



Algoritma Pendeteksi Kerusakan Struktur Bangunan Berbasis Korelasi Jarak dan Metode Kuadrat Terkecil Parsial

Jimmy Tjen^{#1}

[#]Jurusan Informatika Universitas Widya Dharma Pontianak
Jalan Hos Cokroaminoto No.445, Pontianak, Kalimantan Barat, 78243

¹jimmy.tjen@mathmods.eu

Abstrak— Sering kali, sebuah kerusakan struktur yang masif terjadi karena pengabaian terhadap kerusakan kecil. Kejadian malang ini kemudian menimbulkan berbagai kerugian, baik secara material maupun korban jiwa. Oleh karena itu, dirasa penting untuk dapat mendeteksi kerusakan dari sebuah struktur sedini mungkin untuk mencegah terjadinya hal yang tidak diinginkan. Penelitian ini menggagas sebuah algoritma pendeteksi kerusakan struktur bangunan berbasiskan pada metode korelasi jarak dan kuadrat terkecil parsial. algoritma ini berfokus pada pemilihan sekelompok sensor yang dapat bekerja secara optimal berdasarkan pada perhitungan korelasi jarak. Berdasarkan pada percobaan pada data experimental dari sebuah struktur jembatan, algoritma yang digagas dapat mengurangi jumlah akselerometer yang diperlukan hingga 80% untuk menyusun model prediktif tanpa mengurangi atau bahkan meningkatkan akurasi dari model prediktif akselerometer sebesar 1 hingga 1,3%. Lebih lanjut, algoritma yang digagas dapat mendekteksi keberadaan kerusakan struktur dengan baik, serta mampu mengkarakterisasi tingkat kerusakan dari struktur berdasarkan pada perubahan standar deviasi dari residu kuadrat.

Kata kunci— Analisis Mahadata, Korelasi Jarak, Metode Kuadrat Terkecil Parsial, Monitoring Kesehatan Struktur, Pembelajaran Mesin.

Abstract— It is often the case, that a catastrophic structure failure occurred due to the neglect of minor faults. This unfortunate event induces a heavy loss of material and even poses safety risks. Thus, it is important to be able to detect failures in a structure as early as possible, such that this unfortunate event can be avoided. This research proposes a new fault detection algorithm based on distance correlation and the partial least square method. The idea behind this algorithm is to generate a subset of sensors which is able to work optimally based on the calculation of the distance correlation. Based on an experimental dataset taken from a bridge structure, the proposed algorithm can reduce the number of accelerometers needed to generate predictive models by up to 80%, without significantly decreasing or even increasing the model predictive accuracy by 1 up to 1.3%. Furthermore, the suggested algorithm is able to detect the presence of faults, smoothly and potentially characterize the

damage level inside a structure based on the rate of change of the squared residual standard deviation.

Keywords— Big Data Analysis, Distance Correlation, Machine Learning, Structural Health Monitoring, Partial Least Square Method.

I. PENDAHULUAN

Berkembangnya revolusi industri 4.0 telah mendorong perkembangan aktivitas teknologi, seperti penerapan berbasis pengolahan mahadata atau *big data processing* dan pembelajaran mesin atau *machine learning* diberbagai aspek kehidupan [1]. Pada era ini, Lazimnya sebuah struktur buatan manusia akan dilengkapi dengan beberapa jenis perangkat sensor yang mampu mengukur berbagai besaran fisis di dalam struktur. Informasi yang dibawa oleh sensor ini sangatlah penting, sebagai contoh sebuah persamaan matematis berbasis data dapat diturunkan untuk mendeteksi apakah terdapat kerusakan di dalam struktur, yang dapat mendatangkan kerugian di berbagai aspek hingga mengancam keselamatan jiwa orang yang berada di dalamnya. Metode tersebut dalam pembahasan monitoring kesehatan struktural (*structural health and monitoring*) dikenal sebagai metode berbasiskan data atau *data driven approach* [2], [3], [4]. Metode ini sangat efisien untuk digunakan, terutama ketika struktur dari sebuah bangunan terlalu rumit untuk dimodelkan secara penjabaran analitis [5]. Salah satu dari metode yang paling umum digunakan adalah metode Kuadrat Terkecil Parsial (KTP) atau *partial least square* [6], [7].

Metode KTP adalah metode regresi yang mirip dengan metode kuadrat terkecil atau *least square*. Metode ini berfokus pada pemodelan dari informasi yang tersimpan dari setiap variabel, sambil menghindari permasalahan multikolinearitas yang menjadi kelemahan dari metode kuadrat terkecil [8].

Meskipun Metode KTP dapat dengan presisi memodelkan sebuah struktur, namun metode ini juga memiliki kelemahan. Metode KTP memiliki kompleksitas waktu yang bergantung dengan jumlah sampel yang

digunakan. Semakin banyak sampel yang digunakan, semakin lama waktu yang dibutuhkan untuk menyelesaikan algoritma. Lebih lanjut, KTP membutuhkan informasi dari variabel terikat dalam jumlah yang banyak. Hal ini berarti dibutuhkan akselerometer atau sensor dalam jumlah yang banyak yang akan meningkatkan biaya pemeliharaan (*maintenance cost*) dari sebuah struktur.

Kontribusi: berdasarkan pada kajian literatur dan permasalahan yang telah disebutkan diatas, maka penelitian ini secara spesifik bertujuan untuk menghasilkan sebuah algoritma baru berbasis metode KTP yang dapat secara efisien mendeteksi kerusakan struktur. Terkait dengan hal tersebut, maka yang menjadi kontribusi utama dari penelitian ini adalah sebagai berikut:

1. Perumusan dari sebuah algoritma pendeteksi kerusakan struktur berbasis korelasi jarak dan KTP yang membutuhkan sampel lebih sedikit dari metode KTP standar namun, tetap memiliki akurasi model prediktif yang serupa atau bahkan lebih baik dari metode KTP standar.
2. Penurunan dari sebuah parameter baru untuk mengkarakterisasi tingkat kerusakan struktur berdasarkan pada tinjauan statistik.
3. Penyusunan sebuah ide untuk memilih sub himpunan data berdasarkan pada korelasi jarak, yang kemudian dapat digunakan sebagai referensi dalam posisi pemasangan sensor atau akselerometer di dalam sebuah struktur.

Artikel ini disusun dengan formulasi sebagai berikut: Bab II akan membahas mengenai penelitian terkait dan dasar-dasar teori yang dibutuhkan, meliputi: metode KTP, korelasi jarak dan pemilihan sub himpunan berbasis entropi yang telah diajukan pada [13] dan [14]. Bab III akan membahas alur dan persamaan matematis dari algoritma yang diajukan pada penelitian ini dan data yang digunakan untuk memvalidasi performa dari algoritma tersebut. Bab IV akan menampilkan hasil terkait dengan penelitian yang dilakukan. Terakhir, pada Bab V akan disimpulkan hasil dari penelitian beserta dengan analisa penerapan yang mungkin dari algoritma yang digagas, dalam kehidupan nyata.

II. PENELITIAN TERKAIT DAN KAJIAN LITERATUR

Pada bagian ini akan ditunjukkan keterbaharuan dari topik penelitian yang diajukan, lebih lanjut akan dijabarkan secara singkat beberapa konsep utama yang digunakan dalam penelitian ini, yakni metode kuadrat terkecil, konsep pemilihan sub himpunan sensor berdasarkan pada konsep entropi dari teori informasi yang telah digagas dalam [13] dan dikembangkan di [14] serta korelasi jarak yang merupakan variasi nonlinear dari korelasi Pearson (korelasi linear).

A. Keterbaharuan Topik (*State of the art*)

Metode KTP adalah metode regresi yang mirip dengan metode kuadrat terkecil atau *least square*. Metode ini berfokus pada pemodelan dari informasi yang tersimpan dari setiap variabel, sambil menghindari

permasalahan multikolinearitas yang menjadi kelemahan dari metode kuadrat terkecil [8]. Metode ini telah digunakan dalam penelitian pendeteksian kerusakan, seperti pada [9-12].

