APLIKASI TEKNIK *PRINCIPAL COMPONENT ANALYSIS* (PCA) UNTUK PENYEPAKATAN DATA PADA KOMPRESOR MULTI TAHAP

Wielianto Yudha, Tri Partono Adhi

Program Studi Teknik Kimia, Fakultas Teknologi Industri, ITB

Labtek X, Jl. Ganesha 10 Bandung – 40132

E-mail : Wieliano@Yahoo.com

Abstrak

Dalam perkembangannya, deteksi kesalahan nyata dilakukan dengan teknik statistik multivariabel untuk mendapatkan hasil yang lebih akurat dan benar dalam penentuan variabel yang menyebabkan kesalahan nyata tersebut. Teknik statistik multivariabel dikenal handal dalam mengatasi permasalahan dimana terdapat korelasi di antara variabel proses yang diukur. PCA sebagai salah satu teknik analisis statistik kuantitatif multivariabel dikenal baik untuk mendeteksi kesalahan nyata pada pengukuran. Penelitian ini dilaksanakan dengan tujuan untuk menerapkan PCA dalam deteksi dan identifikasi kesalahan nyata pada data pengukuran, menerapkan teknik penyepakatan data untuk memperoleh data pengukuran yang tersepakati, dan mengkaji penerapan PCA pada penyepakatan data secara dinamik. Dalam penelitian ini, operasi kompresor multi tahap dalam pabrik pupuk urea dipilih sebagai studi kasus. Hasil dari penelitian ini membuktikan bahwa teknik PCA dapat diterapkan untuk deteksi dan identifikasi kesalahan nyata pada data pengukuran laju alir, tekanan, dan temperatur. Aplikasi teknik PCA pada penyepakatan data keadaan tunak dapat digunakan untuk mendeteksi letak kesalahan nyata pada data pengukuran sehingga eliminasi atau koreksi terhadap data yang mengandung kesalahan nyata tersebut dapat dilakukan dengan tepat. Selain itu, teknik PCA dapat digunakan untuk mendeteksi dan melokalisir kesalahan dengan cepat sehingga teknik ini memungkinkan untuk diterapkan pada penyepakatan data dinamik. Akan tetapi, aplikasi teknik PCA ini perlu diawali dengan pengumpulan pengetahuan proses yang lengkap sehingga tingkat keyakinan terhadap keputusan penerimaan atau penolakan suatu data lebih tinggi.

Kata Kunci : Data Pengukuran, Kesalahan Nyata, Penyepakatan Data, PCA

Abstract

The latest development of gross error detection has been focused on the use of multivariate statistical techniques. These techniques give more accurate and correct detection in determining the variables suspected to contain gross error. These techniques are reliable to solve the problem of correlation present in measurement data. As one of multivariate statistical technique, PCA can be used to detect gross error in the measurement. The objectives of this research are to apply PCA for detection and identification of gross error in measurement data, to apply data reconciliation technique to obtain reconciled data, and to investigate application of PCA leading to its use in the dynamic data reconciliation. In this research, the operation of multistage compressor in urea fertilizer plant is chosen as the studied case. The results of this research prove that PCA can be applied for detection and identification of gross error in measurement data of mass flow, pressure, and temperature. PCA application on steady state data reconciliation can be used for detection the location of gross error in the measurement data. Gross error in the data can be eliminated and corrected exactly. Moreover, PCA can be used to detect and isolate fault quickly leading to its use in the dynamic data reconciliation. Nevertheless, process knowledge base must be collected first before the application of PCA. If the process knowledge is more complete, the confidence of the inference becomes higher.

Keywords: Measurement Data, Gross Error, Data Reconciliation, PCA

1. Pendahuluan

Dalam industri kimia, gangguan terhadap proses yang berlangsung di dalam pabrik merupakan hal yang perlu diperhatikan secara serius. Gangguan terhadap proses dapat menyebabkan satu atau beberapa unit operasi tidak dapat memenuhi fungsi spesifiknya di dalam proses. Kondisi operasi di dalam proses dapat berubah jika gangguan tersebut tidak segera diatasi. Perubahan kondisi operasi dalam proses kimia sangat tidak dikehendaki karena dapat menyebabkan kegagalan proses. Kegagalan proses dapat menyebabkan kapasitas produksi dan kualitas produk yang diperoleh tidak sesuai dengan target. Jika hal ini terus berlanjut, maka penghentian operasi beberapa unit operasi bahkan seluruh unit operasi akan terjadi. Hal ini sering disebut sebagai *unscheduled shutdown* (USD) yang menimbulkan banyak kerugian. Untuk mencegah beberapa hal tersebut, teknik atau metode diagnosis gangguan dan kesalahan selama proses menjadi sangat diperlukan.

