DETECÇÃO DE FALHAS BASEADA EM DADOS EM SISTEMA DE ARREFECIMENTO DE MOTOR DIESEL

Daniel Augusto do Carmo*, Celso J. Munaro†

*Av. Comandante Álvaro Martins, 71 Universidade Federal do Espírito Santo Vitória, ES, Brasil

> †DEL-CT-UFES Av FernandoFerrari, 545 Vitória, ES

Emails: danielaugustoc@gmail.com, celso.munaro@ufes.br

Abstract— This work presents a method to detect faults in the cooling system of railroad diesel engines. This type of equipment operates at different rotational speeds, which affect the variables used in its control and protection systems. A classifier is proposed to identify the operating point from rotation measurements. Data-based models are identified to estimate the internal pressure of the cooling system and to generate residuals which are used for fault detection. The multiple operation modes require the use of multiple models and different standardization of residuals at these operation points. Data were collected on a fault-free situation and a leaking situation, through a specific data acquisition system developed, installed in a real diesel engine. The application of the proposed methodology allowed to detect a fault caused by water leakage with great anticipation, which is a very desirable feature in fault detection systems. The proposed methodology is easily applicable to other diesel engines.

Keywords— Diesel Engine, Fault diagnosis.

Resumo— Este trabalho apresenta um método para detecção de falhas no sistema de arrefecimento de um motor diesel ferroviário. Este tipo de equipamento opera com diferentes rotações, que afetam as variáveis usadas em seu sistema de controle e proteção. Propõe-se um classificador para identificar o ponto de operação a partir de medidas de rotação. Modelos baseados em dados são identificados para estimar a pressão interna do sistema de arrefecimento, e com elas gerar resíduos usados para a detecção da falha. Os múltiplos modos de operação requerem o uso de múltiplos modelos e diferentes normalizações dos resíduos nestes pontos de operação. Foram coletados dados de operação normal e de falha de um motor diesel real, através de um sistema desenvolvido para a aquisição dos dados. A aplicação da metodologia proposta permitiu detectar uma falha causada por vazamento de água com grande antecipação, sendo esta uma caraterística muito desejável em sistemas de detecção de falhas. A metodologia proposta é facilmente aplicável a outros motores diesel.

 ${\bf Palavras\text{-}chave} - \quad {\rm Motor\ Diesel,\ Diagn\'ostico\ de\ falhas}.$

1 Introdução

Motores diesel são amplamente empregados em muitos setores da economia, tais como: indústria automotiva, geração de energia, máquinas de operação pesada, ferroviária, indústria naval, amplificadores de força. Por sua importância, faz-se necessário monitorar sua operação de modo a gerar alertas em caso de falhas ou eminência de falhas, evitando assim, paradas inesperadas de grande impacto econômico.

Contribuições ao problema de detecção e diagnóstico de falhas em motores diesel têm sido apresentadas com diferentes abordagens, em sua maioria são utilizados modelos paramétricos. Alguns desses trabalhos detectam vazamentos no sistema de ar de motores tipo otto e diesel empregando filtros de Kalman (Nyberg, 2002), observadores de estado (Ceccarelli et al., 2009), equações de paridade (Kimmich et al., 2005) e análise de resíduos dos modelos (Khelil et al., 2012).

O objetivo deste trabalho é propor uma metodologia para detecção de vazamento de água em um sistema de arrefecimento do motor diesel, usando modelos baseados em dados. Na aplicação locomotiva considerada, o motor trabalha em vários pontos de operação (potências), o que limita os métodos que podem ser aplicados. A opção de usar um modelo baseado em dados representa uma alternativa de menor complexidade quando comparada ao uso de modelos fenomenológicos.

Uma breve revisão do método para detecção de falhas para o caso de múltiplos pontos de operação é apresentada na Seção 2. Na Seção 3 é apresentada uma breve descrição do sistema de arrefecimento de motores diesel e as consequências que vazamentos acarretam ao sistema. Na Seção 4 é apresentada a metodologia utilizada para a construção e seleção dos modelos não paramétricos, geração dos resíduos e método de detecção de vazamento. As discussões e resultados são apresentadas na Seção 5 e, por fim, na Seção 6 são feitas as conclusões sobre o trabalho.