Tim peneliti pada [9] menggunakan metode KTP untuk memodelkan getaran yang dirasakan oleh sayap pesawat terbang. Peneliti pada [10] melakukan monitoring pada bendungan air dengan menggunakan metode KTP. Berdasarkan pada percobaan yang telah dilakukan, tim peneliti tersebut menyimpulkan bahwa metode KTP dapat dengan presisi memodelkan tingkat deformasi dari bendungan dan tingkat ketinggian air di dalam bendungan. Pada tahun 2017 telah dilakukan penelitian untuk memodelkan deformasi dari bendungan beton dengan menggunakan metode KTP dan *Extreme Learning Machine* (ELM). Metode ini terbukti dapat memodelkan deformasi bendungan dengan tingkat akurasi yang tinggi dan estimasi dapat dilakukan dalam waktu (kompleksitas) yang singkat [11]. Peneliti pada [12] menyatakan bahwa metode *kernel partial least square* dapat digunakan untuk memonitoring sistem pembuangan air limbah secara presisi.

Penelitian ini dilakukan dengan mengikuti garis besar penelitian yang telah dilakukan pada [13] dan [14]. Pada [13] dan [14], penyusunan sub himpunan didasarkan pada perhitungan entropi dari teori informasi yang digunakan untuk memaksimalkan linearitas antar variabel dalam sub himpunan data. Perbedaan utama dari penelitian ini terhadap kedua penelitian tersebut adalah pada metode penyusunan sub himpunan data dan pendeteksian kerusakan. Pada penelitian ini, yang menjadi tinjauan utama adalah korelasi nonlinear dari variabel dan dampaknya pada metode KTP. Oleh karena itu, parameter yang digunakan adalah korelasi jarak yang merupakan parameter yang dapat mengidentifikasi korelasi linear dan non linear antar variabel secara sekaligus. Lebih lanjut, pada [13] dan [14] pendeteksian kerusakan lebih difokuskan pada seberapa cepat metode yang dicetuskan mampu mendeteksi kerusakan di dalam struktur. Sedangkan pada penelitian ini, yang menjadi fokus adalah karakterisasi tingkat kerusakan yang terdapat pada sebuah struktur.

B. Metode Kuadrat Terkecil Parsial

Metode Kuadrat Terkecil Parsial (KTP) merupakan sebuah metode yang menyusun serangkaian kombinasi linear dari input untuk menghasilkan sebuah model regresi dari variabel output. Metode KTP hampir serupa dengan metode Analisis Komponen Utama (AKU) atau *Principal component Analysis* (PCA) (lihat: [13], [15] dan [16] untuk pemahaman dasar terkait dengan metode PCA). Perbedaan utamanya terletak pada metode KTP akan mentransformasi variabel bebas menuju pada ruang baru, dimana varians dari variabel bebas akan dimaksimalkan terhadap variabel terikat. Hal ini menjadi perbedaan utama terhadap AKU yang berfokus pada proyeksi data (secara keseluruhan) ke dalam ruang ortogonal yang bersesuaian [17].

Secara umum, untuk sebuah matriks $X = [\mathbf{x}_1 \ \mathbf{x}_2 \ \dots \ \mathbf{x}_n]$; $X \in \mathbb{R}^{m \times n}$ yang merupakan matriks yang berisi data pengukuran/ variabel bebas, dengan $\mathbf{x}_i =$

$[x_i(1) x_i(2) \dots x_i(m)]^T; x_i \in \mathbb{R}^m$ merupakan hasil pengukuran dari variabel bebas ke- i dengan $i = 1, 2, \dots, n$ dan $\mathbf{y} = [y(1) y(2) \dots y(m)]^T; y \in \mathbb{R}^m$ yang merupakan variabel terikat, maka metode KTP akan memodelkan \mathbf{y} sebagai $\hat{\mathbf{y}} = X\mathbf{B} + c$ dimana \mathbf{B} adalah parameter regresi dari matriks bobot (*weight matrix*) \mathbf{W} , dari variabel terikat Y , dan c adalah konstanta model. Untuk k buah komponen, algoritma dari metode KTP dinyatakan dalam Algoritma 1 [18].

Algoritma 1: Metode Kuadrat Terkecil Parsial

Masukan : X, \mathbf{y}, k
Keluaran : $\hat{\mathbf{y}}$
Proses:
 $X(1) = X$
 $\mathbf{w}(1) = X(1)^T \mathbf{y} / \|X(1)^T \mathbf{y}\|$
 untuk $i = 1: k$
 $\mathbf{z}(i) = \frac{X(i)\mathbf{w}(i)}{[X(i)\mathbf{w}(i)]^T [X(i)\mathbf{w}(i)]}$
 $\mathbf{p}(i) = X(i)^T \mathbf{z}(i)$
 $\mathbf{q}(i) = \mathbf{y}^T \mathbf{z}(i)$
 Jika $\mathbf{q}(i) = 0$
 $i = k$; berhenti
 Lainnya
 $X(i+1) = X(i) - [X(i)\mathbf{w}(i)]^T [X(i)\mathbf{w}(i)] \mathbf{z}(i) \mathbf{p}(i)^T$
 $\mathbf{w}(i+1) = X(i+1)^T \mathbf{y}$
 akhiri jika
 akhiri untuk
 $\mathbf{W} = [\mathbf{w}(1) \mathbf{w}(2) \dots \mathbf{w}(k)]$
 $\mathbf{P} = [\mathbf{p}(1) \mathbf{p}(2) \dots \mathbf{p}(k)]$
 $\mathbf{q} = [q(1) q(2) \dots q(k)]^T$
 $\mathbf{B} = \mathbf{W}(\mathbf{P}^T \mathbf{W})^{-1} \mathbf{q}$
 $c = q(1) - \mathbf{P}^T \mathbf{B}$
 $\hat{\mathbf{y}} = X\mathbf{B} + c$
Selesai

Berdasarkan pada kompleksitas waktu, metode KTP yang dilakukan sebanyak T kali berada dalam orde kompleksitas $O(k \times m \times T)$ dengan m menyatakan jumlah observasi dari sampel dan k menyatakan jumlah komponen KTP yang digunakan [19]. Terkait dengan ini, dapat terlihat bahwa untuk sampel yang semakin besar, maka waktu yang dibutuhkan untuk menjalankan 1 kali perhitungan dari metode ini akan menjadi semakin lama.

C. Pemilihan Sub Himpunan Berbasis Entropi

Pada [13] dan [14], telah ditunjukkan potensi dari algoritma pemilihan sub himpunan berbasis entropi atau *entropy based sensor subset selection* (e-ss) yang dapat meningkatkan performa dari akurasi model dan kemampuan pendeteksian kerusakan berbasiskan pada algoritma AKU dan pohon regresi. Sebagai rujukan literatur awal terhadap bahasan pada sub bab ini, direkomendasikan kepada pembaca untuk mencermati [20] dan [21], terkait dengan dasar dari teori informasi, terutama: entropi, entropi kondisional dan *Information Gain* (IG). Secara spesifik, ide utama yang terdapat pada penelitian [13] dan [14] adalah memilih sub himpunan dari sebuah

himpunan data, yang dimana sub himpunan ini memiliki varians maksimal untuk setiap anggota sub himpunan. Pemilihan ini akan memastikan bahwa ada linearitas dari tiap anggota sub himpunan.

Diberikan $X = [x_1 x_2 \dots x_n]; X \in \mathbb{R}^{m \times n}$ sebagai matriks yang berisi data pengukuran, dengan $x_i = [x_i(1) x_i(2) \dots x_i(m)]^T; x_i \in \mathbb{R}^m$ merupakan hasil pengukuran dari variabel acak X_i dengan $i = 1, 2, \dots, n$, sehingga x_i adalah sampel pengukuran (*random outcome*) dari variabel acak X_i . Misalkan bahwa persoalan yang diberikan adalah untuk memilih sub himpunan yang dapat memprediksi nilai variabel i . Pertama-tama, definisikan $Z_{ij}, \forall j \neq i$ yang merupakan bilangan acak binari sebagai

$$Z_{ij} = \begin{cases} 1 & \text{jika } |X_i - \beta_{ij} X_j| < \theta \cdot \sigma(X_i), \\ 0 & \text{jika lainnya} \end{cases}, \quad (2.1)$$

