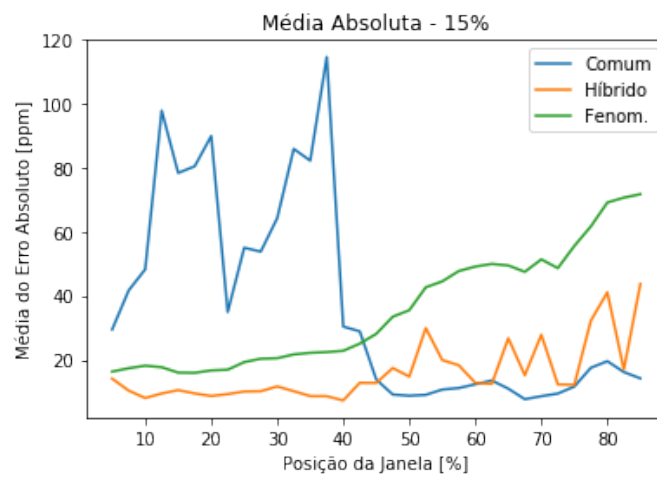
A segunda característica interessante diz respeito a variação do erro, que pode ser observado pelo desvio padrão nas Figuras 18, 19 e 20 e é maior no modelo híbrido do que

Figura 16: Média absoluta do ensaio de janela móvel de 10%.



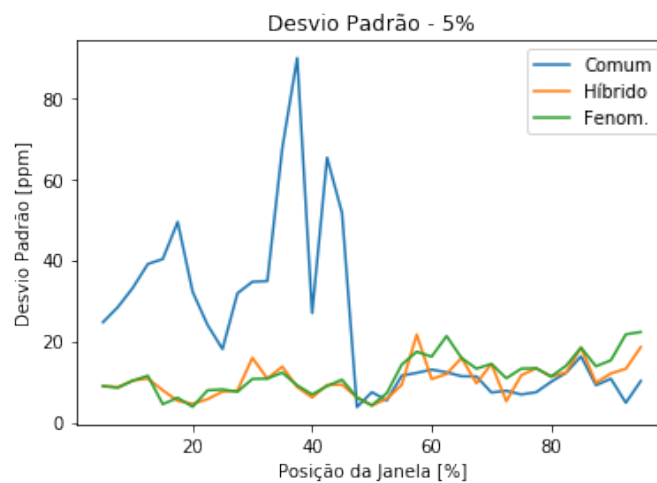
Fonte: Autor.

Figura 17: Média absoluta do ensaio de janela móvel de 15%.



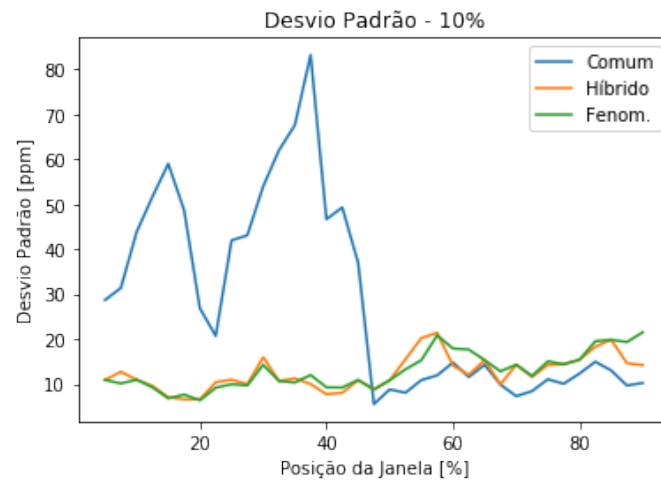
Fonte: Autor.

Figura 18: Desvio padrão do ensaio de janela móvel de 5%.



