

ANALISIS PROGNOSTIK TERHADAP KERUSAKAN BANTALAN PADA POROS KECEPATAN TINGGI TURBIN ANGIN MENGGUNAKAN *MACHINE LEARNING* DENGAN PENDEKATAN *SUPPORT VECTOR REGRESSION* (SVR)

*David Titian Wiranata¹, Toni Prahasto², Achmad Widodo³

¹Mahasiswa Jurusan Teknik Mesin, Fakultas Teknik, Universitas Diponegoro

²Dosen Jurusan Teknik Mesin, Fakultas Teknik, Universitas Diponegoro

³Dosen Jurusan Teknik Mesin, Fakultas Teknik, Universitas Diponegoro

Jl. Prof. Sudharto, SH., Tembalang-Semarang 50275, Telp. +62247460059

*E-mail: titian341@gmail.com

Abstrak

Produksi energi dari sumber terbarukan, seperti dari sumber air, udara dan sumber lainnya semakin meningkat. Tenaga angin adalah salah satu energi terbarukan yang paling berkembang di seluruh dunia. Untuk menghasilkan energi dari aliran angin diperlukan mesin turbin. Salah satu komponen dalam turbin adalah bearing yang terdapat pada bagian sudu (blade). Bearing atau bantalan merupakan komponen penting dalam suatu mesin berputar seperti pada turbin. Kegagalan pada bantalan adalah salah satu jenis kegagalan yang paling sering dijumpai. Kegagalan ini dapat menyebabkan malfungsi atau kerusakan lainnya pada sistem mesin yang berakibat pada terjadinya downtime yang lama dan maintenance cost yang tinggi. Hal inilah yang mendorong dilakukannya proses prognostik untuk memprediksi potensi terjadinya kegagalan dan menentukan estimasi dari Remaining Useful Life (RUL) komponen untuk beroperasi sebelum terjadi kegagalan. Prediksi RUL dilakukan menggunakan data getaran bantalan pada poros kecepatan tinggi turbin angin, kemudian dianalisis secara statistik dan dengan Spectral Kurtosis untuk diperoleh fitur. Fitur tersebut kemudian dipilih yang terbaik dan dilakukan sejumlah proses untuk digunakan sebagai input untuk proses machine learning dengan algoritma Support Vector Regression (SVR). Hasil penelitian menghasilkan 20 variasi model dengan 70% model dengan akurasi prediksi yang beragam namun mampu memberikan estimasi kapan bantalan akan mengalami kegagalan yang diukur dari waktu awal terjadinya kerusakan.

Kata kunci: bantalan; prognostik; spectral kurtosis; statistik; SVR

Abstract

Energy production from renewable sources, such as from air, air and other sources is increasing. Wind power is one of the most developed renewable energies worldwide. To produce energy from wind flow a turbine engine is needed. One of the components in the turbine is the bearing on the blade. Bearings are an important component in a rotating machine such as a turbine. Bearing failure is one of the most common types of failure. This failure can cause malfunctions or other damage to the machine system resulting in long downtime and high maintenance costs. This is what causes a prognostic process to predict potential failures and determine estimates of Remaining Useful Life (RUL) components to operate before failure. The RUL prediction was carried out using bearing vibration data at high speed wind turbines, then statistically analyzed and by Spectral Kurtosis to obtain these features. The best feature is then selected and carried out a number of processes to be used as input for the machine learning process with the Support Vector Regression (SVR) algorithm. The results of the study resulted in 20 variations of the model with 70% of the model with varying prediction accuracy but able to provide an estimate of the time it will experience failure which is measured from the beginning of the damage.

Keywords: bearing; prognostic; spectral kurtosis; statistics; SVR

1. Pendahuluan

Produksi energi dari sumber terbarukan menjadi semakin penting secara global untuk memenuhi permintaan akan energi yang terus meningkat dan mengurangi dampak produksi energi terhadap lingkungan, seperti air dan udara, sistem ekologi dan perubahan iklim [1]. Tenaga angin adalah salah satu teknologi energi terbarukan yang paling berkembang di seluruh dunia. Turbin adalah mesin yang mengubah energi rotasi dari fluida yang diambil oleh sistem rotor menjadi kerja atau energi yang dapat digunakan. Salah satu komponen dalam turbin adalah *bearing* yang terdapat pada bagian sudu (*blade*). *Bearing* atau *rolling element* merupakan komponen penting dalam suatu mesin berputar (*rotary machine*) seperti

dalam turbin, penggerak motor, dan kendaraan [2]. Kegagalan pada *bearing* dalam suatu mesin merupakan salah satu jenis kegagalan yang paling sering dijumpai. Salah satunya pada kegagalan pada *asynchronous motor* dimana 45-55% disebabkan oleh adanya kegagalan pada komponen bearing [3]. Kegagalan pada *bearing* diantaranya terjadi pada bagian *outer-race*, *inner-race*, *roller*, dan *cage*. Kegagalan pada bearing dapat menyebabkan malfungsi atau kerusakan lainnya pada sistem mesin yang berakibat pada terjadinya *downtime* yang lama dan *maintenance cost* yang tinggi [4]. Hal inilah yang mendorong dilakukannya proses prognostik untuk memprediksi potensi terjadinya kegagalan dan menentukan estimasi dari sisa masa efektif komponen untuk beroperasi sebelum terjadi kegagalan (*remaining useful life*). Prognostik sendiri merupakan bagian dari disiplin ilmu baru yaitu *Prognostics Health Management* (PHM) yang juga menjadi strategi baru dalam bidang *maintenance* mesin maupun komponen mesin [5].

Dalam melakukan analisis Prognostik, setidaknya ada tiga aspek utama yang menjadi perhatian yaitu penyusunan atau penentuan indikator kesehatan dari komponen, prediksi sisa masa manfaat (*remaining useful life*) dan manajemen kesehatan dari mesin atau komponen [6]. Penyusunan atau penentuan indikator kesehatan (*features*) dari komponen bertujuan untuk mengevaluasi kondisi saat ini dari sistem mesin dan komponen kritisnya, yang kemudian digunakan untuk sebagai dasar dalam memprediksi *remaining useful life* (RUL) [7]. Fitur yang sesuai dapat menyederhanakan tahap pelacakan dan meningkatkan efektivitas prediksi [8]. Berdasarkan dua aspek pertama, disusun manajemen kesehatan yang optimal direncanakan untuk meminimalkan biaya dan mencegah kecelakaan yang tidak terduga. Penyusunan atau penentuan indikator kesehatan adalah kunci prediksi RUL karena memberikan indikator kesehatan untuk prediksi [3].