dengan: $\theta \in [1, +\infty) \subset \mathbb{R}$, $\sigma(\cdot): \mathbb{R}^m \rightarrow \mathbb{R}$ menyatakan standar deviasi dan $\beta_{ij} = (\mathbf{x}_j^T \mathbf{x}_i)^{-1} \mathbf{x}_j^T \mathbf{x}_i$ merupakan solusi metode kuadrat terkecil (*least square*) dari X_i dan X_j . Lebih lanjut, misalkan bahwa $p_{ij} = \mathbb{P}(Z_{ij} = 1)$ merupakan peluang agar variabel acak Z_{ij} menghasilkan nilai 1. Maka *Information Gain* (IG) dari Z_{ij} yang dikondisikan terhadap $\beta_{ij}, I(Z_{ij}; \beta_{ij})$ dapat dinyatakan sebagai

$$I(Z_{ij}; \beta_{ij}) = H(Z_{ij}) - H(Z_{ij} | \beta_{ij}),$$

$$H(Z_{ij}) = -[p \cdot \log_2(p) + (1-p) \cdot \log_2(1-p)], \quad (2.2)$$

dengan $H(Z_{ij})$ adalah entropi dari variabel acak Z_{ij} dan $H(Z_{ij} | \beta_{ij})$ adalah entropi kondisional dari Z_{ij} terhadap β_{ij} . Sekarang, persoalan dapat diselesaikan dengan mencari pasangan variabel yang menghasilkan nilai $I(Z_{ij}; \beta_{ij})$ maksimum. Namun, karena $H(Z_{ij})$ dan $H(Z_{ij} | \beta_{ij})$ tidak diketahui (meskipun secara pasti, nilai dari $H(Z_{ij})$ adalah sebuah konstanta), maka jalan yang paling dimungkinkan adalah dengan memilih j^* sehingga dihasilkan $H(Z_{ij} | \beta_{ij})$ terkecil. Pemilihan dari j^* dimungkinkan berkat Teorema 1 (pembuktian dari teorema ini dapat dilihat di [13] dan [14])

Teorema 1: diberikan

$$\hat{H}(Z_{ij} | \beta_{ij}) \doteq -[r_\theta(m) \cdot \log_2(r_\theta(m)) + (1-r_\theta(m)) \cdot \log_2(1-r_\theta(m))],$$

$$r_\theta(m) = 1 - \theta^{-2} + \theta^{-2} r_{ij}^2(m), \quad (2.3)$$

Seiring dengan $m \rightarrow \infty$ dan untuk $r_{ij}^2 \in [1 - \frac{\theta^2}{2}, 1]$, maka

$$H(Z_{ij} | \beta_{ij}) < \hat{H}(Z_{ij} | \beta_{ij}), \quad (2.4)$$

dimana, r_{ij}^2 merupakan koefisien determinasi (lihat: [22]) dari variabel i dan j .

Persamaan (2.4) dari Teorema 1 menjelaskan bahwa $H(Z_{ij} | \beta_{ij})$ dapat diestimasi batas atasnya dengan (2.3). Hal ini yang kemudian memungkinkan untuk memaksimalkan $I(Z_{ij} | \beta_{ij})$. Sekumpulan dari variabel yang memaksimalkan nilai IG terhadap variabel i kemudian dikenal sebagai sub himpunan berbasis entropi. Algoritma dari metode ini dinyatakan dalam algoritma 2.

Terhadap kompleksitas waktu, terlihat dari Algoritma 2, proses utama dari algoritma ini bergantung hanya pada perhitungan $\hat{H}(Z_{ij}|\beta_{ij})$ yang dimana, proses ini dapat diselesaikan dalam orde $O(n)$, dengan n menyatakan banyaknya variabel. Sehingga algoritma dari e-ss sangat ringan dan tidak akan memberatkan proses perhitungan.

Algoritma 2: *Pemilihan Sub Himpunan Berbasis Entropi*

Masukan : X ; variabel yang akan diprediksi, $i \in n$; banyaknya variabel yang di inginkan $n^* < n$; θ
Keluaran : sekumpulan indeks $S \subset n$, $|S| = n^*$
Proses:
 $S = \{i\}$
 untuk $k = 1: n^* - 1$
 $j^* = \operatorname{argmin}_{j \in n \setminus S} \hat{H}(Z_{ij}|\beta_{ij})$
 $S = S \cup \{j^*\}$
 Akhiri untuk
 Selesai

D. Korelasi Jarak

Pada tinjauan teori probabilitas, korelasi jarak atau *distance correlation* adalah sebuah besaran yang mengukur hubungan (tidak harus berupa linear) antara dua buah variabel [23]. Berbeda dari korelasi Pearson yang seutuhnya mengukur hubungan linear antar variabel, korelasi jarak akan bernilai 0 jika dan hanya jika 2 variabel bebas (independen) satu sama lain [24] dan bernilai 1 jika dan hanya jika terdapat konstanta fungsi, sehingga kedua variabel dapat terhubung secara linear secara sempurna dengan konstanta tersebut.

Seperti pada sebelumnya, misalkan $X \in \mathbb{R}^{m \times n}$ adalah himpunan data yang tersusun dari vektor pengukuran $x_i; i = 1, 2, \dots, n$. Misalkan bahwa terdapat permasalahan untuk menentukan korelasi jarak dari 2 buah variabel, yakni x_d dan x_e dimana $x_d, x_e \in X$. Dengan sedikit ‘‘pemaksaan notasi’’ (*abuse of notation*), misalkan

$$D = \begin{bmatrix} d_{1,1} & d_{1,2} & \dots & d_{1,m} \\ d_{2,1} & d_{2,2} & \dots & d_{2,m} \\ \vdots & \vdots & \ddots & \vdots \\ d_{m,1} & d_{m,2} & \dots & d_{m,m} \end{bmatrix}, \quad (2.5)$$

sebagai matriks jarak dari x_d , dengan

$$d_{j,k} = \|x_d(j) - x_d(k)\|_2, \quad (2.6)$$

dimana $\|\cdot\|_2: \mathbb{R}^m \rightarrow \mathbb{R}$ menyatakan *Euclidean norm* atau *L₂ norm* [25]. Lebih lanjut, misalkan

$$\mathfrak{D} = \begin{bmatrix} \delta_{1,1} & \delta_{1,2} & \dots & \delta_{1,m} \\ \delta_{2,1} & \delta_{2,2} & \dots & \delta_{2,m} \\ \vdots & \vdots & \ddots & \vdots \\ \delta_{m,1} & \delta_{m,2} & \dots & \delta_{m,m} \end{bmatrix} \quad (2.7)$$

sebagai sebuah matriks jarak terpusat (*centered distance matrix*) dari D dengan elemen baris ke j dan kolom ke k dari matriks ini dinyatakan sebagai

$$\begin{aligned} \delta_{j,k} &= d_{j,k} - \bar{d}_{j,\cdot} - \bar{d}_{\cdot,k} + \bar{d}_{\cdot,\cdot}; \\ \bar{d}_{j,\cdot} &= \frac{1}{m} \sum_{p=1}^m d_{j,p}; \\ \bar{d}_{\cdot,k} &= \frac{1}{m} \sum_{p=1}^m d_{p,k}; \\ \bar{d}_{\cdot,\cdot} &= \frac{1}{m^2} \sum_{p=1}^m \sum_{q=1}^m d_{p,q}; \end{aligned} \quad (2.8)$$

dengan: $\bar{d}_{j,\cdot}$; $\bar{d}_{\cdot,k}$ dan $\bar{d}_{\cdot,\cdot}$ secara berurutan menyatakan rerata baris ke j , kolom ke k dan rerata keseluruhan (rerata semua elemen pada matriks D). Dengan cara yang sama, dapat diturunkan pula matriks jarak dan matriks jarak terpusat untuk x_e , yang dalam penelitian ini akan disebut sebagai E untuk matriks jarak dan \mathcal{E} untuk matriks jarak terpusat.