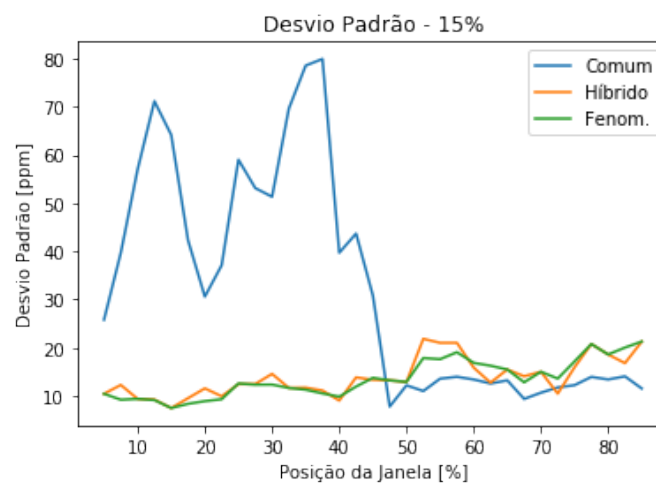
Fonte: Autor.

Figura 19: Desvio padrão do ensaio de janela móvel de 10%.



Fonte: Autor.

Figura 20: Desvio padrão do ensaio de janela móvel de 15%.



Fonte: Autor.

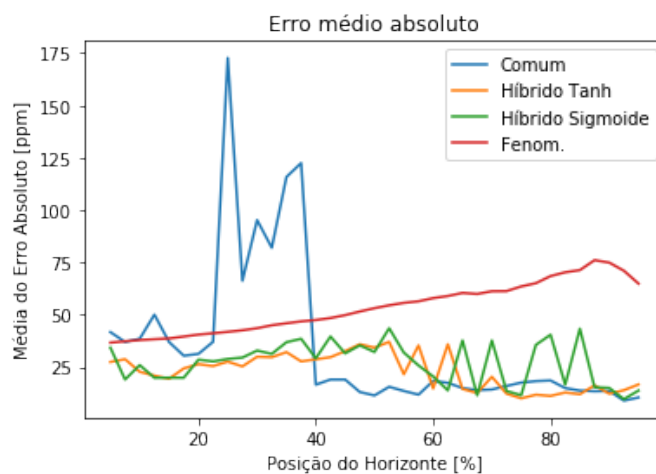
o comum para casos onde a janela esta no fim de seu curso (acima de 75%), sendo que ela tende inclusive a se aproximar do modelo fenomenológico. Por essa razão, o modelo comum acaba tendo menor variabilidade do que o modelo híbrido na condição da janela no fim de curso.

4.4 Ensaio da Janela Móvel - Rede Híbrida Alternativa

O fato dos modelos híbrido e fenomenológico tenderem ao mesmo desvio padrão indica que o modelo híbrido ainda não está conseguindo extrapolar completamente o resultado, e está tendo um erro induzido pelo modelo fenomenológico. Por esta razão, decidiu-se adequar a rede híbrida utilizada para uma com estrutura desenvolvida para prever o erro do modelo fenomenológico, ao invés de utilizar uma estrutura igual ao da rede comum. Para construir a nova rede híbrida, alterou-se a função de ativação da última camada por uma função tangente hiperbólica, já que os erros podem ser negativos ou positivos, então a função *tanh* acaba sendo mais indicada por ter um intervalo [-1, 1]. Todos os outros parâmetros da rede permaneceram iguais.

Testou-se os resultados com essa nova rede híbrida. Estes podem ser observados nas Figuras 21-28.

Figura 21: Resultados da média absoluta do horizonte deslizante para o novo modelo híbrido.



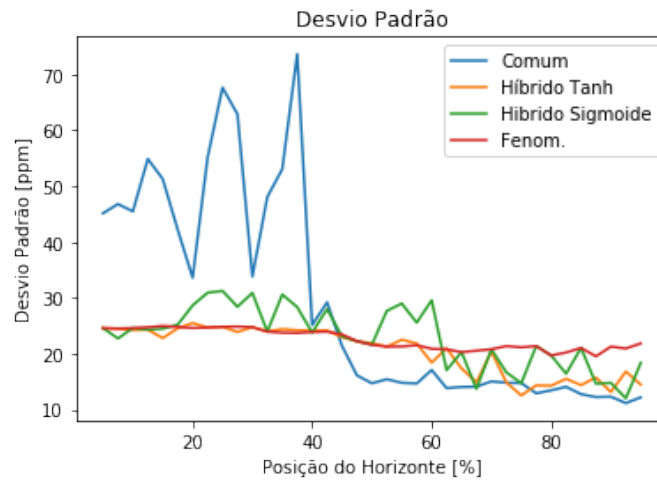
Fonte: Autor.