Data merupakan komponen yang penting dalam melakukan analisis prognostic khususnya pada *bearing* karena menjadi sumber dalam mengidentifikasi kondisi komponen serta melakukan analisis. Data-data tersebut diperoleh dari sensor yang dipasang pada alat. Terdapat beberapa jenis sensor yang digunakan diantaranya sensor getaran, suhu, kimia, emisi akustik, dan tekanan suara. Data dari sensor tersebut perlu diolah terlebih dahulu sebelum digunakan untuk analisis seperti proses *filtering*, penghilangan *noise*, dll [9]. Data-data dari sensor tersebut dipilih yang paling dapat mengindikasikan terjadinya *fault* pada *bearing*. Proses ini disebut dengan *feature selection*.

Langkah penting dalam analisis prognostic yang efektif adalah memilih algoritme yang paling tepat, berdasarkan properti sistem, karakteristik data, dan penerapan algoritme yang sesuai, dan akhirnya memiliki hasil yang optimal sesuai dengan algoritme yang dipilih [10]. Untuk menentukan algoritme yang optimal, maka diperlukan kriteria pemilihan yang terukur dan juga sistem penilaian (*benchmarking*) yang akurat. Dalam penelitian ini, beberapa pendekatan digunakan pada data yang sama untuk menentukan hasil analisis yang paling optimal. Hasil dari tiap pendekatan akan dibandingkan dan akan dipilih dengan nilai akurasi yang paling tinggi. Dalam proses analisis, data yang telah melalui *preprocessing* akan dibagi menjadi dua dengan rasio tertentu dimana satu bagian digunakan sebagai *train data* yaitu digunakan untuk menyusun *model* sesuai dengan algoritme yang dipilih dan *test data* untuk mengukur keberhasilan dari *model* yang telah disusun terhadap data baru. Antara *train data* dan *test data* tidak ada data yang sama.

Analisis prognostik dalam *Prognostics Health Management* (PHM) dapat meningkatkan keandalan dan ketersediaan dari sistem mesin atau komponen sekaligus mengurangi kesalahan yang tidak terjadwal dan biaya pemeliharaan sistem salah satunya pada turbin [5]. Hal ini yang mendorong mulai banyaknya aplikasi PHM ini dalam dunia industri. Perkembangan berbagai alat, algoritme, dan teknologi juga menjadi keuntungan bagi PHM ini sehingga dapat menghasilkan analisis dan prediksi yang lebih akurat.

Adapun tujuan dari tugas akhir ini adalah sebagai berikut:

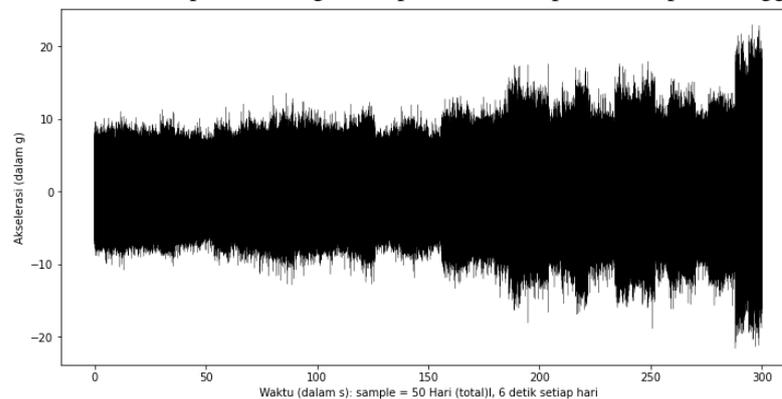
1. Menganalisa parameter yang sesuai sebagai *health indicator* dari *bearing* pada turbin.
2. Menganalisa pendekatan algoritme *machine learning Support Vector Regression* (SVR) dalam melakukan analisis prognostik yaitu memprediksi estimasi sisa masa manfaat (*remaing useful life*) dari bantalan pada poros kecepatan tinggi yang terdapat di turbin angin.
3. Mengetahui pengaruh perbedaan metode pengolahan sinyal dari fitur yang diekstraksi dari data getaran bantalan poros kecepatan turbin angin terhadap performa model SVR yang dihasilkan dalam memprediksi *Remaining Usefu Life* (RUL) bantalan pada poros kecepatan tinggi turbin angin.
4. Mengevaluasi performa model yang dihasilkan dengan pendekatan algoritme *Support Vector Regression* (SVR) yang digunakan dalam membuat prediksi pada *Remaining Usefu Life* (RUL) bantalan pada poros kecepatan tinggi turbin angin serta parameter model yang mempengaruhi performa dalam melakukan prediksi.

2. Bahan dan Metode Penelitian

2.1. Pengumpulan Data Getaran Bantalan pada Turbin Angin

Data yang digunakan untuk memprediksi nilai *Remaining Useful Life* (RUL) dari bantalan turbin angin, diperoleh secara daring yang dipublikasikan oleh Mathworks. Data ini diambil dari turbin angin berkapasitas 2 MW pada bantalan

yang terpasang di poros kecepatan tinggi (*high-speed shaft*). Data diambil berupa waktu pengambilan data, getaran bantalan, dan kecepatan putar bantalan (RPM), setiap hari selama 50 hari dengan pengambilan sampel data selama 6 detik. Pada akhir pengambilan data, bantalan mengalami kegagalan pada bagian *inner race*. Pada analisis ini hanya digunakan data getaran. Gambar 1 merupakan data getaran pada bantalan poros kecepatan tinggi turbin angin



Gambar 1. Grafik Getaran Bantalan pada Poros Kecepatan Tinggi Turbin Angin

2.2. Spectral Kurtosis dari Data Getaran Bantalan Turbin Angin

Sinyal getaran yang ditemui dalam sistem mesin putar, seperti pada turbin khususnya pada bagian bantalan dapat secara luas diklasifikasikan sebagai stasioner atau nonstasioner [11]. Sinyal getaran bantalan pada bantalan poros kecepatan tinggi pada turbin dianalisis sebagai sinyal stationer dan non-stationer. Untuk analisis getaran sebagai sinyal non-stationer dilakukan dengan pendekatan analisis *time-frequency* dengan metode *Spectral Kurtosis*. Metode analisis Spectral Kurtosis merupakan teknik analisis *time-frequency* yang menggabungkan proses transformasi sinyal dengan menggunakan *Short-Time Fourier Transform* (STFT) dan perhitungan nilai *Kurtosis* untuk tiap nilai frekuensi [12].

Pada tahap STFT, ditentukan panjang *window* yang optimal dikarenakan mempengaruhi resolusi nilai *Spectral Kurtosis* yang dihasilkan dari data getaran. Pada analisis ini, digunakan variasi panjang *window* sebesar 16, 32, 64, 128, 256, dan 512.