Berdasarkan pada definisi yang telah diberikan diatas, maka korelasi jarak antara variabel x_d dan x_e dapat dinyatakan sebagai:

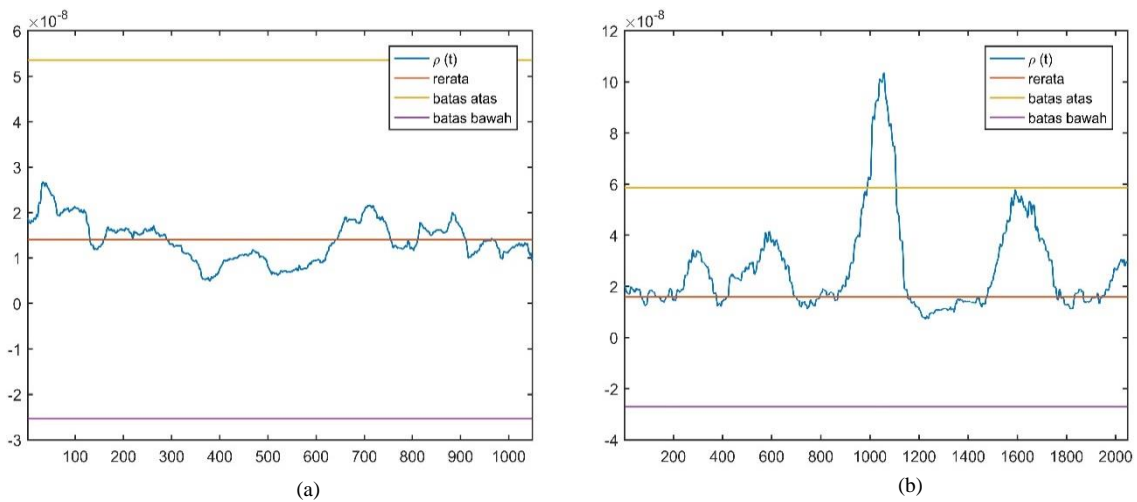
$$\begin{aligned} R_d(x_d, x_e) &= \frac{\sigma_{x_d, x_e}}{\sqrt{\sigma_{x_d}^2 \sigma_{x_e}^2}}, \\ \sigma_{x_d, x_e} &= \frac{1}{m^2} \sum_{p=1}^m \sum_{q=1}^m \delta_{p,q} \epsilon_{p,q}, \\ \sigma_{x_d}^2 &= \frac{1}{m^2} \sum_{p=1}^m \sum_{q=1}^m \delta_{p,q}^2, \\ \sigma_{x_e}^2 &= \frac{1}{m^2} \sum_{p=1}^m \sum_{q=1}^m \epsilon_{p,q}, \end{aligned} \quad (2.9)$$

dengan σ_{x_d, x_e} menyatakan kovarians jarak dari x_d dan x_e , $\sigma_{(\cdot)}^2: \mathbb{R}^m \rightarrow \mathbb{R}$ adalah fungsi varians jarak dan $\epsilon_{j,k}$ menyatakan elemen baris ke- j kolom ke- k dari matriks \mathcal{E} .

Meskipun korelasi jarak merupakan besaran yang efektif untuk mengkuantifikasi hubungan antar variabel, namun perlu diperhatikan bahwa korelasi jarak memiliki kompleksitas sebesar $O(m^2)$. Dimana, proses perhitungan dari korelasi jarak dapat menjadi ‘‘berat’’ ketika jumlah sampel menjadi semakin besar [26].

III. METODOLOGI

Bagian ini akan berpusat pada penjabaran algoritma pendeteksi kerusakan struktur bangunan berbasis korelasi jarak dan KTP yang telah dipaparkan pada bab sebelumnya.



Gambar 1 Contoh dinamika dari $\rho(t)$ pada kondisi nominal (a) dan kondisi terdapat kerusakan (b)

Pada bagian ini, akan dibahas pula struktur data yang digunakan yang berasal dari situs penyimpanan data gratis yang dikelola oleh Los Alamos National Laboratory (LANL).

Secara umum, algoritma yang digagas memiliki 3 tahapan utama: pemilihan sub himpunan data, pemodelan matematis dan pendeteksian kerusakan. Pada pemilihan sub himpunan, dari sejumlah sensor yang ada akan dipilih sekelompok sensor yang memiliki korelasi jarak yang “baik” satu sama lain. Kemudian, akan diturunkan sebuah model matematis yang dapat merepresentasikan dinamik dari sensor tersebut. Barulah kemudian, pemeriksaan terhadap kerusakan struktur dapat dilakukan untuk mengidentifikasi keberadaan dari kerusakan. Detil dari algoritma ini akan dipaparkan dalam sub bagian A.

A. Algoritma Pendeteksi Kerusakan Struktur Berbasis Korelasi Jarak dan KTP

Untuk melakukan proses pendeteksian kerusakan struktur berbasis korelasi jarak dan KTP, ada 3 tahapan utama yang perlu dilalui: penentuan sub himpunan data, penyusunan model prediktif dan diagnosis kerusakan struktur. Untuk menjabarkan isi dari sub bagian ini dengan baik, seperti pada bab sebelumnya, misalkan bahwa $X = [x_1 \ x_2 \ \dots \ x_n]$; $X \in \mathbb{R}^{m \times n}$ adalah sebuah matriks yang berisikan data pengukuran dinamik (respon dinamis) yang diperoleh dari pembacaan respon vibrasional oleh akselerometer yang dipasang pada sebuah struktur, dengan $x_i \in \mathbb{R}^m$; $x_i = [x_i(1) \ x_i(2) \ \dots \ x_i(m)]^T$ merupakan hasil pengukuran respon dinamis dari akselerometer ke- i dengan $i = 1, 2, \dots, n$.

Pemilihan sub himpunan: Misalkan bahwa terdapat permasalahan untuk menentukan sub himpunan data yang memiliki n^* buah akselerometer untuk mengestimasi dinamika dari akselerometer i . Permasalahan ini dapat diselesaikan dengan memilih sejumlah akselerometer lain dalam X yang memiliki korelasi jarak yang tinggi dengan i .

Berdasarkan pada X yang diberikan diatas, dan alur perhitungan dari (2.4) hingga (2.9), maka dapat ditentukan

matriks korelasi dari X . Dalam kasus ini, setiap elemen dari matriks korelasi jarak, D dapat dinyatakan sebagai

$$D_{i,j} = R_d(x_i, x_j) = \frac{\sigma_{x_i, x_j}}{\sqrt{\sigma_{x_i}^2 \sigma_{x_j}^2}} \quad (3.1)$$

Secara konstruksi, variabel i pasti memiliki korelasi sempurna dengan dirinya sendiri, oleh karena itu, hanya dibutuhkan $n^* - 1$ variabel lain untuk membentuk sub himpunan data. Proses pemilihan variabel untuk sub himpunan ini dinyatakan dalam Algoritma 3.

Algoritma 3: Pemilihan Sub Himpunan Berbasis Korelasi Jarak

Masukan : X ; variabel yang akan diprediksi, $i \in n$; banyaknya variabel yang di inginkan $n^* < n$; θ

Keluaran : sekumpulan indeks $S \subset n, |S| = n^*$

Proses:

$$D_{i,j} = \frac{\sigma_{x_i, x_j}}{\sqrt{\sigma_{x_i}^2 \sigma_{x_j}^2}}$$

$$S = \{i\}$$

untuk $k = 1 : n^* - 1$

$$j^* = \underset{j \in n \setminus S}{\operatorname{argmax}} D_{i,j}$$

$$S = S \cup \{j^*\}$$

Akhiri untuk

Selesai

Pemodelan Matematis: Setelah sub himpunan data untuk variabel i diperoleh, langkah selanjutnya adalah menyusun model matematis untuk mengestimasi nilai dari variabel i untuk setiap saat dengan metode KTP.

Diberikan $X_i \in \mathbb{R}^{m \times n^* - 1}$ sebagai sub himpunan data dari X , dengan setiap variabel terkandung di dalam X_i memiliki korelasi jarak yang tinggi terhadap variabel i . Berdasarkan pada definisi diatas, estimasi dari i , yang disebut sebagai \hat{i} dapat diperoleh dengan mengikuti alur perhitungan sesuai pada Algoritma 1.

Diagnosis kerusakan: Tahap pemodelan matematis akan menghasilkan estimasi dari variabel i , yang bergantung pada variabel yang terkandung dalam X_i . Misalkan

$$\mathbf{r} = (\mathbf{x}_i - \hat{\mathbf{x}}_i)^2 \tag{3.2}$$

sebagai vektor residual kuadrat dari i dan \hat{i} . Secara probabilistik, maka setiap residual kuadrat dari i dan \hat{i} dapat dinyatakan dalam distribusi

$$r = \bar{r} \pm k \cdot \sigma(\mathbf{r}), \tag{3.3}$$

Dimana $k \in \mathbb{Z}^+$ adalah sebuah parameter *tuning*. Jika $k = 3$, maka secara statistik terdapat peluang sebesar 99,7% sampel acak berada dalam rentang distribusi tersebut [27]. Secara umum, semakin besar nilai n maka semakin besar pula rentang data yang bisa di cakup, dan begitu pula sebaliknya.