Nesses resultados, observa-se que o desvio padrão da nova rede híbrida está tendendo ao da rede comum nas condições finais de janela. Além disso, a média absoluta do erro também está mais próxima entre os dois modelos, sendo inclusive melhor do que a comum em diversas posições da janela. Isso é realizado pelo modelo sem perder a característica de obter boas médias na condição de posição da janela menor que 40 %.

Na Figura 29, é feita uma predição com o modelo híbrido alternativo (treinado com 10% iniciais dos dados), e é comparado com o modelo fenomenológico e os dados reais da planta. Observa-se um erro menor no modelo híbrido, e que ele foi capaz de corrigir a variação temporal do erro do modelo fenomenológico. Um gráfico de dispersão desse mesmo ensaio também foi feito, de forma a exemplificar os resultados da rede neuronal híbrida e o modelo fenomenológico. Esse resultado se encontra na Figura 30.

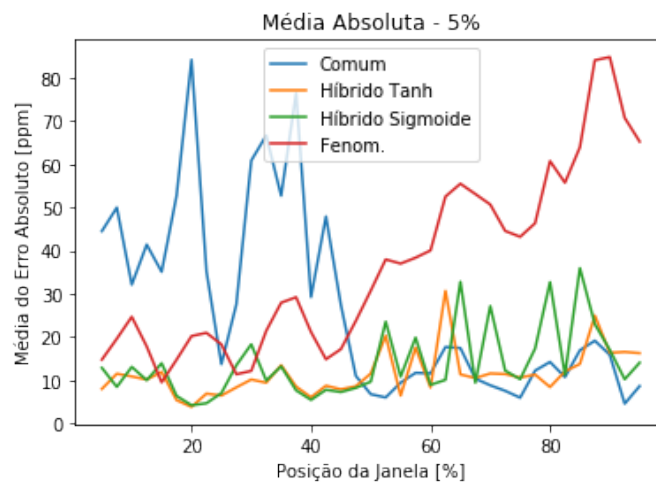
Um outro teste realizado para esse ensaio é o cálculo do coeficiente de Pearson. Esse

Figura 22: Resultados do desvio padrão do horizonte deslizante.



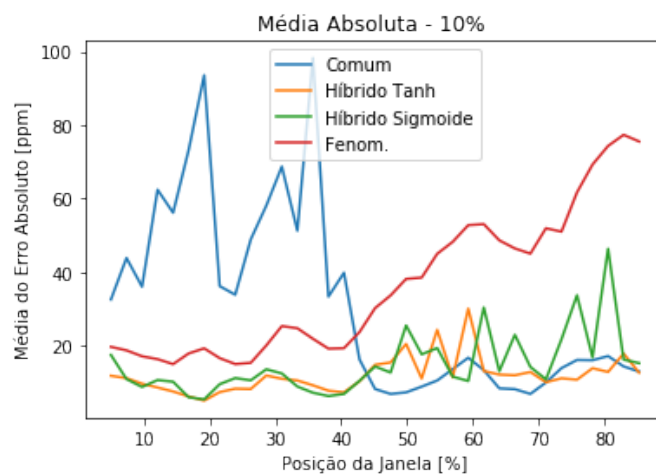
Fonte: Autor.

Figura 23: Resultados do ensaio de janela móvel de 5% para o modelo híbrido alternativo.



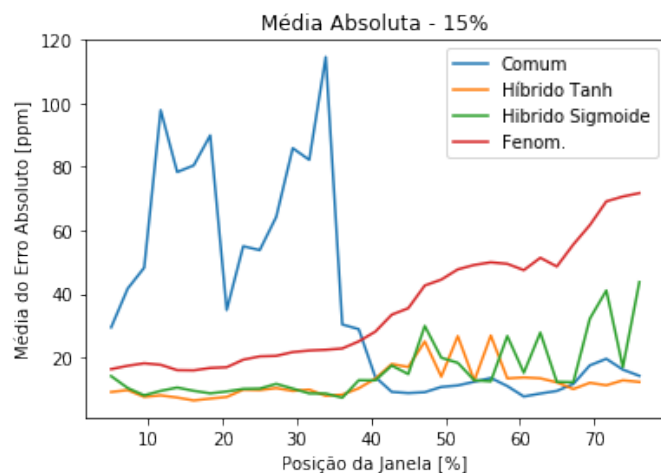
Fonte: Autor.