2 Revisão Bibliográfica

2.1 Detecção de falhas via análise de resíduos

Um dos grandes desafios no universo de sistemas de controle tem sido garantir a alta performance de equipamentos com amplo espectro de utilização e cada vez mais complexos. A detecção e diagnostico de falhas tem importante contribuição neste sentido (Ding, 2008).

O uso de modelos para estimar variáveis reduz o custo com redundância de equipamentos. Estes métodos são aqui considerados para estimar variáveis, compará-las com as variáveis medidas, e a partir dos resíduos gerados detectar e isolar falhas (Patan, 2008). No caso univariado, os resíduos podem ser monitorados com uma simples carta de Shewhart. Para o caso multivariado, a estatística T^2 de Hotelling deve ser utilizada (Chiang et al., 2000).

O primeiro passo para avaliar o modelo gerador dos resíduos é a escolha das métricas da qualidade dos modelos. Modelos que geram menores resíduos têm maior sensibilidade a falhas, sendo assim preferidos (Patan, 2008). Dois exemplos de métricas para isso são a raiz quadrada do erro quadrático médio (RMSE), denotado aqui por erro, e coeficiente de Spearman, (Kuhn and Johnson, 2013), dado por:

$$\rho = 1 - \frac{\sum_{i=1}^{n} d_i^2}{n(n^2 - 1)} \tag{1}$$

onde ρ é coeficiente de Spearman, d_i a diferença do posto da observação com o posto da estimativa e n o número de observações.

Na escolha dos modelos, pode-se considerar o uso de modelos lineares, não-lineares ou múltiplos modelos (lineares ou não), dependendo da aplicação.

Para a isolação da falha, usa-se algoritmos de classificação tais como *fuzzy*, redes neurais e árvores de decisão (Aldrich and Auret, 2016).

2.2 Operação em múltiplos pontos de operação

A maioria dos processos industriais opera em diferentes pontos de operação. Próximo a estes pontos seu comportamento é aproximadamente linear, e tende a tornar-se não-linear à medida que deles se afasta. Os comportamentos nestas regiões podem ser representados por modelos não lineares ou conjuntos de modelos lineares.

Para processos com estas características, deve-se investigar a metodologia mais adequada afim de obter os modelos que irão gerar os resíduos. O uso de modelos não lineares tem a desvantagem da maior dificuldade na escolha dos modelos e da obtenção de seus parâmetros. A vantagem é usar um único modelo para calcular os resíduos. O uso de modelos lineares é mais simples pelo fato

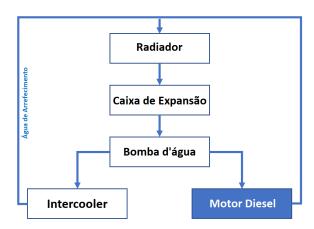


Figura 1: Esquema básico de um sistema de arrefecimento de motores.

de haver estruturas bem conhecidas e relações lineares de seus parâmetros. A desvantagem é o uso de um classificador para indicar o ponto de operação a cada instante e o modelo correspondente a ser usado. No caso de múltiplos modelos, podese considerar também o uso de múltiplos modelos não-lineares.

3 Sistema de arrefecimento do motor diesel

3.1 Descrição do sistema

O sistema de arrefecimento no motor diesel é responsável por manter a sua temperatura dentro dos parâmetros de projeto, evitando que os componentes atinjam limites que causem falhas catastróficas (são aquelas que impossibilitam a recuperação do equipamento). Caso esse sistema apresente falha e essa seja detectada, ela deve ser prontamente tratada. Caso não seja, a parada é inevitável.

O sistema de arrefecimento é usualmente constituído de bomba, radiadores, válvulas e tubulações externas e internas do motor. A bomba tem a função de garantir a pressão para a circulação do fluido de arrefecimento, usualmente água com aditivo. O radiador deve retirar o calor desse fluido que foi aquecido durante o arrefecimento do motor. As válvulas e tubulações têm a função de direcionar o fluxo para os locais corretos que necessitam perder calor.