2.3. Ekstraksi Fitur dari Sinyal Getaran dan Spektral Kurtosis

Ekstraksi fitur dilakukan dengan metode *time domain* berdasarkan perhitungan statistik pada data getaran bantalan secara langsung dan hasil *Spectral Kurtosis* dan diperoleh tren dari spektrum *Spectral Kurtosis* sehingga dapat dianalisis pola yang mampu mendeteksi adanya cacat bantalan yang baru muncul [13]. Pada perhitungan statistik secara langsung pada data getaran bantalan turbin angin digunakan 12 perhitungan statistik yang terangkum dalam Tabel 1 sedangkan untuk perhitungan statistik pada nilai *Spectral Kurtosis* digunakan 11 perhitungan statistik yang juga terangkum pada Tabel 1 tanpa melakukan perhitungan nilai *entropy*. Data bantalan turbin angin yang dianalisis secara *time domain* dan *time-frequency domain* dihasilkan sebanyak 23 fitur yang digunakan sebagai dasar analisis prognosis.

Tabel 1. Daftar Fitur yang Diekstrak

Fitur	Rumus Perhitungan
Rata-Rata Absolut	$\bar{x} = \frac{1}{N} \sum_{i=1}^N x_i $
Standar Deviasi (STD)	$\sigma = \sqrt{\frac{1}{N} \sum_{i=1}^N (x_i - \bar{x})^2}$
Skewness	$SK = \frac{1}{N} \sum_{i=1}^N \frac{(x_i - \bar{x})^3}{\sigma^3}$
Kurtosis	$K = \frac{1}{N} \sum_{i=1}^N \frac{(x_i - \bar{x})^4}{\sigma^4}$
Entropi	$H(X) = - \sum_{i=1}^N P(x_i) \log P(x_i)$
Root Mean Square (RMS)	$x_{rms} = \sqrt{\frac{1}{N} \sum_{i=1}^N x^2}$

Fitur	Rumus Perhitungan
Maksimum Absolut	$x_{max} = \max value $
Peak to Peak (P2P)	$x_p = \max value - \min value$
Faktor Puncak	$x_{crest} = \frac{\max value}{x_{rms}}$
Clearance Factor	$x_{clear} = \frac{x_p}{\left(\frac{1}{N} \sum_{i=1}^N \sqrt{ x_i }\right)^2}$
Shape Factor	$Shape\ factor = \frac{x_{rms}}{\bar{x}}$
Impuls	$Impulse = \frac{\max value}{\bar{x}}$

2.4. Pemilihan Fitur dari Data Getaran Bantalan

Pemilihan dilakukan dengan perhitungan metrik menggunakan tiga komponen yaitu *correlation* (Corr), *monotonicity* (Mon), dan *robustness* (Rob). Untuk menghitung metrik tersebut, fitur didekomposisi menjadi tren rata-rata dan bagian acak menggunakan metode pemulusan (*smoothing*). Dengan persamaan $X(t_k) = X_T(t_k) + X_R(t_k)$. Rumus perhitungan untuk ketiga metrik tersebut terangkum pada Tabel 2. Ketiga metrik tersebut digunakan untuk mengukur nilai kebaikan fitur (J) dalam merepresentasikan terjadinya degradasi pada bantalan yang terpasang pada poros kecepatan tinggi turbin angin. Nilai kebaikan dihitung dengan persamaan $\max_{x=\Omega} J = \omega_1 Corr(X) + \omega_2 Mon(X) + \omega_3 Rob(X)$ dimana ω adalah berat signifikansi dari masing-masing metrik dalam perhitungan nilai kebaikan fitur. Dalam analisis ini akan digunakan nilai $\omega_1 = 0.2$ untuk *correlation*, $\omega_2 = 0.5$ untuk *monotonicity*, dan $\omega_3 = 0.3$ untuk *robustness* [14].

Tabel 2. Rumus Perhitungan Metrik Kebaikan Fitur

Fitur	Rumus Perhitungan
<i>Correlation</i> (Corr)	$Corr(X, T) = \frac{ K \sum_k X_T(t_k) t_k - \sum_k X_T t_k \sum_k t_k }{\sqrt{[K \sum_k X_T(t_k)^2 - (\sum_k X_T(t_k))^2][K \sum_k t_k^2 - (\sum_k t_k)^2]}}$
<i>Monotonicity</i> (Mon)	$Mon(X) = \frac{1}{K-1} \left \sum_k \delta(X_T(t_{k+1}) - X_T(t_k)) - \sum_k \delta(X_T(t_k) - X_T(t_{k+1})) \right $
<i>Robustness</i> (Rob)	$Rob(X) = \frac{1}{K} \sum_k \exp\left(-\left \frac{X_R(t_k)}{X(t_k)}\right \right)$

2.5. Proses Smoothing pada Fitur Terbaik yang Dipilih

Fitur terbaik yang telah dipilih dilakukan proses *smoothing* atau penghalusan dikarenakan nilai pada fitur-fitur tersebut memiliki *noise* tinggi. *Noise* yang tinggi pada fitur menyebabkan tren pada fitur tidak terlihat secara jelas yang bisa mempengaruhi hasil dari prediksi nilai RUL dari bantalan. Proses *smoothing* yang dilakukan bertujuan untuk mengurangi *noise* pada fitur sehingga tren yang dibentuk oleh data pada masing-masing fitur dapat terlihat dengan jelas. Pada analisis ini digunakan empat algoritma *smoothing* yaitu *simple moving average*, *exponential moving average*, *simple exponential smoothing*, dan *Savitzky-Golay filter* dengan parameter masing-masing.

2.6. Pembagian Data untuk Train dan Test

Data yang digunakan setelah proses *smoothing*, dibagi menjadi dua bagian yaitu untuk proses *train* model dan *test*. Untuk *training* model, digunakan 40% dari keseluruhan total data, mulai dari data hari ke 1 hingga hari ke 20. Sedangkan untuk proses *test* dari model, digunakan 60% dari keseluruhan data yaitu mulai dari hari ke 21 hingga hari ke 50.

2.7. Standarisasi Nilai Fitur Terbaik yang Dipilih

Proses standarisasi dan penyesuaian skala dilakukan dengan fungsi *StandardScaler* dari *library preprocessing* Scikit-learn. Proses ini dilakukan karena algoritma *machine learning* yang mampu bekerja lebih baik ketika fitur berada pada skala yang relatif sama dan mendekati distribusi normal. Standarisasi dengan *StandardScaler* dilakukan pada data *train* untuk memperoleh koefisien. Koefisien tersebut digunakan untuk melakukan penyesuaian skala dan standarisasi pada fitur yang dipilih baik untuk data *train* dan *test*.