Figura 24: Resultados do ensaio de janela móvel de 10% para o modelo híbrido alternativo.



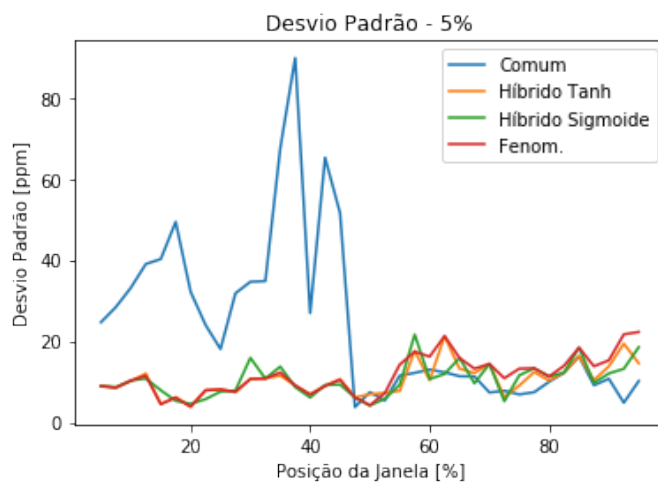
Fonte: Autor.

Figura 25: Resultados do ensaio de janela móvel de 15% para o modelo híbrido alternativo.



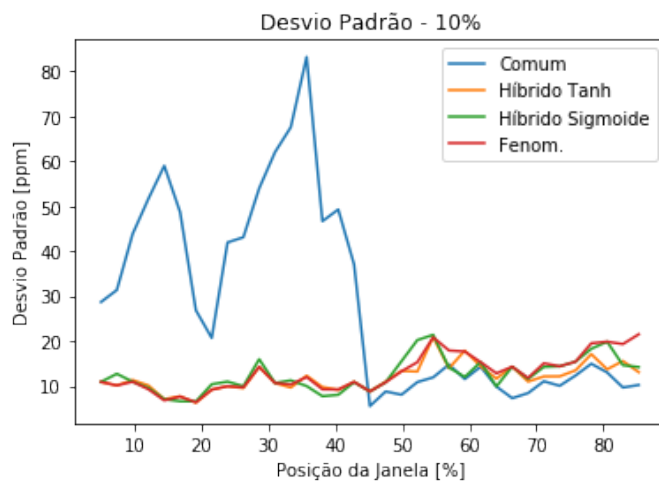
Fonte: Autor.

Figura 26: Desvio Padrão do ensaio de janela móvel de 5% para o novo modelo híbrido.



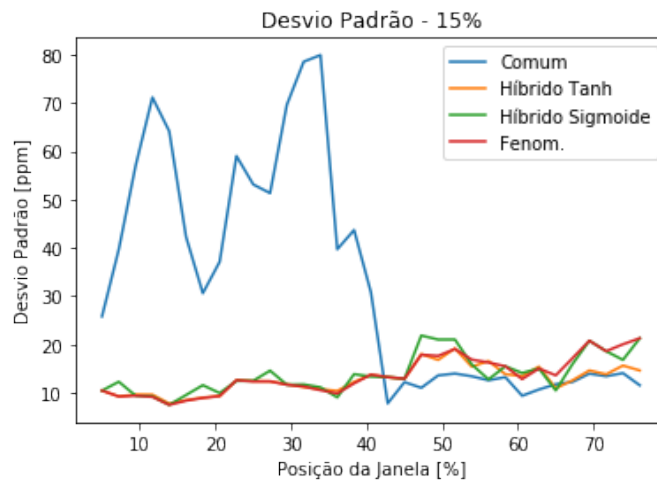
Fonte: Autor.

Figura 27: Desvio Padrão do ensaio de janela móvel de 10% para o novo modelo híbrido.



Fonte: Autor.