Pendeteksian kerusakan struktur pada penelitian ini akan didasarkan pada fakta tersebut. Diasumsikan bahwa jika model yang dihasilkan dari data sebuah struktur yang dinyatakan dalam kondisi baik memiliki akurasi model yang tinggi, maka akan ada $k^* \in \mathbb{Z}^+$ sehingga semua residual kuadrat akan terkandung dalam \mathbf{r} . Jika terdapat kerusakan pada struktur, maka model tidak lagi dapat mengestimasi sistem dengan baik, sehingga residual kuadrat yang dihasilkan jauh lebih besar daripada yang didefinisikan pada (3.3).

Dalam kasus tertentu, bisa saja terjadi *bug* atau malfungsi yang mengakibatkan kesalahan pembacaan dinamika oleh akselerometer. Hal ini akan berakibat pada status sebuah struktur dinyatakan sebagai terdapat kerusakan, meskipun nyatanya dalam keadaan baik. Kasus demikian disebut juga sebagai negatif palsu (*false negative*), dan untuk meminimalisir kasus negatif palsu, maka perlu didefinisikan sebuah parameter baru

$$\rho(t) = \left(\sum_{j=-\lfloor \frac{1}{2}N \rfloor}^{\lfloor \frac{1}{2}N \rfloor} l(t+j) \right)^{-1} \cdot \sum_{i=-\lfloor \frac{1}{2}N \rfloor}^{\lfloor \frac{1}{2}N \rfloor} r(t+i),$$

$$l(t) = \begin{cases} 1 & \text{jika } l(t) \in \mathbf{r} \\ 0 & \text{jika lainnya} \end{cases}, \tag{3.4}$$

dimana, $\rho(t)$ menyatakan rerata bergerak (*moving average*) sebanyak N elemen dari \mathbf{r} dan $\lfloor \cdot \rfloor : \mathbb{R} \rightarrow \mathbb{R}$ merupakan fungsi pembulatan kebawah. Berdasarkan pada (3.4), sebuah struktur dinyatakan dalam kondisi nominal atau baik apabila

$$\bar{r} - k \cdot \sigma(\mathbf{r}) \leq \rho \leq \bar{r} + k \cdot \sigma(\mathbf{r}), \tag{3.5}$$

dan rusak, apabila sebaliknya.

Gambar 1(a) menunjukkan contoh dinamika dari sebuah akselerometer dalam kondisi nominal dan jika terdapat kerusakan struktur (*faulty*) dengan $k = 2$. Pada kondisi nominal, terlihat bahwa garis biru yang merupakan representasi $\rho(t)$ berada dalam batas yang direpresentasikan dalam garis berwarna kuning dan ungu. Dalam kasus ini, disimpulkan bahwa tidak terdapat kerusakan pada struktur. Sedangkan, pada Gambar 1(b) terlihat bahwa garis berwarna biru telah melewati batas atas

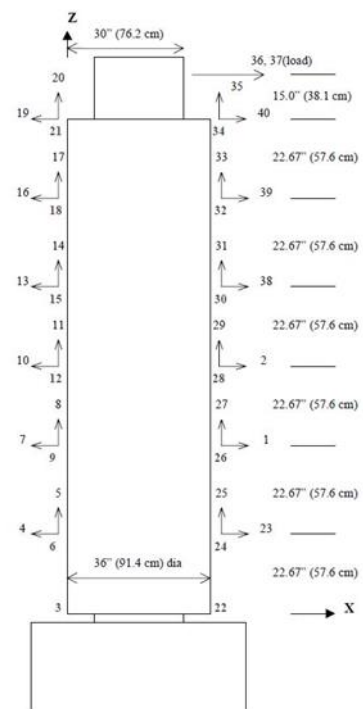
dari residual. Sehingga, disimpulkan bahwa terdapat kerusakan pada struktur.

Terkait dengan tingkat kerusakan struktur, apabila kerusakan struktur semakin masif, maka $\rho(t)$ akan semakin jauh melewati batas yang diizinkan. Untuk menyelidiki ini, maka dapat didefinisikan sebuah besaran

$$T_k = \frac{\max(\rho) - \bar{r}}{\sigma(\mathbf{r})}, \tag{3.6}$$

dimana T_k adalah sebuah besaran skalar, serupa dengan k pada (3.3) yang menjelaskan seberapa besar kuadrat residu telah menyimpang terhadap reratanya. Semakin besar T_k maka semakin tinggi tingkat kerusakan yang ada di sebuah struktur. Klaim ini akan dibuktikan pada Bab IV.

B. Studi kasus: Tiang Jembatan Irvine, California.



Gambar 2. Posisi akselerometer pada tiang jembatan Irvine.

Metode yang telah dijelaskan pada Sub Bab sebelumnya kemudian akan diujikan pada kumpulan data untuk menguji tingkat kepresisian dari algoritma. Pada penelitian ini, himpunan data yang digunakan adalah pengukuran respon dinamik dari sebuah tiang jembatan yang dilakukan oleh *University California Irvine* (UCI), dimana data ini dapat diakses secara gratis dari repositori data pada [28].

Tiang jembatan ini merupakan struktur yang tersusun dari silinder beton berdiameter 0,61 meter yang dimasukan kedalam silinder beton lain yang berdiameter 0,91 meter. Kedua silinder memiliki tinggi 3,45 meter dan dipasangkan diatas fondasi beton setinggi 0,635 meter. Pada tiang tersebut, dipasang 40 buah akselerometer yang berguna untuk mengukur respon dinamik dari struktur. Gambar 2 mengilustrasikan penampang lintang dari tiang jembatan beserta posisi dari 40 akselerometer yang dipasangkan di

dalamnya. Struktur tersebut digetarkan dengan menggunakan *shaker* yang di pasang di bagian fondasi dengan frekuensi yang berkisar dari 0 hingga 400 Hz, selama 8 sekon dan menghasilkan 2048 sampel pengukuran. Percobaan tersebut diulang sebanyak 2 kali sehingga dihasilkan 2 himpunan data nominal.

Untuk mensimulasikan kerusakan pada struktur, maka di bagian atas tiang jembatan kemudian dipasang dengan seperangkat aktuator hidrolik. Aktuator tersebut akan memberikan gaya tambahan pada tiang, yang mengakibatkan deformasi lateral pada tiang. Besar deformasi yang diberikan adalah sebesar 0,25 Δy; 0,4 Δy dan 0,7 Δy, dengan Δy menyatakan deformasi lateral yang dialami oleh tiang, relatif terhadap arah sumbu y dari tiang. Percobaan ini diulangi sebanyak beberapa kali dan menghasilkan 6 buah himpunan data dengan tingkat kerusakan yang berbeda-beda (setiap himpunan data memiliki 2048 sampel).

Pada penelitian ini, untuk memvalidasi performa dari algoritma yang telah digagas, maka akan digunakan beberapa kumpulan data yang disediakan oleh UCI. Secara spesifik, sebanyak 3072 sampel dari himpunan data nominal akan digunakan untuk menyusun model matematis.

Lebih lanjut, 1024 sampel yang tersisa akan digunakan untuk menguji akurasi prediktif dari model yang telah dibangun. Himpunan data yang bersesuaian dengan deformasi 0,25 Δy kemudian akan diasosiasikan dengan data tingkat kerusakan I. Sedangkan data dengan deformasi 0,4 Δy dan 0,7 Δy secara berurutan dikategorikan sebagai tingkat kerusakan II dan III (Semakin tinggi tingkat kerusakan, semakin parah kerusakan yang terjadi di dalam struktur). Data tersebut kemudian akan digunakan untuk menguji performa dari algoritma dalam mendeteksi kerusakan dari sebuah struktur.

IV. HASIL DAN PEMBAHASAN

Pada Bab ini akan ditampilkan hasil percobaan dari algoritma yang diusulkan terhadap himpunan data yang disediakan oleh UCI. Terdapat 2 hal yang menjadi pokok bahasan pada Bab ini: tingkat kepresisian dari model yang dibangun, dan kemampuan metode yang diusulkan dalam mendeteksi kerusakan struktur.