2.8. Pembuatan Fitur dengan Principal Component Analysis (PCA)

Pada pembuatan model untuk prediksi nilai RUL pada bantalan poros kecepatan tinggi turbin angin, terdapat variasi yang menggunakan input komponen hasil *Principal Component Analysis* (PCA). PCA adalah teknik statistik non-parametrik yang bersifat *unsupervised* dan digunakan untuk pengurangan dimensi dalam pembelajaran mesin. Proses ini dilakukan dengan perintah PCA dari *library preprocessing* Scikit-learn. Dari proses ini dihasilkan dua komponen PCA.

2.9. Prediksi Nilai RUL dengan Metode Support Vector Regression (SVR)

Untuk melakukan prediksi terhadap nilai RUL dari bantalan turbin angin, dilakukan pembuatan model dengan metode *Support Vector Regression* (SVR). Pada analisis ini akan dilakukan 20 variasi pembuatan model sesuai dengan fitur yang digunakan sebagai input. Model yang dibentuk dibagi menjadi dua jenis yaitu model dasar dengan kernel linier dan model dengan parameter terbaik. Untuk mencari parameter terbaik dari algoritma SVR, digunakan *library* Optuna. Dalam analisis ini parameter dari model SVR yang akan dioptimasi terangkum pada Tabel 3

Tabel 3. Parameter yang Dioptimasi pada Model SVR untuk Prediksi Nilai RUL Bantalan pada Poros Kecepatan Tinggi Turbin Angin

Parameter	Keterangan	Nilai
Kernel	Jenis kernel yang digunakan pada model untuk mempelajari data <i>train</i> dan melakukan prediksi	linear, poly, rbf, sigmoid
Degree	Derajat polynomial pada kernel poly, pada kernel lain nilainya diabaikan	1, 2, 3
Gamma Coef0	Nilai koefisien untuk kernel jenis rbf, poly, dan sigmoid Fungsi independen dalam kernel yang hanya signifikan pada kernel poly dan sigmoid	scale, auto 0.01 hingga 0.4, dengan langkah 0.01
Epsilon	Epsilon dalam model epsilon-SVR. Nilai Ini menentukan tabung epsilon di mana tidak ada penalti yang terkait fungsi kerugian pada proses <i>training</i> dengan poin yang diprediksi dalam jarak epsilon dari nilai aktual.	0.01 hingga 0.4, dengan langkah 0.01

Untuk menilai performa dari tiap model dalam melakukan prediksi nilai RUL, digunakan perhitungan skor dengan menggunakan *Root Mean Square Error* (RMSE) dan Koefisien Determinasi (R^2). Nilai RMSE sendiri mencerminkan performa model dalam melakukan prediksi dimana semakin kecil nilai RMSE maka semakin baik performa dari model. Tabel 4 merangkum rumus perhitungan untuk RMSE dan R^2

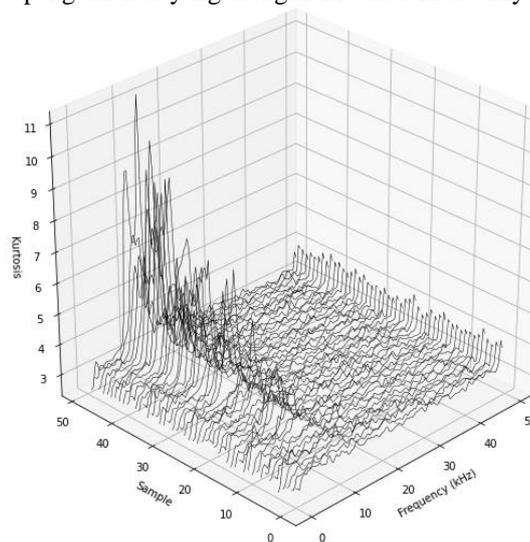
Tabel 4. Rumus Perhitungan *Root Mean Square Error* (RMSE) dan Koefisien Determinasi (R^2)

Metrik	Rumus Perhitungan
<i>Root Mean Square Error</i> (RMSE)	$RMSE = \sqrt{\frac{1}{N} \sum_{i=1}^N (Prediksi_i - Aktual_i)^2}$
Koefisien Determinasi (R^2)	$R^2 = \left(\frac{N(\sum xy) - (\sum x)(\sum y)}{\sqrt{[N\sum x^2 - (\sum x)^2][N\sum y^2 - (\sum y)^2]}} \right)^2$

3. Hasil dan Pembahasan

3.1. Nilai *Spectral Kurtosis* dari Getaran Bantalan pada Poros Kecepatan Tinggi Turbin Angin

Hasil dari perhitungan nilai *Spectral Kurtosis* dari data getaran bantalan poros kecepatan tinggi turbin angin ditampilkan dalam bentuk grafik tiga dimensi pada Gambar 2. Dari gambar tersebut, terlihat bahwa nilai SK mengalami tren kenaikan mulai dari hari ke 20 pengambilan yang mengindikasikan munculnya kerusakan pada bantalan.



Gambar 2. Hasil *Spectral Kurtosis* dari Data Getaran Bantalan Turbin Angin

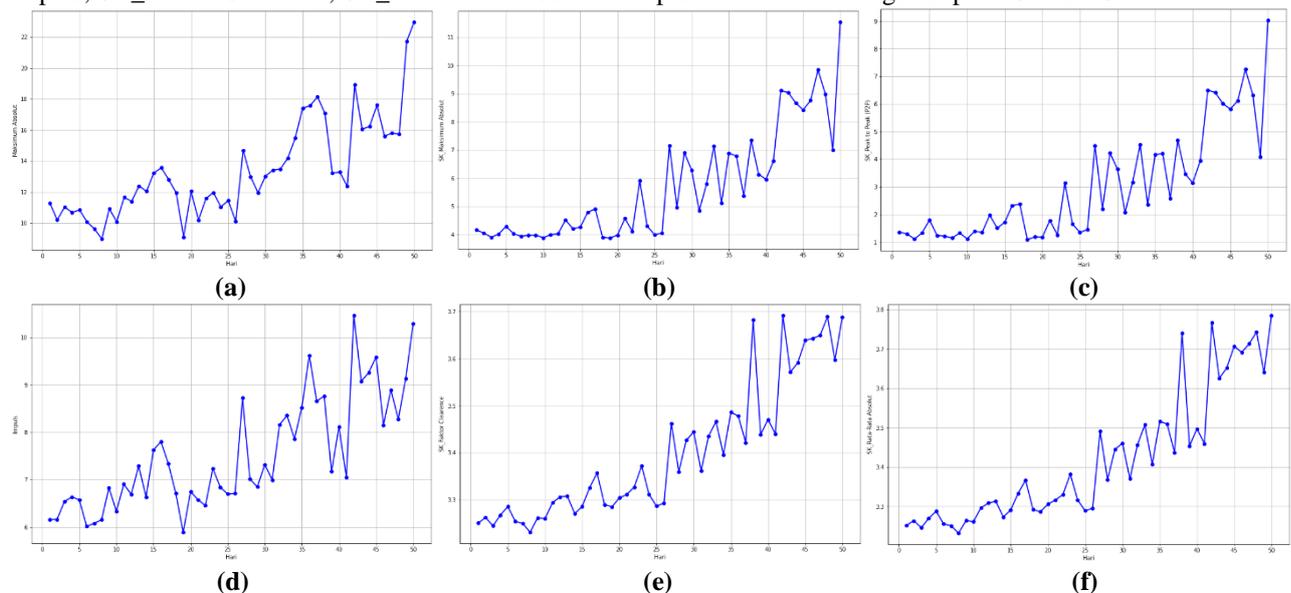
3.2. Pemilihan Fitur Terbaik dari Hasil Ekstraksi Fitur pada Data Getaran Bantalan Secara Langsung dan dari Hasil *Spectral Kurtosis*