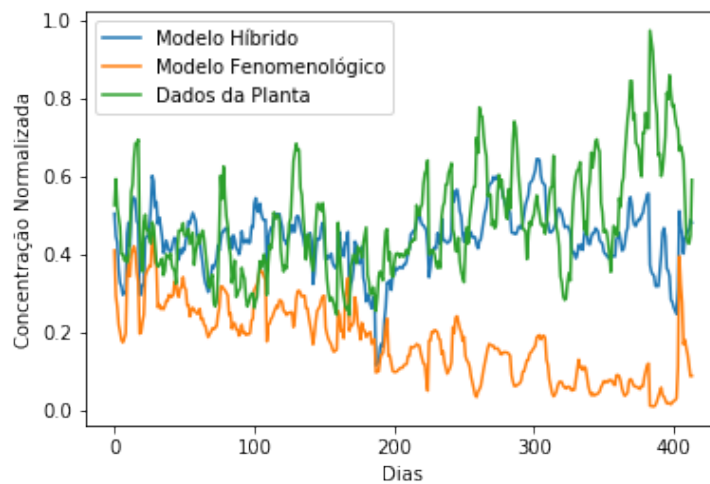
Figura 28: Desvio Padrão do ensaio de janela móvel de 15% para o novo modelo híbrido.



Fonte: Autor.

coeficiente indica as correlações entre o modelo e os dados, onde o valor "1" representa uma correlação perfeita. Os resultados dessa análise para o teste de janela móvel de 15% se encontram na Tabela 3. Os resultados para as demais janelas, assim como para o horizonte deslizando se encontram no Apêndice A.

Figura 29: Predição do modelo híbrido e fenomenológico frente aos dados da planta.



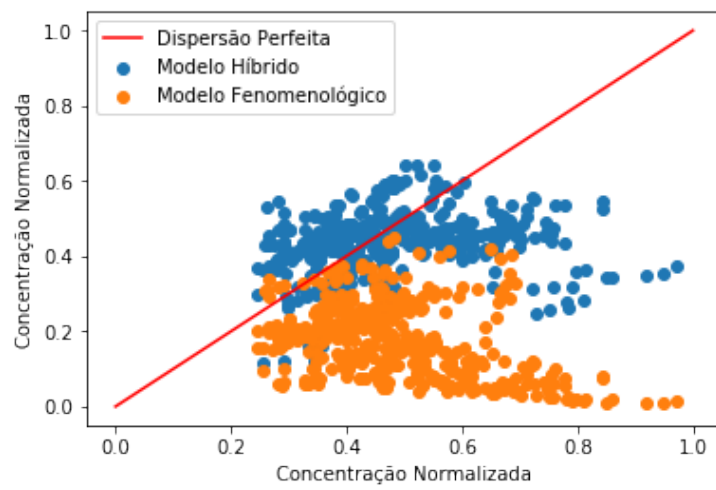
Fonte: Autor.

Tabela 3: Coeficientes de Pearson para as diferentes posições da janela e os diferentes modelos no ensaio de 15%.

	Comum	Fenom.	Hibrido Sig.	Hibrido <i>Tanh</i>
5.0%	0.548	0.748	0.748	0.748
7.5%	-0.148	0.505	0.191	0.503
10.0%	-0.002	0.469	0.469	0.438
12.5%	-0.415	0.585	0.568	0.515
15.0%	0.157	0.560	0.519	0.535
17.5%	0.707	0.519	0.336	0.506
20.0%	0.493	0.476	0.204	0.467
22.5%	0.843	0.423	0.414	0.423
25.0%	0.555	0.395	0.424	0.395
27.5%	0.717	0.287	0.259	0.271
30.0%	0.531	0.387	0.370	0.387
32.5%	-0.199	0.448	0.448	0.433
35.0%	-0.403	0.526	0.461	0.534
37.5%	-0.502	0.574	0.529	0.547
40.0%	0.022	0.282	0.439	0.229
42.5%	0.275	-0.123	0.055	-0.160
45.0%	0.057	-0.346	-0.133	-0.350
47.5%	0.409	-0.271	-0.271	-0.271
50.0%	0.107	-0.329	-0.330	-0.218
52.5%	0.112	-0.624	-0.647	-0.649
55.0%	0.042	-0.775	-0.252	-0.255
57.5%	-0.055	-0.792	-0.294	-0.783
60.0%	0.186	-0.679	-0.053	0.103
62.5%	0.824	-0.628	0.228	-0.608
65.0%	0.235	-0.082	-0.082	0.243
67.5%	-0.014	0.029	0.196	0.254
70.0%	-0.177	-0.184	-0.184	-0.238
72.5%	-0.310	-0.100	0.478	0.414
75.0%	-0.403	-0.198	0.342	0.619
77.5%	-0.342	-0.302	-0.302	0.683
80.0%	-0.375	-0.517	-0.517	0.552
82.5%	-0.256	-0.246	0.110	0.366
85.0%	0.457	-0.315	-0.315	0.548
87.5%	0.476	-0.357	0.346	0.446
90.0%	0.164	-0.376	0.456	0.584
92.5%	0.268	-0.321	0.743	0.574
95.0%	0.445	-0.216	0.580	0.663