A figura 1 apresenta o esquema básico do sistema de arrefecimento de um motor. Este sistema raramente é um sistema aberto de circulação de fluido de arrefecimento e pelas leis da termodinâmica pode-se determinar o estado do fluido medindo duas propriedades, no caso, a temperatura e a pressão.

3.2 Vazamento de água no sistema de arrefecimento

As falhas nesse sistema reduzem a capacidade de troca de calor com o ambiente, reduzindo sua eficiência. As causas mais comuns são sujeira nos trocadores de calor, entupimentos, vazamento do fluido de arrefecimento. Neste trabalho a falha será induzida via vazamento de água de arrefecimento, sendo feita apenas a detecção da falha, sem o diagnóstico.

Em motores diesel de grande porte que apresentam tanque de expansão antes da bomba, o vazamento reduz a coluna de água interna, reduzindo a pressão do sistema de arrefecimento. A temperatura medida subirá e oscilará mais, até que o sistema não seja mais capaz de retirar a quantidade de necessária calor, levando à parada obrigatória do motor. Como proteção, os motores são equipados com termostatos e pressostados, indicando falha ao usuário somente quando não há mais condições operacionais.

As variações na pressão podem ser causadas pela variação de temperatura, que por sua vez, é causada por variação da temperatura ambiente, ciclos de funcionamento do ventilador do radiador e principalmente a potência produzida pelo motor naquele determinado ponto de operação.

3.3 Detecção de vazamentos

No sistema de arrefecimento, a vazão pode ser diretamente aferida por medidores de vazão ou estimada a partir da pressão e da temperatura do fluido (estados termodinâmicos). Quedas de pressão nos sistemas em que as condições termodinâmicas são conhecidas indicam possíveis vazamentos

O uso de modelos pode indicar desvios da pressão medida em relação à prevista, e este desvio ser usado para detectar falhas. A detecção pode ser assim feita através de cartas de controle estatístico ou por meio de cartas de soma cumulativa. As cartas de soma cumulativa identificam mudanças da média dos resíduos em uma determinada janela de tempo. É sugerido que os limiares da carta sejam calibrados para que não ocorram falsos alarmes, (Eliades and Polycarpou, 2012; Misiunas et al., 2005).

4 Metodologia

4.1 Descrição do motor utilizado

O motor diesel utilizado neste trabalho tem aplicação ferroviária com as seguintes características: 4100HP, 16 cilindros de 10,5 litros cada, rotação máxima de 1150 rotações por minuto, peso 19 toneladas. Sua função principal é gerar a energia cinética para que o alternador da locomotiva produza a energia elétrica necessária para os motores

de tração. A figura 2 apresenta o motor diesel utilizado nos testes.

Este motor opera em 9 níveis distintos de potência com 5 velocidades diferentes de rotação (medida em RPM), que nesse caso, definem os pontos de operação. Neste motor, a bomba de água de arrefecimento é impulsionada pelo próprio eixo virabrequim. Isto acarreta 5 níveis diferentes de pressão, visto que, a rotação da bomba depende diretamente da rotação do motor.



Figura 2: Motor diesel utilizado nos testes Fonte: www.get.com

4.2 Aquisição dos dados

Os dados são coletados a partir do barramento de comunicação interna do sistema de controle do motor diesel. O barramento recebe os dados de sensores internos do motor diesel. São eles a rotação, posição do eixo virabrequim, pressão de ar, temperatura de ar, pressão de água, temperatura e pressão de água e óleo.

Foi desenvolvido hardware e software específicos para realizar a coleta desses dados, já devidamente tratados pelo sistema. A frequência de coleta dos dados é de 20Hz.

Três coletas de dados foram realizadas, sendo duas para treinamento e validação do classificador e para determinação dos modelos (total de 54.957 amostras), e uma com dados da falha causada pelo vazamento inserido no sistema de arrefecimento(12.136 amostras).

Na figura 3 são mostrados os sinais de um trecho dos dados de treinamento, nos quais observase a relação direta entre rotação e pressão. Observa-se as variações de pressão decorrentes da variação de temperatura, evidenciadas entre os instantes 300s e 700s. A rotação apresenta pequenas variações em torno de seu valor nominal.