Fitur yang telah diekstraksi dengan perhitungan statistik terhadap data getaran bantalan poros kecepatan tinggi secara langsung dan hasil *Spectral Kurtosis* yaitu sebanyak 23 fitur, dihitung nilai kebaikan fitur sebagai indikator kesehatan yang mampu menggambarkan kondisi kerusakan bantalan. Dari hasil perhitungan tersebut, dipilih enam fitur dengan nilai *goodness metric* atau nilai kebaikan fitur tertinggi. Hasilnya ditampilkan pada Tabel 5.

Tabel 5. Hasil Perhitungan Nilai Kebaikan Fitur Sebagai Indikator Kesehatan

Fitur	Correlation	Monotonicity	Robustness	Goodness Metric
Maksimum Absolut	0.7952	0.2857	0.9400	0.5839
SK_Maksimum Absolut	0.8495	0.1494	0.9220	0.5212
SK_Peak to Peak (P2P)	0.8412	0.1553	0.8265	0.4939
Impuls	0.7995	0.0933	0.9501	0.4916
SK_Faktor Clearance	0.9005	0.0095	0.9903	0.4819
SK_Rata-Rata Absolut	0.8929	0.0113	0.9890	0.4809

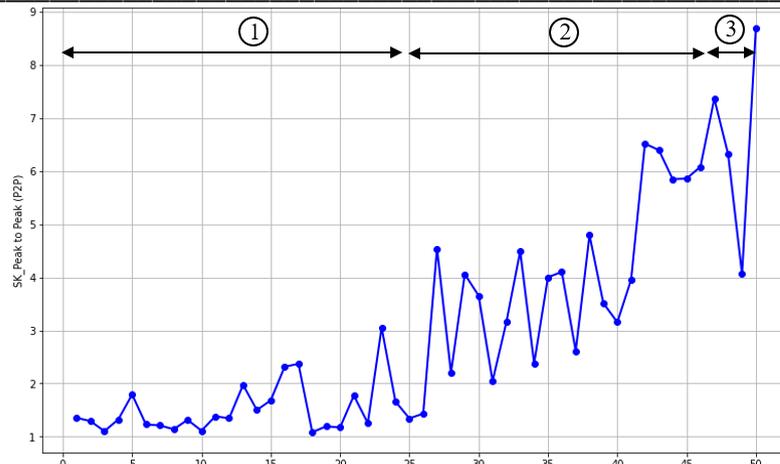
Nilai dari keenam fitur yang dipilih yaitu Maksimum Absolut, SK_Maksimum Absolut, SK_Peak to Peak (P2P), Impuls, SK_Faktor Clearance, SK_Rata-Rata Absolut ditampilkan dalam bentuk grafik pada Gambar 3.



Gambar 3. Fitur Terbaik dari Hasil Ekstraksi Fitur pada Data Getaran Bantalan dan *Spectral Kurtosis* (SK)
 (a) Maksimum Absolut, (b) SK_Maksimum Absolut, (c) SK_Peak to Peak, (d) Impuls, (e) SK_Faktor Clearance,
 (f) SK_Rata-rata Absolut

3.3. Hasil Fitur dari Proses PCA dari Fitur Terbaik yang Dipilih

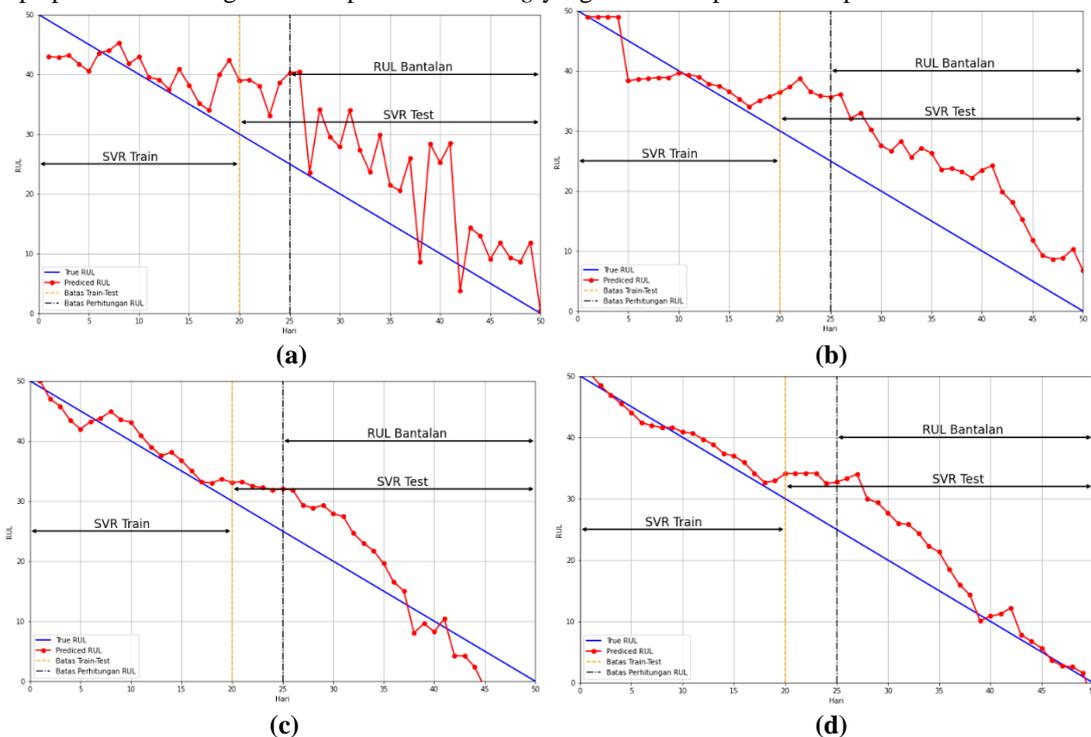
Fitur-fitur yang diekstrak dari data getaran bantalan secara langsung dengan domain waktu, dinilai tidak dapat menggambarkan kondisi bantalan secara baik akibat adanya *noise* pada data getaran. Hal ini ditunjukkan oleh nilai fitur dengan osilasi tinggi sehingga degradasi awal pada bantalan tidak terdeteksi dengan baik. *Noise* yang tinggi pada fitur menyebabkan sulitnya pendeteksian awal terjadi degradasi atau kerusakan pada bantalan. Oleh karena itu, diperlukan fitur yang lebih sensitif dalam menggambarkan perkembangan kegagalan. Penggunaan metode *Spectral Kurtosis* ditujukan untuk menghasilkan fitur dengan *noise* rendah sehingga bisa diperoleh fitur yang mampu mendeteksi awal terjadinya kerusakan pada bantalan sehingga fitur terbaik yang dipilih dapat digunakan sebagai indikator kesehatan yang menggambarkan kondisi dari bantalan poros kecepatan tinggi pada turbin angin. Salah satu fitur yang dipilih yaitu *Peak to Peak* dari nilai *Spectral Kurtosis* yang ditampilkan pada Gambar 8. Grafik tersebut menunjukkan fluktuasi tertinggi juga terjadi pada hari ke-25 dan nilai fitur terkecil antara hari ke-1 dan ke-22. Berdasarkan hal tersebut dapat ditetapkan tiga zona dari kondisi bantalan yaitu: zona normal (1), zona degradasi (2), dan zona runtuh (3) yang memiliki nilai tertinggi. Perhitungan RUL dari bantalan dihitung sejak terjadinya degradasi atau kerusakan awal yaitu pada hari ke-25. Oleh karena itu nilai aktual dari RUL yang dimiliki oleh bantalan bernilai 25 hari karena pada data faktual disebutkan bahwa pada akhir pengambilan data terjadi kegagalan pada inner race bantalan. Data train untuk pembuatan model adalah data bantalan poros kecepatan tinggi turbin angin yang bebas dari adanya kegagalan. Pada Gambar 4, pengukuran dengan nilai fitur terkecil ditemukan untuk pengukuran antara hari ke-1 dan ke-22 sehingga pada area ini bisa ditetapkan bahwa kondisi bantalan normal. Namun untuk analisis ini akan digunakan data 20 hari pertama.