Fonte: Autor

Figura 30: Gráfico de dispersão dos modelos híbrido e fenomenológico.



Fonte: Autor.

5 CONCLUSÃO

Este trabalho apresenta uma modelagem híbrida para reatores de hidrotratamento. É mostrado que estas redes apresentam um resultado melhor que redes convencionais em condições com poucos dados, e são capazes de obter um resultado equivalente a redes comuns quando esta tem uma quantidade suficiente de dados para serem treinadas. Além disso, é apresentado que ambas as redes são capazes de reduzir o erro da predição do processo relativo ao modelo fenomenológico.

É mostrado também que a rede híbrida tem dificuldade de extrapolar os resultados em situações que o modelo fenomenológico muda de comportamento. Isso implica que para essa abordagem de predição funcionar, o erro do modelo fenomenológico não pode variar abruptamente.

Esses resultados indicam que o modelo híbrido é capaz de aprender corretamente o erro gerado pelo modelo fenomenológico enquanto este não divergir de seu padrão normal. De forma que essa abordagem se torna interessante no momento que é possível garantir uma confiabilidade mínima do modelo fenomenológico.

As características apresentadas pelo modelo híbrido o torna de grande interesse para situações onde é importante haver uma estimativa do comportamento físico do sistema, aplicações *online* ou para dar início a um novo projeto. Tendo uma pequena quantidade de dados ele já se mostra capaz de otimizar a resposta do modelo fenomenológico e gerar resultados, ao contrário de um modelo comum, que necessita de mais informação para ser treinado. Além disso, o fato de ambos os modelos apresentarem resultados semelhantes com maiores quantidades de dados significa que após o modelo híbrido ser implementado, não há necessidade de substituir modelos.

BIBLIOGRAFIA

AQUINO GONÇALVES, B. T. de. **Modelagem, Simulação e Estimação de Parâmetros de um Reator de Hidrodessulfurização**. 2017. Diss. (Mestrado) – Universidade Federal do Rio Grande do Sul.

AQUINO GONÇALVES, B. T. de. Modelagem, Simulação e Estimação de Parâmetros de um Reator de Hidrotratamento (HDT) com Análise de Identificabilidade. **CICAP**, 2017.

CHAICHANA, K. et al. Neural network hybrid model of a direct internal reforming solid oxide fuel cell. **International Journal of Hydrogen Energy**, 2012.

CHOLLET, F. et al. **Keras**. [S.l.: s.n.], 2015. Disponível em: <<https://keras.io>>.

CYBENKOT, G. Approximation by Superpositions of a Sigmoidal Function. **Mathematics of Control, Signals, and Systems**, 1989.

DIETZ, A. C. **Modelagem e Simulação de Reatores de Hidrotratamento (HDT) de Correntes de Diesel**. 2014. Diss. (Mestrado) – UFRGS.

FACURE, M. **Introdução às Redes Neurais Artificiais**. Português. [S.l.: s.n.], 2017. Disponível em: <<https://matheusfacure.github.io/2017/03/05/ann-intro/>>.

GLOROT, X.; BORDES, A.; BENGIO, Y. Deep Sparse Rectifier Neural Networks. In: INTERNATIONAL Conference on Artificial Intelligence and Statistics. [S.l.: s.n.], 2011.

HAYKIN, S. **Neural Networks and Learning Machines**. Edição: Inc. Pearson Education. [S.l.]: Pearson Education, Inc., 2009.