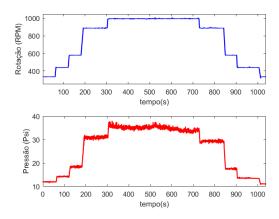


Figura 3: Relação entre a rotação e a pressão de água do motor.

4.3 Detecção do ponto de operação

Descreve-se aqui a metodologia que será usada para detectar o ponto de operação do motor. Ela é necessária para a aplicação e avaliação do desempenho do método proposto.

As transições de rotação do motor usado são rápidas quando comparadas ao tempo em que esse permanece nas rotações pré-estabelecidas do sistema de controle, que são: 335, 440, 580, 888 e 995 RPM. Nesses pontos de operação os dados são considerados estacionários, são nessas condições que se deseja detectar as falhas do sistema de arrefecimento. Durante as transições o classificador não pode indicar nenhum destes pontos, e o diagnóstico somente deve ser feito quando um ponto de operação for detectado.

Dentre os muitos classificadores existentes, utilizou-se o tipo k-NN, dada sua simplicidade. Definiu-se para seu treinamento 6 classes, 5 são as rotações pré-estabelecidas e uma para a transição entre rotações.

Os dados usados pelo classificador são a rotação e sua variação. Para rotular os dados, admitese como pertencente a cada classe rotações com valores de ± 10 da rotação nominal da classe. A variação da rotação é calculada pela diferença da rotação no instante menos a rotação medida 20 amostras antes (lembrando que a frequência de amostragem é $20 \, \mathrm{Hz}$).

A rotina knnsearch do Matlab com método de procura kd-tree será utilizada pelo classificador, com 10 vizinhos mais próximos medidos por distância euclidiana.

O desempenho do classificador deve atender uma especificidade acima de 95% (fração dos dados não pertencentes à uma classe corretamente classificados) e uma sensibilidade acima de 95% (fração dos dados pertencentes à uma classe corretamente classificados).

4.4 Modelos e geração dos resíduos

As técnicas de modelagem para determinar os modelos que representem um ou todos os pontos de operação testados neste trabalho com seus respectivos parâmetros, são:

- Rede neural de regressão generalizada (*GRNN*), com fator de suavização de 0.2.
- Regressão por máquinas de vetores de suporte (SVM) usando função gaussiana.
- Regressão Polinomial (POLY) com polinômios de até quinto grau.
- Árvore de regressão (*TREE*) com modelo de regressão particionado.

Estes modelos foram obtidos através das rotinas do *Matlab*: newgrnn, fitrsvm, fit e fitrtree respectivamente. Os parâmetros não citados acima foram os default dessas rotinas.

As técnicas de modelagem acima serão utilizadas para obtenção de modelos únicos que representam todos os pontos de operação e modelos múltiplos, um por ponto de operação. No primeiro caso, as entradas do modelo são a rotação e temperatura e a saída a pressão. O segundo caso, a saída é a mesma e a entrada é apenas a temperatura, pois a rotação do ponto de operação é conhecida.

Serão selecionados os modelos que obtiverem o maior coeficiente de Spearman com a condição de que o erro seja menor que o erro médio dentre os modelos gerados.

A validação cruzada é feita aplicando o repeated K-fold (Kuhn and Johnson, 2013; Refaeilzadeh et al., 2009) que utiliza sucessivos conjuntos (folds) gerando uma maior massa para análise estatística, tornando a seleção do modelo mais confiável. Os dados para gerar cada fold são selecionados aleatoriamente a cada iteração do repeated K-fold. São testadas todas as técnicas de determinação de modelos apresentadas.

Faz-se 50 repetições com 10 folds, resultando em 500 modelos a cada teste. Um maior número de repetições permite caracterizar melhor as estatísticas usadas para a escolha dos modelos. O número de repetições foi escolhido assegurando um baixo esforço computacional. A rotina crossval do Matlab é utilizada nessa etapa da modelagem.