Keberadaan gangguan dalam proses dapat diamati dari data pengukuran variabel proses. Tentu saja, data pengukuran yang benar dan akurat sangat diperlukan dalam pengamatan kemungkinan terjadinya gangguan pada proses. Data pengukuran merupakan bahan dasar dari metode diagnosis gangguan dan kesalahan selama proses. Kemampuan deteksi dan diagnosis gangguan dan kesalahan dalam proses sangat bergantung pada data yang diperoleh.

Secara umum, proses dalam industri kimia memiliki data yang berjumlah banyak dan saling berkorelasi. Data tersebut diperoleh dari setiap unit operasi untuk setiap variabel proses yang terdapat dalam pabrik. Banyak penelitian telah dilakukan untuk menemukan dan menerapkan metode deteksi dan diagnosis gangguan dan kesalahan. Penelitian-penelitian tersebut membuktikan metode analisis statistik multivariabel mampu digunakan untuk mengolah data yang berjumlah banyak dan berkorelasi (Amand, Heyen, dan Kalitventzeff, 2001; Rotem, Wachs, dan Lewin, 1999; Venkatasubramanian, Rengaswamy, Kavuri, dan Yin, 2003; Klančar, 1998; Narasimhan dan Shah, 2004).

Salah satu metode analisis statistik multivariabel yang dapat digunakan untuk deteksi dan diagnosis gangguan dan kesalahan pada proses adalah PCA (*Principal Component Analysis*). Secara khusus, metode ini baik untuk deteksi kesalahan nyata (*gross error*) pada data pengukuran, diagnosis variabel proses yang memberikan pengaruh terhadap kesalahan, dan menemukan penyebab kesalahan pada data pengukuran. Pengolahan data pengukuran variabel

proses dengan PCA merupakan hal yang menarik untuk dilakukan karena dapat memberikan informasi berharga bagi suatu proses.

Tuntutan kehandalan (*reliability*) proses dan ketangguhan (*robustness*) sistem diagnosis pada proses membutuhkan data pengukuran variabel proses yang berjumlah banyak. Data pengukuran tersebut diantaranya tersedia dari sistem pengukuran yang berlebih (*redundant*). Keberadaan data pengukuran yang berlebih cenderung mengakibatkan ketidaksesuaian di antara data tersebut sehingga dapat mengantarkan pada pemahaman yang tidak sempurna terhadap informasi yang dikandungnya (*misleading*). Oleh karena itu, diperlukan metode untuk mengolah dan menganalisis data sehingga diperoleh data yang disepakati dan hanya mengandung kesalahan acak (*random error*).

Data pengukuran variabel proses selalu mengandung kesalahan. Kesalahan dalam data pengukuran dapat berupa kesalahan nyata (*gross error*) dan kesalahan acak (*random error*). Kesalahan nyata dapat disebabkan oleh gangguan proses, kebocoran proses, dan kegagalan pemakaian atau kesalahan kalibrasi instrumen. Deteksi kesalahan nyata pada data pengukuran variabel proses perlu dilakukan untuk mengetahui gangguan proses itu terjadi karena masalah proses atau kegagalan fungsi sistem pengukuran.

Aktivitas penyepakatan data digunakan untuk memproses data pengukuran yang mengandung kesalahan (acak) menjadi informasi yang konsisten. Tujuan penyepakatan data adalah menyelesaikan permasalahan kontradiksi antara pengukuran dan batasan operasionalnya. Data pengukuran yang tersepakati merupakan data yang diperlukan dalam pemantauan dan pengendalian sistem proses. Aktivitas penyepakatan data dapat dilakukan dengan baik jika kesalahan nyata tidak terdapat dalam data pengukuran. Kesalahan nyata perlu dideteksi dan kemudian dieliminasi atau dikoreksi sebelum diperoleh data tersepakati yang benar.