Pada penelitian ini, akurasi dari model prediktif diukur dengan menggunakan besaran akar rerata kuadrat ternormalisasi atau *Normalized Root Mean Square Error* (NRMSE) yang dinyatakan sebagai

$$NRMSE(\mathbf{y}, \hat{\mathbf{y}}) = \frac{1}{\bar{y}} \sqrt{\frac{\sum_{i=1}^m [y(i) - \hat{y}(i)]^2}{m}}, \quad (4.1)$$

dengan $\mathbf{y} \in \mathbb{R}^m$ merupakan besaran yang ingin di prediksi, $\hat{\mathbf{y}} \in \mathbb{R}^m$ merupakan prediksi dari \mathbf{y} dan \bar{y} merupakan rerata dari \mathbf{y} . NRMSE memiliki kisaran nilai dari 0 hingga 1. Semakin kecil NRMSE, semakin kecil pula proporsi varians yang tidak dapat diprediksi oleh model. Untuk mempermudah pembacaan akurasi, maka hasil yang akan ditampilkan pada tabel adalah $1 - NRMSE \times 100\%$.

Dimana, besaran ini menyatakan seberapa banyak (dalam persen) varians dari variabel terikat yang dapat diestimasi oleh model. Pada tahap ini, dari total 40 sub himpunan data yang dibentuk, hanya akan diambil sub himpunan data yang memiliki akurasi diatas akurasi rerata untuk dipakai dalam proses pendeteksian kerusakan. Hal ini dimaksudkan agar hanya model dengan akurasi tinggi saja yang digunakan dalam proses pendeteksian kerusakan struktur.

TABEL I
AKURASI MODEL PREDIKTIF DARI METODE KTP. MENGINDIKASIKAN NILAI TERTINGGI DARI KETIGA METODE

N. Sensor	KTP-jarak	KTP-entropi	KTP-standar
1	94,77	94,38	95,20*
2	84,70	85,05*	84,78
3	93,30*	93,25	92,91
4	91,19	90,20	91,48*
5	87,87*	87,78	87,86
6	95,95*	94,61	94,77
7	95,41	95,73	95,74*
8	89,99*	89,08	89,66
9	95,78*	95,08	95,03
10	96,19*	95,51	95,76
11	84,99	84,93	84,59
12	94,25	93,15	94,81*
13	97,08*	96,82	96,29
14	89,43*	87,39	88,63
15	95,57*	93,57	94,25
16	96,57*	95,14	95,49
17	89,00	89,42*	88,87
18	96,37*	95,39	95,89
19	95,66	96,01*	94,50
20	89,45	89,49*	88,77
21	96,09*	95,86	94,77
22	92,17	92,17	92,58*
23	92,32*	91,05	92,04
24	95,49*	95,04	94,14
25	88,30*	88,21	87,83
26	94,74*	94,74*	94,58
27	88,54	88,34	88,85*
28	95,11*	94,83	93,86
29	86,32	84,92	86,48*
30	95,90	95,90	96,13*
31	89,12	85,49	90,09*
32	95,86	94,26	96,36*
33	89,63	88,36	89,68*
34	94,70*	94,70*	94,28
35	92,50*	90,95	92,50*
36	90,20	90,23	90,63*
37	84,69	84,79	85,10*
38	96,83*	96,74	96,51
39	83,94*	83,45	83,51
40	83,85	83,70	83,99*
Rerata	92,00*	91,39	91,73

TABEL III
NILAI T_k UNTUK 4 KASUS: NOMINAL, KERUSAKAN TINGKAT I, II DAN III UNTUK KETIGA MODEL KTP. WARNA MERAH KASUS

N. sensor	Nominal			Kerusakan Tk. I			Kerusakan Tk. II			Kerusakan Tk. III		
	Jarak	Entropi	Standar	Jarak	Entropi	Standar	Jarak	Entropi	Standar	Jarak	Entropi	Standar
1	1,7	1,3	0,8	0,7	4,1	1,4	20,0	15,9	55,4	25,7	14,8	77,8
3	0,4	0,5	1,2	3,8	2,1	9,0	330,0	328,6	388,4	318,5	316,3	422,2
6	0,3	0,3	0,1	1,2	2,4	4,5	10,7	93,8	78,7	14,3	116,4	128,9
7	0,6	1,4	0,6	3,4	2,4	2,7	26,6	8,9	53,7	33,5	10,7	72,7
9	0,3	1,2	0,8	2,3	4,5	4,0	266,6	48,8	158,0	236,2	39,4	161,0
10	1,6	1,0	1,2	3,8	3,1	1,6	3,8	62,4	4,7	4,5	90,0	5,3
12	0,9	0,3	0,6	1,2	3,8	1,8	76,6	32,4	94,1	74,3	53,2	83,2
13	0,1	0,1	1,0	3,9	2,5	3,0	14,0	8,7	47,0	19,8	5,4	69,7
15	3,8	2,6	1,6	0,1	4,4	1,0	10,8	24,6	11,0	14,1	22,8	15,4
16	1,1	0,9	0,8	1,2	1,3	0,9	6,7	96,8	19,8	5,9	131,1	22,3
18	2,3	0,8	2,0	0,5	0,7	1,9	87,5	34,6	107,4	90,0	35,7	129,4
19	1,7	2,1	1,5	0,5	1,1	0,2	10,3	42,7	10,5	14,4	52,5	16,3
21	1,1	0,2	0,1	1,3	1,8	1,4	54,1	96,9	44,6	63,5	95,7	64,3
22	0,7	0,7	0,4	3,7	3,7	11,4	294,9	294,9	368,7	282,7	282,7	403,3
24	2,5	1,4	1,0	0,9	3,3	7,5	9,4	130,3	91,6	11,9	133,2	131,2
26	1,0	1,0	0,1	3,2	3,1	3,8	91,2	91,2	153,7	80,9	80,9	143,6
28	0,8	0,8	0,3	4,0	6	3,0	78,1	28,8	124,2	67,3	22,3	104,7
30	0,3	0,3	0,3	4,0	4	5,8	21,8	21,8	36,8	26,9	26,9	49,4
32	0,8	0,4	0,5	7,9	1,1	5,2	61,7	72,6	121,5	53,3	65,3	129,0
34	1,0	1,0	0,7	23,2	23,2	16,7	48,6	48,6	34,1	52,1	52,1	48,0
38	0,4	0,1	1,0	1,5	2	1,3	15,2	10,1	56,7	18,6	11,5	77,1
Rerata	1,1	0,9	0,8	3,4	3,8	4,2	73,3	75,9	98,1	71,8	79,0	112,1
F_p	-	-	-	52%	48%	52%	0%	0%	0%	0%	0%	0%
F_n	5%	0%	0%	-	-	-	-	-	-	-	-	-

Pada tahap pendeteksian struktur bangunan, ada 2 parameter yang akan diamati. Pertama, tingkat dari positif palsu, yakni jumlahan dari data yang terindikasi memiliki kerusakan struktur, namun dilaporkan oleh algoritma sebagai nominal. Persentase positif palsu dinyatakan sebagai

$$F_p = \frac{f_t}{n_a} \times 100\% , \quad (4.2)$$

dengan F_p adalah tingkat positif palsu (dalam persen), n_a adalah jumlah sub himpunan yang memiliki akurasi model prediktif diatas rata rata, f_t adalah jumlah sub himpunan data yang tidak dapat mendekteksi adanya kerusakan struktur berdasarkan pada algoritma yang dirumuskan ketika diberi input data yang terdapat kerusakan struktur. Secara spesifik, f_t bernilai 0 jika tidak ada penyimpangan terhadap kondisi pada (3.5) dan 1 jika lainnya. Parameter berikutnya yang akan diteliti adalah negatif palsu. Negatif palsu merupakan sebuah kondisi dimana struktur yang dalam keadaan nominal dilaporkan sebagai rusak. Persentase dari negatif palsu dinyatakan sebagai

$$F_n = \frac{f_n}{n_a} \times 100\% , \quad (4.2)$$

dengan F_n adalah persentase negatif palsu dan f_n adalah jumlah sub himpunan yang melaporkan adanya kerusakan struktur, pada saat diberikan masukan data nominal, atau data yang tidak mengandung informasi kerusakan struktur.