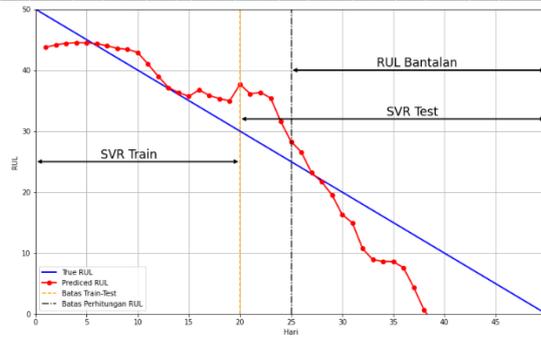


Gambar 4. Nilai *Peak To Peak* dari Hasil Perhitungan *Spectral Kurtosis* Data Getaran Bantalan pada Poros Kecepatan Tinggi Turbin Angin

3.4. Hasil Pembuatan Model SVR untuk Prediksi RUL Bantalan Poros Kecepatan Tinggi Turbin Angin

Model *Support Vector Regression* (SVR) yang dihasilkan dapat dibagi menjadi dua bagian berdasarkan jenis inputnya yaitu enam fitur terbaik dan satu komponen hasil PCA. Pada masing-masing jenis input, terdapat dua variasi model yang tidak dilakukan proses *smoothing* dan delapan variasi yang dilakukan proses *smoothing* dengan empat metode berbeda. Model SVR yang dibentuk menggunakan model dasar berupa kernel linier dan parameter terbaik. Masing-masing variasi model dilakukan proses standarisasi dan penyesuaian skala dengan *StandardScaler* pada input yang digunakan. Pengukuran performa dari model dilakukan dengan perhitungan RMSE dan koefisien determinasi (R^2). Hasil prediksi RUL oleh model secara garis besar dapat menjadi dua yaitu model yang mampu memprediksi kapan terjadinya kerusakan pada bantalan karena pada prediksi menghasilkan $RUL=0$ dan tidak mampu melakukan prediksi karena $RUL \neq 0$. Pada model dengan menggunakan parameter terbaik memiliki performa yang lebih baik dibanding model dasar dengan kernel linier. Penggunaan parameter terbaik juga mampu mengatasi masalah *overfitting* pada model. Pada model yang tidak dilakukan proses *smoothing* memiliki fluktuasi yang tinggi dibandingkan dengan model yang dihasilkan menggunakan input fitur dengan proses *smoothing*. Masing-masing model yang terbentuk menghasilkan nilai RUL yang beragam namun, sebagian besar menghasilkan prediksi RUL dengan nilai lebih rendah dibanding nilai aktualnya. Gambar 5 menampilkan hasil prediksi dari model dengan parameter algoritma SVR terbaik dan input menggunakan enam fitur terbaik tanpa proses PCA dengan variasi proses *smoothing* yang dilakukan pada fitur input.