Selanjutnya, berkaitan dengan statistika inferensial, uji terhadap kesalahan nyata secara konvensional, seperti uji *chi-square* pada pengolahan data pengukuran untuk mendapatkan selisih kuadrat terkecil maupun uji pada nilai residual dari beberapa batasan (*constraints*), tidak sepenuhnya handal dalam mengidentifikasi kesalahan nyata dengan benar. Hal ini muncul karena adanya korelasi dalam data pengukuran dan di antara hasil pengaturan dan residual. PCA sebagai salah satu metode analisis statistik kuantitatif multivariabel dapat digunakan untuk mendeteksi kesalahan nyata pada pengukuran. Selain itu, metode ini juga dapat digunakan untuk

mendeteksi dan mendiagnosis gangguan dan kesalahan pada proses.

Penelitian ini dilaksanakan dengan tujuan untuk menerapkan PCA dalam deteksi dan identifikasi kesalahan nyata pada data pengukuran, menerapkan teknik penyepakatan data untuk memperoleh data pengukuran yang tersepakati, dan mengkaji penerapan PCA pada penyepakatan data secara dinamik. Dalam penelitian ini, operasi kompresor multi tahap dalam pabrik pupuk urea dipilih sebagai studi kasus.

2. Fundamental

Aktivitas pengukuran variabel-variabel proses dilakukan pada setiap proses kimia. Pada hakikatnya, semua kuantitas yang diperoleh dari pengukuran mengandung kesalahan dan besarnya tidak pasti. Oleh karena itu, permasalahan kesalahan tersebut merupakan bagian penting dari perlakuan dan interpretasi terhadap data hasil pengukuran. Kesalahan pada pengukuran adalah kesalahan yang terkandung di dalam data pengukuran yang dapat berupa kesalahan acak (*random error*), kesalahan sistematik (*systematic error*), dan kesalahan nyata (*gross error*). Kesalahan acak merupakan kesalahan yang besarnya bervariasi dan memiliki tanda yang berbeda-beda. Kesalahan ini memiliki nilai yang berosilasi di sekitar nol dan rata-ratanya nol. Masing-masing kesalahan acak tidak memperlihatkan adanya keteraturan dan saling independen. Nilai tiap kesalahan acak tidak dapat diprediksi dan tidak dapat dijelaskan (Madron dan Sharp, 1992). Kesalahan sistematik adalah kesalahan yang dapat ditandai oleh adanya hasil pengukuran yang memiliki kecenderungan selalu lebih kecil ataupun selalu lebih besar dari nilai sejatinya. Kesalahan sistematik memberikan informasi adanya kesalahan yang konstan terhadap waktu atau kesalahan yang kelakuannya dapat diperkirakan (Madron dan Sharp, 1992). Kesalahan ini dapat dijumpai pada kasus seperti: pengaturan instrumen yang tidak sempurna/baik (memiliki kesalahan konstan), kesalahan yang memiliki hubungan linier terhadap waktu akibat pergeseran nilai nol pada instrumen, kesalahan dengan variasi periodik mengikuti temperatur ambien harian, dan sebagainya. Kesalahan nyata merupakan kesalahan yang dapat muncul secara tiba-tiba dan tidak disadari. Kesalahan nyata dapat disebabkan oleh gangguan proses, kebocoran proses, dan kegagalan pemakaian atau kesalahan kalibrasi instrumen (Madron dan Sharp, 1992).

Teori yang membahas kesalahan pengukuran dibagi menjadi teori umum dan teori khusus (Madron dan Sharp, 1992). Penerapan teori kesalahan diawali dengan diperkenalkannya

istilah kesalahan absolut. Kesalahan absolut pada pengukuran e (untuk selanjutnya hanya berkenaan dengan kesalahan pada pengukuran) didefinisikan sebagai:

$$x^+ = \hat{x} + e \quad (1)$$

dimana x^+ adalah nilai hasil pengukuran, $\hat{x}$ adalah nilai sejati pengukuran, dan e adalah kesalahan pada nilai hasil pengukuran.