Como o coeficiente de Spearman é calculado pelo posto dos dados estimados em relação ao posto dos dados reais, espera-se que o modelo assim selecionado apresente os menores erros de estimativas para novas entradas, mesmo não tendo necessariamente produzido o menor erro para o conjunto de dados usado.

4.5 Análise dos resíduos e detecção do vazamento

Uma forma usual para detecção de falhas é o emprego de estatística univariada, onde o vetor dos resíduos será avaliado utilizando cartas de controle de Shewhart. Se os resíduos gerados pelos modelos apresentarem diferentes médias e variâncias para cada ponto de operação, a normalização dos resíduos deve ser diferente em cada região, para que se possa utilizar um único limiar.

No caso de utilização de cartas de soma cumulativa, o vazamento é detectado no momento em que o limite inferior é atingido. Esse limite é calculado iterativamente no vetor de treinamento até que não sejam observados falsos positivos. A soma da carta é feita usando uma janela móvel. O tamanho da janela determina a sensibilidade da carta a desvios de médias, sendo de 3000 amostras neste trabalho.

Os modelos geram os resíduos durante a operação do motor. Caso o classificador indique que o motor esteja entre os pontos de operação (transição), deve-se atribuir ao resíduo valor zero.

5 Discussão e resultados

5.1 Pontos de operação

O classificador k-NN foi treinado para detectar os pontos de operação e regiões de transição. Foram selecionadas para o treinamento 20000 amostras (amostras 34000 a 54000 no gráfico superior da figurar 4). Essas amostras continham dados de todas as classes que se deseja treinar. As amostras foram rotuladas conforme descrito na seção de metodologia e obteve-se sensibilidade e especificidade de 100% usando os dados de validação (amostras 12000 a 32000 no gráfico inferior da figura 4).

Como a variação da rotação é calculada usando um atraso de 20 amostras, o classificador identifica a classe da rotação com atraso. Entretanto, dada a grande quantidade de amostras em estado estacionário, este erro é desprezível.

Na figura 5 são mostrados os dados de validação de rotação e a saída do classificador (gráfico inferior). As 6 classes de rotação são identificadas e durante as transições a saída do classificador é zero.

Os gráficos de dispersão de temperatura versus pressão para cada uma das classes são mostrados na figura 6. Observa-se se a existência da relação entre estas variáveis, destacando também o aumento da dispersão da pressão com o aumento da rotação.

Com a utilização do classificador é possível normalizar os resíduos por ponto de operação, viabilizando a detecção de falha por estatística univariada, possibilitando assim, a analise do resíduo com limiares constantes.

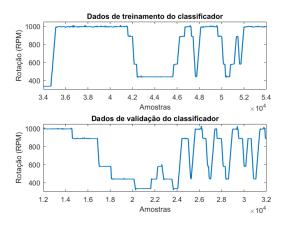


Figura 4: Dados utilizados para treinamento e validação do classificador k-NN.

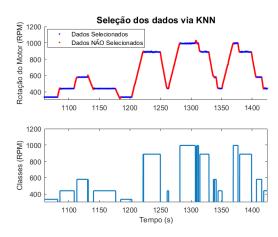


Figura 5: Classificação feita via k-NN.

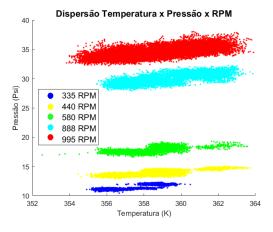


Figura 6: Dispersão por ponto de operação.

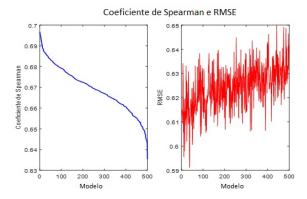


Figura 7: Comparação entre o coeficiente de Spearman e RMSE para regressão obtida via SVM no ponto de operação de 995 rotações por minuto.

5.2 Seleção dos modelos

Todos os 500 modelos gerados na validação cruzada, para modelos únicos e múltiplos, foram avaliados usando o coeficiente de Spearman e erro (RMSE). A figura 7 apresenta o resultado para 500 modelos, onde um deles é selecionado para uma destas situações.