Pada penelitian ini, digunakan nilai $k = 3$ dan $N = 100$ untuk parameter identifikasi pada (3.4) dan (3.5). Hal ini

dimaksudkan untuk menghindari kesalahan deteksi karena negatif palsu. Lebih lanjut, berdasarkan pada data yang diuntuk setiap sub himpunan akan memiliki elemen sebanyak 5 atau $n^* = 5$. Tentu saja, nilai ini dapat dipilih oleh pengguna, menyesuaikan pada konteks kasus yang akan digunakan. Sebagai perbandingan terhadap algoritma yang diusulkan, akan ditampilkan juga akurasi dari model prediktif dan hasil pendeteksian kerusakan dari 2 metode lainnya, yakni: metode berbasis entropi yang dijabarkan dalam [13] dan [14] dan metode PLS dasar yang menggunakan seluruh variabel yang ada dalam 1 model.

A. Akurasi Model Prediktif

Tabel 1 menampilkan akurasi model prediktif yang diukur sebagai $1 - NRMSE \times 100\%$ dari metode KTP korelasi jarak, KTP entropi dan KTP standar untuk 40 akselerometer yang terpasang didalam struktur.

Terlihat dari Tabel 1, bahwa metode KTP korelasi jarak memiliki tingkat akurasi yang sedikit lebih baik ketimbang 2 metode KTP lainnya. Hal ini dikarena dari korelasi jarak mampu menunjukkan hubungan nonlinear dari 2 buah variabel, yang kemudian dapat dimodelkan dengan baik oleh metode KTP. Perlu dicermati bahwa metode KTP korelasi jarak dan KTP entropi hanya menggunakan informasi 4 buah akselerometer (setiap sub himpunan memiliki 5 anggota, namun 1 digunakan sebagai variabel terikat) namun memiliki akurasi yang mendekati atau bahkan lebih baik daripada KTP dengan semua variabel. Hal ini menunjukkan bahwa algoritma pemilihan sub

himpunan data berbasis korelasi jarak mampu dengan baik memilih variabel bebas untuk memodelkan variabel terikat tanpa menurunkan (atau bahkan meingkatkan) kualitas model prediktif.

Keuntungan lain dari pemilihan sub himpunan data berbasis korelasi jarak terletak pada kompleksitas algoritma. Perlu diketahui bahwa metode KTP berada pada $O(k \times m \times T)$ dengan k menyatakan banyaknya komponen KTP, m menyatakan banyaknya sampel dan T adalah satuan pengulangan dari algoritma. Terlihat bahwa untuk m yang semakin meningkat, maka waktu yang dibutuhkan untuk menjalankan 1 kali metode KTP akan semakin lama. Dengan melakukan pemilihan sub himpunan terlebih dahulu, maka jumlah sampel dapat dikurangi. Sehingga, algoritma masih dapat menghasilkan model yang akurat, dan efisien dari segi waktu.

Berdasarkan dari analisa akurasi model prediktif, terlihat bahwa akselerometer no. 1, 3, 6, 7, 9, 10, 12, 13, 15, 16, 18, 19, 21, 22, 24, 26, 28, 30, 32, 34 dan 38 adalah akselerometer yang memiliki akurasi model diatas rerata. Sehingga, hanya akselerometer tersebut yang akan diuji tingkat kesensitivitasnya dalam mendeteksi kerusakan di dalam struktur. Perlu diperhatikan bahwa, sebagai contoh akselerometer no. 23 memiliki akurasi diatas rata-rata untuk metode KTP standar dan korelasi jarak. Namun, akselerometer no. 23 tidak diikuti karena akselerometer tersebut memiliki akurasi model dibawah rerata untuk metode KTP entropi. Sehingga, untuk membuat perbandingan yang adil, hanya akselerometer yang ketiga modelnya memiliki akurasi model prediktif diatas rerata yang bersesuaian saja yang akan digunakan dalam tahap pendeteksian kerusakan struktur.

B. Pendeteksian Kerusakan Struktur.

Tabel 2. Merepresentasikan hasil analisa kerusakan struktur dari metode KTP standar, entropi dan korelasi jarak. Tabel 2 memuat informasi nilai T_k sesuai dengan (3.6) yang menunjukkan simpangan terjauh dari data. Lebih lanjut, Tabel 2 juga memuat informasi terkait dengan tingkat negatif dan positif palsu.

Terlihat bahwa ketiga metode dapat dengan sempurna memprediksi kerusakan struktur di tingkat II dan III, serta dapat mengenali data struktur nominal dengan baik. Permasalahan secara spesifik terlihat dari pendeteksian kerusakan struktur pada tingkat kerusakan I, dimana ketiga metode memiliki tingkat kesalahan hingga 50% dalam mendeteksi kerusakan kecil. Perlu diketahui bahwa metode pendeteksi kerusakan yang diusulkan bergantung pada nilai k . Semakin besar nilai k , maka semakin sulit bagi algoritma untuk mendeteksi kerusakan akibat batas atas dan bawah yang terlewat besar. Namun, memilih k yang kecil, sangat rentan dengan kasus negatif palsu pada pembacaan data nominal. Hingga saat ini, belum terdapat informasi bagaimana cara memilih rentang yang baik. Sebagian besar metode masih berfokus pada cara heuristik, yang mengandalkan *insting* atau "perasaan" dari pengguna. Terkait dengan kasus ini, adapun parameter k yang dirasa paling tepat adalah 2, bukan 3. Namun,

penelitian ini bertujuan untuk memperlihatkan kemampuan algoritma untuk mendeteksi kerusakan dan bukan menentukan parameter optimal. Sehingga, bahasan terkait nilai parameter *tuning* yang optimal tidak akan dibahas pada saat ini.

Hal menarik lain yang bisa diamati dari Tabel 2 adalah bagaimana rerata dari ke-4 kasus meningkat sesuai dengan tingkat kerusakan struktur. Terlihat bahwa nilai T_k meningkat dari sekitar 1 hingga 2 berubah menjadi 70 hingga diatas 100. Hal ini sejalan dengan teori probabilitas, bahwa setiap data memiliki distribusinya masing masing. Pada saat data yang mengalami kerusakan di prediksi dengan model yang dibangun dengan data nominal, maka residu yang dihasilkan akan tinggi, sehingga nilai standar deviasinya berubah secara signifikan. Hal ini menunjukkan bahwa parameter T_k dapat digunakan sebagai suatu parameter yang mengkarakterisasi tingkat kerusakan dari sebuah struktur. Semakin tinggi nilai T_k , semakin fatal tingkat kerusakan yang terdapat didalam struktur.

Meskipun berada pada struktur yang sama, setiap akselerometer akan memberikan respon yang berbeda satu sama lain. Hal ini disebabkan oleh perubahan dinamika yang terbaca oleh akselerometer, relatif terhadap posisi kerusakan yang terjadi. Berdasarkan Tabel 2, terlihat bahwa akselerometer yang paling sensitif akan deformasi lateral ditunjukkan oleh akselerometer 3, 9 dan 22, dengan nilai T_k yang berada diatas 100. Dari Gambar 2, diketahui bahwa akselerometer 3, 9 dan 22 adalah akselerometer yang terletak di bagian fondasi dari tiang jembatan. Tekanan yang diberikan oleh aktuator mengakibatkan deformasi lateral pada tiang, yang mengakibatkan perubahan pada arah longitudinal. Akibat adanya fondasi yang menopang yang bersifat kaku, maka perubahan tekanan maksimum akan dirasakan pada bagian dasar dari tiang [29]. Oleh karena itu, pembacaan akselerometer di dekat dasar tiang mengalami fluktuasi tertinggi relatif terhadap model nominal.

Informasi berguna terakhir adalah terkait dengan posisi pemasangan sensor pada struktur. Algoritma korelasi jarak dan entropi menawarkan pemilihan sub himpunan data yang menjanjikan bahwa setiap anggota sub himpunan memiliki hubungan linear atau non linear yang baik terhadap anggota lain. Dalam penerapan riilnya, sub himpunan ini dapat diterjemahkan sebagai posisi dimana sekelompok akselerometer atau sensor harus dipasangkan di dalam sebuah struktur, untuk menjamin kinerja dari mereka berjalan dengan baik. Pada tahap awal, sensor dapat dipasangkan di seluruh arah. Kemudian melalui pengolahan data sesuai dengan alur yang dijelaskan pada Bab III, dapat diperoleh sekelompok sensor yang memiliki korelasi yang baik antara satu sama lain. Kemudian sensor lain dapat dilepas, menyisakan sensor yang memiliki korelasi baik saja. Dengan cara ini, maka biaya pemeliharaan dari sensor dapat ditekan, tanpa harus menumbalkan kualitas prediksi dari akselerometer tersebut.