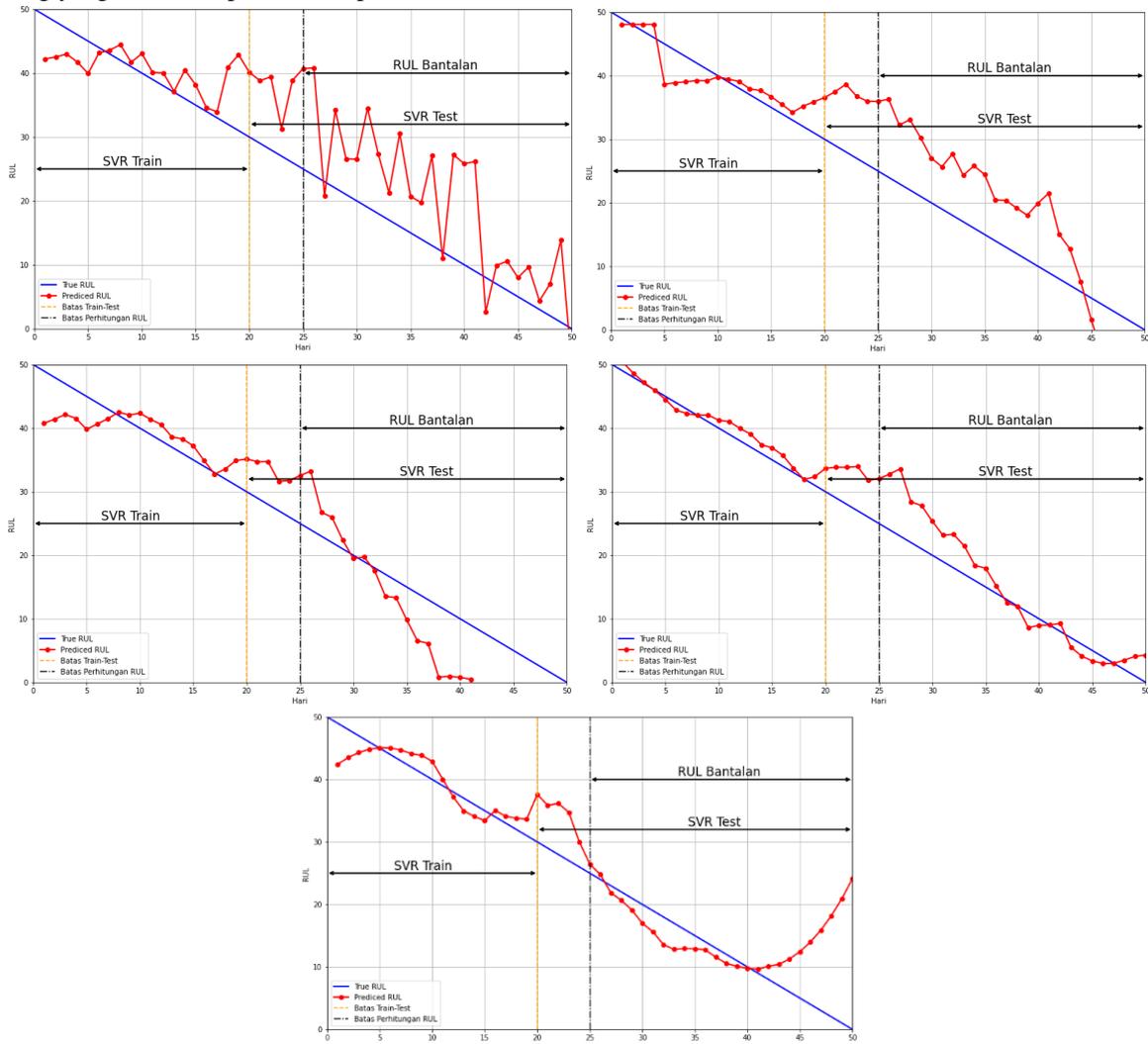




(e)

Gambar 5. Hasil Prediksi Nilai RUL dengan Input Enam Fitur Terbaik Menggunakan Parameter Terbaik dan *Smoothing* pada Fitur Menggunakan Metode (a) Tanpa Proses *Smoothing*, (b) *Simple Moving Average*, (c) *Exponential Moving Average*, (d) *SimpleExponential Smoothing*, dan (e) *Savitzky-Golay Filter*

Gambar 6 menampilkan hasil prediksi dari model dengan parameter algoritma SVR terbaik dan input menggunakan satu kompone PCA yang diperoleh dari enam fitur terbaik tanpa proses PCA dengan variasi proses *smoothing* yang dilakukan pada fitur input.



(a)

(b)

Gambar 6. Hasil Prediksi Nilai RUL dengan Input Satu Komponen PCA dari Enam Fitur Terbaik Menggunakan Parameter Terbaik dan *Smoothing* pada Fitur Menggunakan Metode (a) Tanpa Proses *Smoothing*, (b) *Simple Moving Average*, (c) *Exponential Moving Average*, (d) *SimpleExponential Smoothing*, dan (e) *Savitzky-Golay Filter*

3.5. Hasil Pembuatan Model SVR untuk Prediksi RUL Bantalan Poros Kecepatan Tinggi Turbin Angin

Dari 20 variasi model SVR dengan beberapa input dan variasi parameter yang digunakan untuk membentuk model tersebut yang digunakan untuk memprediksi nilai RUL bantalan pada poros kecepatan tinggi turbin angin, dapat dikelompokkan menjadi dua bagian berdasarkan hasil dari prediksinya. Kelompok tersebut adalah model yang dapat memprediksi kapan terjadinya kegagalan total dari bantalan yaitu ketika nilai RUL mencapai nilai 0 dan model yang tidak mampu memprediksi kapan bantalan akan gagal karena dari hasil prediksi model tidak ada yang mencapai nilai $RUL=0$. Untuk model yang mampu memprediksi kapan terjadinya kegagalan total dari bantalan dirangkum dalam Tabel 6 beserta dengan nilai RMSE dan R^2 -nya. Berdasarkan dari Tabel 6 dari 20 variasi model yang dibentuk dengan menggunakan algoritma Support Vector Regression (SVR) menggunakan beberapa konfigurasi parameter dan sejumlah variasi input dengan pemrosesan yang berbeda, 70 % model yang dihasilkan dapat memprediksi waktu sisa (RUL) dari bantalan pada poros kecepatan tinggi turbin angin sebelum mengalami kegagalan sepenuhnya yaitu ketika RUL bernilai nol. Dari model yang dapat melakukan prediksi tersebut lima model mengalami *overfitting*. *Overfitting* mengacu pada model yang memodelkan data pelatihan (train) terlalu baik.

Tabel 6. Hasil Prediksi oleh Variasi Model yang Mampu Menentukan Waktu Terjadinya Kegagalan pada Bantalan

Variasi Model	RMSE Train	RMSE Test	RUL
Model SVR parameter terbaik dengan input enam fitur tanpa proses <i>Smoothing</i>	4.5785	9.6796	25
Model SVR kernel linier dengan input enam fitur dengan proses <i>smoothing Simple Exponential</i>	1.7344	6.2716	19
Model SVR parameter terbaik dengan input enam fitur dengan proses <i>smoothing Simple Exponential</i>	1.7604	5.3712	19
Model SVR kernel linier dengan input enam fitur dengan proses <i>smoothing Holt</i>	1.0032	6.2716	18
Model SVR parameter terbaik dengan input enam fitur tanpa proses <i>Smoothing</i>	4.5785	9.6796	25
Model SVR kernel linier dengan input enam fitur dengan proses <i>smoothing Simple Exponential</i>	1.7344	6.2716	19
Model SVR parameter terbaik dengan input enam fitur dengan proses <i>smoothing Simple Exponential</i>	1.7604	5.3712	19
Model SVR kernel linier dengan input enam fitur dengan proses <i>smoothing Holt</i>	1.0032	6.2716	18
Model SVR parameter terbaik dengan input enam fitur tanpa proses <i>Smoothing</i>	4.5785	9.6796	25
Model SVR kernel linier dengan input enam fitur dengan proses <i>smoothing Simple Exponential</i>	1.7344	6.2716	19
Model SVR parameter terbaik dengan input enam fitur dengan proses <i>smoothing Simple Exponential</i>	1.7604	5.3712	19
Model SVR kernel linier dengan input enam fitur dengan proses <i>smoothing Holt</i>	1.0032	6.2716	18
Model SVR kernel linier dengan input satu PCA dari enam fitur dengan proses <i>smoothing Holt</i>	0.9617	9.1203	18
Model SVR kernel linier dengan input satu PCA dari enam fitur dengan proses <i>smoothing Savitzky-Golay Filter</i>	3.0933	29.3766	8