A figura 8 apresenta os modelos obtidos pelas 4 técnicas de modelagem para cada um dos 5 pontos de operação. Os melhores modelos, baseados no coeficiente de Spearman, são indicados por setas. Observa-se que técnicas diferentes, obtiveram resultados distintos entre os 5 pontos de operação.

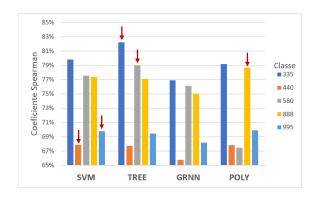


Figura 8: Comparativo entre o coeficiente de Spearman entre todas as técnicas de modelagem por ponto de operação.

O uso de múltiplos modelos foi então comparado ao uso de um único modelo para todos os pontos de operação. Para cada uma das 4 técnicas de modelagem obteve-se um modelo único e um modelo por ponto de operação. O resultado é mostrado na figura 9. Os resíduos dos 5 modelos foram concatenados para o cálculo do coeficiente de Spearman, de forma que o comprimento do vetor fosse igual ao caso de modelo único. Na figura 10 o mesmo foi feito para a análise do erro.

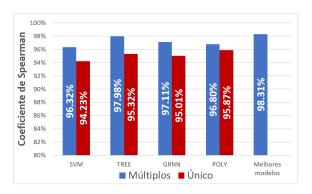


Figura 9: Comparativos entre os coeficientes de Spearman das técnicas empregadas.

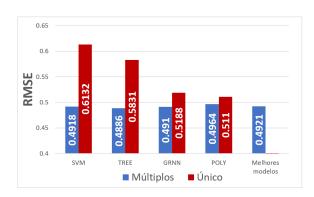


Figura 10: Comparativos entre os erros (RMSE) das técnicas empregadas.

Na técnica de regressão polinomial verificouse que não havia melhorias significativas com grau maior que 3. Para a rede neural, buscou-se melhorar os resultados da regressão generalizada variando a suavização, e o valor de 0.2 utilizado gerou os melhores resultados. Valores menores dificultavam muito a convergência do algoritmo de treinamento.

Todas as técnicas apresentaram resultados melhores com o uso de múltiplos modelos, ou seja, melhor coeficiente de Spearman com erros menores dos resíduos. O resultado denominado "Melhores modelos" foi obtido selecionando os modelos que deram melhor resultado por ponto de operação na etapa anterior (figura 8).

5.3 Tratamento e análise dos resíduos

A figura 11 apresenta a saída contínua dos resíduos não normalizados e normalizados pelas variâncias do ponto de operação, gerados pelos múltiplos modelos (dados concatenados). Fica clara a diferença da variância dos resíduos nos diferentes pontos de operação (rotações). Portanto, a normalização dos resíduos deve ser diferente em cada região para que se tenha um único limiar estatístico de teste. O uso do classificador para identificar estas regiões torna-se fundamental durante o monitoramento das falhas. Na figura 11

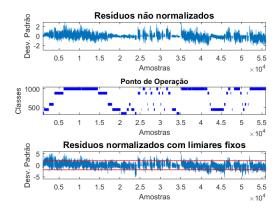


Figura 11: Resíduos não normalizado e resíduos normalizados com limiares fixos.

(inferior) os resíduos foram normalizados obedecendo a saída do classificador. O resíduo é zero nos instantes de transição dos pontos de operação do motor. Os limiares do gráfico inferior na mesma figura são calculados com um intervalo de confiança de 99%.

5.4 Detecção de vazamentos

Da figura 11 (inferior) pode-se concluir que a utilização da carta de controle de Shewhart não fornece bons resultados. Nenhuma das regras de Shewhart para indicar processo fora de controle é atendida, indicando vazamentos com frequência (falsos alarmes). Isso ocorre pela elevada frequência de coleta de dados.

Obteve-se melhor resultado para a detecção do vazamento com a carta de soma cumulativa. Os limiares inferior de -180 e superior de 180 desvios-padrão foram determinados iterativamente para que não houvesse falsos positivos no vetor de treinamento.