V. KESIMPULAN

Pada penelitian ini, telah digagas sebuah algoritma baru berbasis pada metode kuadrat terkecil parsial atau KTP dan korelasi jarak untuk mendeteksi kerusakan struktur bangunan. Algoritma yang ditawarkan terbukti mampu mendeteksi kerusakan struktur dengan baik, serta mampu mengkarakterisasi tingkat kerusakan yang ada. Lebih lanjut, algoritma yang ditawarkan dapat memodelkan dinamika akselerometer menggunakan variabel 80% lebih sedikit dari metode kuadrat terkecil parsial standar, tanpa mengurangi tingkat akurasi atau bahkan meningkatkan akurasi model prediktif sekitar 1-1,3%.

Terkait dengan arah penelitian di masa mendatang, dirasa akan menarik untuk mempelajari secara mendalam, bagaimana parameter pendektaksian ini dapat dikembangkan. Sebagai contoh, dengan mengembangkan lebih lanjut secara teoritis karakterisasi tingkat kerusakan dari parameter T_k dan bagaimana melakukan pemilihan yang efisien terhadap k sehingga dapat dengan optimal menghindari kasus negatif palsu dan positif palsu sekaligus. Diharapkan pula algoritma ini dapat diterapkan di berbagai struktur buatan manusia, agar dapat menjadi sebuah *alarm*, untuk menghindari kerusakan struktur secara masif yang dapat menimbulkan kerugian secara materiil ataupun merenggut korban jiwa.

DAFTAR PUSTAKA

- [1] V. E. Satya, "Strategi Indonesia menghadapi industri 4.0," *Info Singkat*, vol. 10, no. 9, pp. 19-24, 2018.
- [2] D. A. Tibaduiza Burgos, R. C. Gomez Vargas, C. Pedraza, D. Agis and F. Pozo, "Damage identification in structural health monitoring: A brief review from its implementation to the use of data-driven applications," *Sensors*, vol. 20, no. 3, p. 733, 2020.
- [3] M. Azimi, A. D. Eslamlou and G. Pekcan, "Data-driven structural health monitoring and damage detection through deep learning: State-of-the-art review," *Sensors*, vol. 20, no. 10, p. 2778, 2020.
- [4] H. V. Dang, H. Tran-Ngoc, T. V. Nguyen, T. Bui-Tien, G. De Roeck and H. X. Nguyen, "Data-driven structural health monitoring using feature fusion and hybrid deep learning," *IEEE Transactions on Automation Science and Engineering*, vol. 18, no. 4, pp. 2087-2103, 2020.
- [5] F. Smarra, G. D. D. Girolamo, V. Gattulli, F. Graziosi and A. D'Innocenzo, "Learning models for seismic-induced vibrations optimal control in structures via random forests," *Journal of Optimization Theory and Applications*, vol. 187, no. 3, pp. 855-874, 2020.
- [6] H.-M. Lin, M.-H. Lee, J.-C. Liang, H.-Y. Chang, P. Huang and C.-C. Tsai, "A review of using partial least square structural equation modeling in e-learning research," *British Journal of Educational Technology*, vol. 51, no. 4, pp. 1354-1372, 2020.
- [7] L. Wang, "Enhanced fault detection for nonlinear processes using modified kernel partial least squares and the statistical local approach," *The Canadian Journal of Chemical Engineering*, vol. 96, no. 5, pp. 1116-1126, 2018.
- [8] H. Abdi, "Partial least square regression (PLS regression)," *Encyclopedia for research methods for the social sciences*, vol. 6, no. 4, pp. 792-795, 2003.
- [9] M. Ruiz, L. Mujica, X. Berjaga and J. Rodellar, "Partial least square projection to latent structures (PLS) regression to estimate impact localization in structures," *Smart materials and structures*, vol. 22, no. 2, p. 025028, 2013.
- [10] X. Cheng, Q. Li, W. Zhou and Z. Zhou, "External deformation monitoring and improved partial least squares data analysis methods of high core rock-fill dam (HCRFD)," *Sensors*, vol. 20, no. 2, p. 444, 2020.
- [11] F. Kang, J. Liu, J. Li and S. Li, "Concrete dam deformation prediction model for health monitoring based on extreme learning machine," *Structural Control and Health Monitoring*, vol. 24, no. 10, p. e1997, 2017.
- [12] H. Liu, J. Yang, Y. Zhang and C. Yang, "Monitoring of wastewater treatment processes using dynamic concurrent kernel partial least squares," *Process Safety and Environmental Protection*, vol. 147, pp. 274-282, 2021.
- [13] J. Tjen, F. Smarra and A. D'Innocenzo, "An entropy-based sensor selection algorithm for structural damage detection," in *IEEE 16th International Conference on Automation Science and Engineering (CASE)*, Online Virtual Meeting, 2020.
- [14] F. Smarra, J. Tjen and A. D'Innocenzo, "Learning methods for structural damage detection via entropy-based sensors selection," *International Journal of Robust and Nonlinear Control*, vol. 32, no. 10, pp. 6035-6067, 2022.
- [15] S. Karamizadeh, S. M. Abdullah, A. A. Manaf, M. Zamani and A. Hooman, "An overview of principal component analysis," *Journal of Signal and Information Processing*, vol. 4, 2020.
- [16] B. M. S. Hasan and A. M. Abdulazeez, "A review of principal component analysis algorithm for dimensionality reduction," *Journal of Soft Computing and Data Mining*, vol. 2, no. 1, pp. 20-30, 2021.
- [17] H. Abdi and L. J. Williams, "Partial least squares methods: partial least squares correlation and partial least square regression," in *Computational toxicology*, Springer, 2013, pp. 549-579.
- [18] X.-Q. Zeng and G.-Z. Li, "Incremental partial least squares analysis of big streaming data," *Pattern recognition*, vol. 47, no. 11, pp. 3726-3735, 2014.
- [19] A. Jordao, M. Lie, V. H. C. de Melo and W. R. Schwartz, "method, Covariance-free partial least squares: An incremental dimensionality reduction," *Proceedings of the IEEE/CVF Winter Conference on Applications of Computer Vision*, pp. 1421-1429, 2021.
- [20] G. E. Crooks, "On measures of entropy and information," *Tech. Note*, vol. 9, no. 4, 2017.
- [21] W. H. Zurek, *Complexity, entropy and the physics of information*, Boca Raton: CRC Press, 2018.
- [22] D. Zhang, "A coefficient of determination for generalized linear models," *The American Statistician*, vol. 71, no. 4, pp. 310-316, 2017.
- [23] C. Fu, P. Wang, L. Zhao and X. Wang, "A distance correlation-based Kriging modeling method for high-dimensional problems," *Knowledge-Based Systems*, vol. 206, p. 106356, 2020.
- [24] D. Edelmann, K. Fokianos and M. Pitsillou, "An updated literature review of distance correlation and its applications to time series," *International Statistical Review*, vol. 87, no. 2, pp. 237-262, 2019.
- [25] R. Estrin, D. Orban and M. Saunders, "Euclidean-norm error bounds for SYMMLQ and CG," *SIAM Journal on Matrix Analysis and Applications*, vol. 40, no. 1, pp. 235-253, 2019.
- [26] A. Chaudhuri and W. Hu, "A fast algorithm for computing distance correlation," *Computational statistics & data analysis*, vol. 135, no. 2, pp. 15-24, 2019.
- [27] S. J. Novick, W. Zhao and H. Yang, "Setting alert and action limits in the presence of significant amount of censoring in data," *PDA Journal of Pharmaceutical Science and Technology*, vol. 71, no. 1, pp. 20-32, 2017.
- [28] Electrical engineering og Los Alamos National Laboratory (LANL). [Online]. Available: <https://www.lanl.gov/projects/national-security-education-center/engineering/software/shm-data-sets-and-software.php>. [Accessed 5 12 2021].

- [29] S. Basack, B. Indraratna, C. Rujikiatkamjorn and F. Siahaan, "Modeling the stone column behavior in soft ground with special emphasis on lateral deformation," *Journal of Geotechnical and Geoenvironmental Engineering*, vol. 143, no. 6, p. 4017016, 2017.