4. Kesimpulan

Kesimpulan dari penelitian ini adalah sebagai berikut:

1. Dengan melakukan perhitungan metrik yang terdiri dari tiga komponen yaitu *correlation*, *monotonicity*, dan *robustness* dengan nilai berat atau signifikansi masing-masing, dapat ditentukan dan dipilih fitur yang memiliki pengaruh paling besar dalam membentuk model untuk memprediksi nilai *Remaining Useful Life* (RUL). Fitur yang dipilih berperan sebagai indikator kesehatan pada bantalan dimana menunjukkan kondisi bantalan dalam beroperasi.
2. Melalui data getaran pada bantalan yang terpasang pada poros kecepatan tinggi pada turbin angin, berhasil dibuat beberapa variasi model dengan menggunakan algoritma *Support Vector Regression* (SVR) yang mampu melakukan prediksi terhadap *Remaining Useful Life* (RUL) dari bantalan sebelum mengalami kegagalan yang diukur dari awal terjadinya degradasi.
3. Dengan menggunakan data getaran bantalan, dapat dilakukan sejumlah metode pengolahan sinyal yang bertujuan untuk mengekstraksi fitur dari data tersebut. Fitur ini merupakan nilai-nilai yang mampu menggambarkan kondisi dari bantalan mulai dari kondisi normal, mengalami degradasi hingga akhirnya terjadi

kegagalan. Fitur ini digunakan sebagai input dalam membuat model SVR yang memprediksi nilai RUL bantalan. Metode pengolahan sinyal yang dipilih untuk mengekstraksi fitur dari data getaran bantalan akan mempengaruhi hasil prediksi dan performa model SVR yang terbentuk sehingga perlu dipilih metode yang sesuai dengan karakteristik sinyal getaran yang dihasilkan oleh bantalan.

4. Dalam membentuk model untuk melakukan prediksi dengan menggunakan algoritma SVR, model yang terbentuk dapat melakukan prediksi dengan baik namun perlu dimaksimalkan seluruh parameter yang disediakan tersebut dan dicari parameter paling optimal sehingga diperoleh model dengan performa yang paling baik yang dalam hal adalah model untuk melakukan prediksi dari nilai RUL bantalan pada poros kecepatan tinggi yang ada pada turbin angin. Model dengan parameter yang sudah dimaksimalkan dapat menghasilkan prediksi yang lebih akurat. Pada algoritma SVR, terdapat parameter yang bisa disesuaikan mengikuti karakteristik data yang digunakan sehingga model yang terbentuk adalah model yang paling optimal untuk jenis data input.

5. Daftar Pustaka

- [1] Zhang Z. *Automatic Fault Prediction of Wind Turbine Main Bearing Based on SCADA Data and Artificial Neural Network*. Open Journal of Applied Sciences. 2018;08(06):211–25.
- [2] Jin X, Que Z, Sun Y, Guo Y, Qiao W. *A Data-Driven Approach for Bearing Fault Prognostics*. IEEE Trans Ind Appl. 2019 Jul 1;55(4):3394–401.
- [3] Liu X, Song P, Yang C, Hao C, Peng W. *Prognostics and Health Management of Bearings Based on Logarithmic Linear Recursive Least-Squares and Recursive Maximum Likelihood Estimation*. IEEE Transactions on Industrial Electronics. 2017 Jul 29;65(2):1549–58.
- [4] Wang Y, Xu G, Liang L, Jiang K. *Detection of Weak Transient Signals Based on Wavelet Packet Transform and Manifold Learning for Rolling Element Bearing Fault Diagnosis*. Mech Syst Signal Process. 2015 Mar 1;54:259–76.
- [5] Abid K, Sayed-Mouchaweh M, Cornez L. *Adaptive Data-driven Approach for Fault Prognostics based on Normal Conditions-Application to Shaft Bearings of Wind Turbine*. 2019.
- [6] Wang D, Tsui KL, Miao Q. *Prognostics and Health Management: A Review of Vibration Based Bearing and Gear Health Indicators*. IEEE Access. 2017 Nov 19;6:665–76.
- [7] Heng A, Tan ACC, Mathew J, Montgomery N, Banjevic D, Jardine AKS. *Intelligent Condition-Based Prediction of Machinery Reliability*. Mech Syst Signal Process. 2009 Jul;23(5):1600–14.
- [8] Jianming S, Yongxiang L, Gong W, Xuzhi L. *Health Index Synthetization and Remaining Useful Life Estimation for Turbofan Engines Based on Run-To-Failure Datasets*. Eksploatacja i Niezawodnosc. 2016;18(4):621–31.
- [9] Yan W, Qiu H, Iyer N. *Feature Extraction for Bearing Prognostics and Health Management (PHM)-A Survey (Preprint) Stinfo Copy Air Force Research Laboratory Materials and Manufacturing Directorate Wright-Patterson Air Force Base, Oh 45433-7750 Air Force Materiel Command United States Air Force 1. Report Date (DD-MM-YY) 2. Report Type 3. Dates Covered (From-To) Feature Extraction For Bearing Prognostics And Health Management (Phm)-A Survey (Preprint)*. 2008;
- [10] Lee J, Wu F, Zhao W, Ghaffari M, Liao L, Siegel D. *Prognostics and Health Management Design for Rotary Machinery Systems - Reviews, Methodology and Applications*. Mech Syst Signal Process. 2014 Jan;42(1–2):314–34.
- [11] Yan R, Gao RX. *Hilbert-Huang Transform-Based Vibration Signal Analysis for Machine Health Monitoring*. IEEE Trans Instrum Meas. 2006 Dec;55(6):2320–9.
- [12] Leite VCMN, Borges Da Silva JG, Veloso GFC, Borges Da Silva LE, Lambert-Torres G, Bonaldi EL, et al. *Detection of Localized Bearing Faults in Induction Machines by Spectral Kurtosis and Envelope Analysis of Stator Current*. IEEE Transactions on Industrial Electronics. 2015;62(3):1855–65.
- [13] Sun Q, Chen P, Zhang D, Xi F. *Pattern Recognition for Automatic Machinery Fault Diagnosis*. Journal of Vibration and Acoustics, Transactions of the ASME. 2004 Apr;126(2):307–16.
- [14] Zhang B, Zhang L, Xu J. *Degradation Feature Selection for Remaining Useful Life Prediction of Rolling Element Bearings*. Qual Reliab Eng Int. 2016 Mar 1;32(2):547–54.