Como sabe-se da existência do vazamento, era esperado que a falha fosse identificada quando a soma cumulativa ultrapassasse o limiar inferior, como proposto pelas literaturas e como observado na figura 12 inferior. O vazamento foi detectado com um atraso de 220s após seu início. Este resultado é muito bom, pois estima-se que o sistema de proteção do motor levaria uma hora para ser acionado devido a esta falha.

É possível verificar pela figura 12, o momento em que o vazamento é detectado na carta de soma cumulativa, não retornando para o centro da carta mesmo após as sucessivas mudanças do ponto de operação.

6 Conclusões

Foi apresentada uma metodologia baseada em dados para detecção de falhas no sistema de arrefecimento de motores diesel. A aquisição de dados

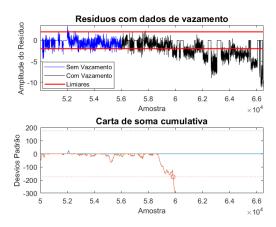


Figura 12: Detecção de Vazamento na Carta de Soma Cumulativa.

dos sensores existentes no motor permitiu identificar as características das variáveis sensíveis às falhas, bem como o efeito das diferentes rotações do motor sobre as mesmas. Com estas informações, diferentes modelos foram explorados para gerar estimativas da pressão do sistema de arrefecimento, que comparadas aos valores medidos, geraram resíduos usados para a detecção da falha.

Um classificador k-NN foi utilizado para identificar os pontos de operação (rotações) do motor, uma vez que os modelos eram distintos e os resíduos tinham diferentes variâncias nos diferentes pontos de operação. A partir da coleta de três conjuntos de dados, o classificador e os modelos foram treinados, validados e testados.

O resultado, usando carta de soma cumulativa mostrou que a falha no sistema de arrefecimento pode ser detectada com grande antecedência, podendo evitar uma parada indesejada no motor diesel.

A próxima etapa deste trabalho será gerar dados para diferentes falhas no mesmo sistema de arrefecimento, aproveitando o sistema de detecção de falhas e integrando-o um a sistema de diagnóstico para indicar qual delas ocorreu.

Referências

Aldrich, C. and Auret, L. (2016). Unsupervised process monitoring and fault diagnosis with machine learning methods, Springer.

Ceccarelli, R., Canudas-de Wit, C., Moulin, P. and Sciarretta, A. (2009). Model-based adaptive observers for intake leakage detection in diesel engines, *American Control Conference*, 2009. ACC'09., IEEE, pp. 1128–1133.

Chiang, L. H., Russell, E. L. and Braatz, R. D. (2000). Fault detection and diagnosis in industrial systems, Springer Science & Business Media.

- Ding, S. X. (2008). Model-based fault diagnosis techniques: design schemes, algorithms, and tools, Springer Science & Business Media.
- Eliades, D. and Polycarpou, M. (2012). Leakage fault detection in district metered areas of water distribution systems, *Journal of Hydroinformatics* **14**(4): 992–1005.
- Khelil, Y., Graton, G., Djeziri, M., Ouladsine, M. and Outbib, R. (2012). Fault detection and isolation in marine diesel engines: A generic methodology, *IFAC Proceedings Volumes* 45(20): 964–969.
- Kimmich, F., Schwarte, A. and Isermann, R. (2005). Fault detection for modern diesel engines using signal-and process model-based methods, *Control Engineering Practice* **13**(2): 189–203.
- Kuhn, M. and Johnson, K. (2013). Applied predictive modeling, Vol. 26, Springer.
- Misiunas, D., Lambert, M. F., Simpson, A. R. and Olsson, G. (2005). Burst detection and location in water transmission pipelines, *Impacts of Global Climate Change*, pp. 1–12.
- Nyberg, M. (2002). Model-based diagnosis of an automotive engine using several types of fault models, *IEEE Transactions on control systems technology* **10**(5): 679–689.
- Patan, K. (2008). Artificial neural networks for the modelling and fault diagnosis of technical processes, Springer.
- Refaeilzadeh, P., Tang, L. and Liu, H. (2009). Cross-validation, *Encyclopedia of database* systems, Springer, pp. 532–538.