JIN, J.; LI, M.; JIN, L. Data Normalization to Accelerate Training for Linear Neural Net to Predict Tropical Cyclone Tracks. **Mathematical Problems in Engineering**, 2014.

KINGMA, D. P.; BA, J. L. Adam: A Method for Stochastic Optimization. In: ICLR - International Conference on Learning Representations. [S.l.: s.n.], 2015.

PRECHELT, L. Neural Networks: Tricks of the Trade. In: edição: Heidelberg Springer Berlin. [S.l.]: Springer, Berlin, Heidelberg, 2012. Early Stopping - but when?, p. 53–67.

PSICHOGIOS, D. C.; UNGAR, L. H. A Hybrid Neural Network-First Principles Approach to Process Modeling. **AIChE Journal**, 1992.

QI, H. et al. A hybrid neural network-”rst principles model for fixed-bed reactor. **Chemical Engineering Science**, 1999.

RAMACHANDRAN, P.; ZOPH, B.; LE, Q. V. **Searching for Activation Functions**. [S.l.], 2017.

SRIVASTAVA, N. et al. Dropout: A Simple Way to Prevent Neural Networks from Overfitting. **Journal of Machine Learning Research**, 2015.

APÊNDICE A TABELAS DE COEFICIENTE DE PEARSON

Se encontram nesse apêndice as tabelas de coeficiente de pearson para o teste de horizonte deslizando, assim como os de janela móvel de 5% e 10%. Estes valores são referentes ao quarto ensaio realizado, descrito na Seção 4.4.

Tabela 4: Coeficientes de Pearson para as diferentes posições da janela e os diferentes modelos no ensaio de 5%.

	Comum	Fenom.	Hibrido Sig.	Hibrido <i>Tanh</i>
5.0%	0.736	0.880	0.880	0.880
7.5%	-0.165	0.607	0.159	0.607
10.0%	0.134	0.618	0.618	0.624
12.5%	-0.664	0.463	0.534	0.404
15.0%	0.201	0.738	-0.288	0.724
17.5%	0.644	0.069	0.317	0.012
20.0%	0.297	0.737	0.701	0.720
22.5%	0.038	0.878	0.885	0.878
25.0%	0.550	0.940	0.897	0.940
27.5%	0.765	0.138	-0.046	0.143
30.0%	0.739	0.164	-0.186	0.165
32.5%	0.180	-0.105	-0.105	-0.104
35.0%	0.032	0.224	0.044	0.453
37.5%	-0.664	0.667	0.661	0.673
40.0%	-0.230	0.936	0.903	0.934
42.5%	-0.006	0.251	0.707	0.246
45.0%	-0.223	-0.254	0.008	-0.278
47.5%	0.808	0.552	0.552	0.552
50.0%	-0.205	0.569	0.571	0.395
52.5%	0.574	0.389	0.673	0.369
55.0%	0.771	-0.898	0.478	0.695
57.5%	-0.296	-0.889	-0.556	-0.895
60.0%	-0.321	-0.727	0.195	0.306
62.5%	0.825	-0.708	0.718	-0.667
65.0%	0.537	-0.831	-0.831	-0.060
67.5%	0.869	-0.692	0.484	0.428
70.0%	0.666	-0.567	-0.567	-0.567
72.5%	0.005	-0.450	0.579	0.405
75.0%	0.059	0.063	0.561	0.692
77.5%	0.682	0.466	0.466	0.512
80.0%	0.265	-0.454	-0.454	0.447
82.5%	-0.234	-0.858	0.167	0.340
85.0%	0.427	-0.367	-0.367	0.393
87.5%	0.656	-0.443	0.590	0.681
90.0%	0.000	-0.491	0.364	0.441
92.5%	0.452	-0.426	0.783	0.580
95.0%	0.701	-0.270	0.589	0.670

Fonte: Autor

Tabela 5: Coeficientes de Pearson para as diferentes posições da janela e os diferentes modelos no ensaio de 10%.

	Comum	Fenom.	Hibrido Sig.	Hibrido <i>Tanh</i>
5.0%	0.462	0.767	0.767	0.768
7.5%	-0.145	0.507	0.307	0.506
10.0%	-0.007	0.428	0.428	0.389
12.5%	-0.178	0.504	0.453	0.367
15.0%	0.055	0.400	0.140	0.379
17.5%	0.729	0.551	0.694	0.541
20.0%	0.458	0.826	0.755	0.831
22.5%	0.568	0.541	0.424	0.541
25.0%	0.601	0.455	0.356	0.455
27.5%	0.819	0.214	0.115	0.233
30.0%	0.645	0.120	0.074	0.119
32.5%	0.048	0.295	0.295	0.301
35.0%	-0.206	0.734	0.567	0.790
37.5%	-0.615	0.577	0.756	0.534
40.0%	-0.138	0.429	0.650	0.353
42.5%	0.137	0.184	0.615	0.191
45.0%	-0.011	-0.143	0.078	-0.160
47.5%	0.482	0.166	0.166	0.166
50.0%	0.410	-0.126	-0.125	-0.016
52.5%	0.663	-0.458	-0.435	-0.487
55.0%	0.035	-0.789	-0.358	-0.270
57.5%	-0.259	-0.799	-0.257	-0.793
60.0%	0.006	-0.723	0.165	0.276
62.5%	0.835	-0.695	0.450	-0.660
65.0%	0.766	-0.642	-0.642	-0.184
67.5%	0.764	-0.568	0.362	0.291
70.0%	0.412	0.011	0.011	0.011
72.5%	-0.213	0.158	0.136	0.314
75.0%	-0.243	-0.147	0.408	0.551
77.5%	0.069	-0.106	-0.106	0.542
80.0%	-0.281	-0.340	-0.340	0.396
82.5%	-0.259	-0.442	0.034	0.335
85.0%	0.455	-0.511	-0.511	0.667
87.5%	0.537	-0.433	0.253	0.464
90.0%	0.000	-0.435	0.491	0.579
92.5%	0.268	-0.321	0.743	0.574
95.0%	0.445	-0.216	0.580	0.663

Fonte: Autor

Tabela 6: Coeficientes de Pearson para as diferentes posições do horizonte deslizante.

	Comum	Fenom.	Hibrido Sig.	Hibrido <i>Tanh</i>
5.0%	0.757	-0.237	-0.237	-0.237
7.5%	0.664	-0.402	-0.316	-0.401
10.0%	0.744	-0.430	-0.430	-0.434
12.5%	0.634	-0.437	-0.421	-0.455
15.0%	0.613	-0.495	-0.482	-0.428
17.5%	0.808	-0.508	-0.542	-0.506
20.0%	0.885	-0.495	-0.534	-0.505
22.5%	0.648	-0.501	-0.565	-0.501
25.0%	-0.153	-0.496	-0.606	-0.496
27.5%	0.370	-0.518	-0.585	-0.491
30.0%	0.248	-0.506	-0.558	-0.506
32.5%	0.025	-0.477	-0.477	-0.461
35.0%	-0.389	-0.471	-0.646	-0.500
37.5%	-0.367	-0.471	-0.470	-0.483
40.0%	0.150	-0.533	-0.488	-0.540
42.5%	0.066	-0.555	-0.519	-0.563
45.0%	-0.047	-0.545	-0.487	-0.512
47.5%	-0.258	-0.493	-0.493	-0.493
50.0%	0.057	-0.463	-0.467	-0.471
52.5%	-0.182	-0.461	-0.618	-0.447
55.0%	-0.110	-0.470	-0.316	-0.345
57.5%	0.160	-0.458	-0.357	-0.493
60.0%	-0.271	-0.404	-0.202	0.137
62.5%	-0.031	-0.373	0.210	-0.381
65.0%	0.007	-0.324	-0.324	0.165
67.5%	-0.082	-0.323	0.601	0.502
70.0%	-0.285	-0.299	-0.299	-0.438
72.5%	-0.288	-0.292	0.398	0.620
75.0%	-0.329	-0.237	0.634	0.734
77.5%	-0.022	-0.204	-0.204	0.655
80.0%	-0.148	-0.284	-0.284	0.529
82.5%	-0.306	-0.270	0.248	0.410
85.0%	0.411	-0.287	-0.287	0.563
87.5%	0.379	-0.357	0.346	0.446
90.0%	0.164	-0.376	0.456	0.584
92.5%	0.268	-0.321	0.743	0.574
95.0%	0.445	-0.216	0.580	0.663

Fonte: Autor