Principal Component Analysis (PCA) adalah salah satu teknik penyelesaian multivariabel dimana sejumlah variabel terkorelasi dapat ditransformasikan menjadi himpunan variabel tak terkorelasi yang berdimensi lebih kecil. Teknik statistik multivariabel ini dikenal sebagai alat yang tangguh untuk meringkas data dan mereduksi dimensinya sehingga dapat diperoleh dan dipertahankan informasi pokoknya. Dengan diperolehnya informasi pokok tersebut, analisis terhadap data akan menjadi lebih mudah dibandingkan analisis terhadap himpunan data awal yang masih berukuran besar. Teknik ini juga dapat digunakan secara efektif untuk mengekstrak informasi yang benar dari himpunan data yang diperoleh dari proses yang mengalami gangguan (Venkatasubramanian, 2003).

Metode analisis berbasis sejarah proses biasanya dilakukan setelah terjadi suatu kegagalan pada operasi pabrik. Hal ini bertujuan untuk mengidentifikasi penyebab kesalahan dan dampak lanjutan dari kesalahan yang ada sehingga dapat ditemukan usulan tindakan penanganan yang tepat supaya kegagalan yang sama dapat dicegah dan dihindari pada masa mendatang. Selain itu, analisis kesalahan dalam tahap perancangan selanjutnya, dilakukan dengan tujuan mengidentifikasi kemungkinan terjadinya kesalahan dan dampak yang ditimbulkan kesalahan tersebut sehingga dapat disusun suatu prosedur operasi dan konfigurasi pengendalian yang mampu menghindari kesalahan tersebut.

Salah satu dampak yang ditimbulkan dari kesalahan yang ada dalam proses adalah variabilitas. Variabilitas merupakan hal yang tidak dikehendaki di dalam proses yang menargetkan kualitas produk. Beberapa proses kimia dipengaruhi oleh variabilitas, tetapi variabel prosesnya tidak berfluktuasi secara acak sepenuhnya. Variabel proses ini dihubungkan oleh satu himpunan pembatas (neraca massa dan energi, kebijakan pengoperasian) yang dapat disusun dalam suatu model proses. Analisis secara statistik dari data hasil pengukuran variabel proses terhadap waktu dapat menyatakan korelasi yang mendasari hubungan antar variabel terukur, jika model matematika yang tepat tidak tersedia.

Tindakan koreksi dapat dilakukan lebih awal jika kesalahan diidentifikasi dan dilokalisasi dengan cepat. Metode analisis PCA memenuhi kedua persyaratan tersebut. Dengan kemampuan deteksi dan diagnosis yang akurat secara dini dan kemampuan membedakan antara satu kesalahan dengan kesalahan lain, PCA telah digunakan untuk mendeteksi dan mendiagnosis kesalahan pada beberapa kasus, seperti: unit operasi reaktor (Shah dkk., 1998), kompresor (Rotem dkk., 1999), sistem lingkaran sintesis amonia (Amand dkk., 2001), dan pabrik fraksinasi *hydrocracker* (Tong dan Crowe, 1996). Selain itu, PCA juga memiliki kemampuan untuk mengakomodasi gangguan (*noise*) dan ketidakpastian (*uncertainties*) yang dapat mengakibatkan kegagalan diagnosis secara tiba-tiba, mengenali dan mengelompokkan kesalahan yang tidak diketahui penyebabnya (*unknown faults, novel faults*).

Analisis statistik menggunakan PCA pada dasarnya adalah dekomposisi ortogonal suatu matriks data proses X (matriks $j \times k$, dimana j adalah jumlah sampel dan k adalah jumlah variabel terukur) ke dalam komponen-komponen utamanya (*principal components, PC*). Tujuan dari analisis statistik ini adalah untuk menemukan faktor yang memberikan pengaruh paling besar terhadap pola dari data proses yang terukur (Venkatasubramanian, 2003). Prinsip dekomposisi suatu matriks data pengukuran dalam PCA disajikan pada persamaan (2).

$$X = \bar{X} + E = \sum_{i=1}^k t_i p_i^T + E = TP^T + E \quad (2)$$

Matriks residual E menyatakan perbedaan antara nilai data terukur (X) terhadap nilai data sejatinya ($\bar{X}$). Konstanta k menyatakan jumlah dari komponen utama yang mempengaruhi data terukur. Vektor t_i adalah nilai dari komponen-komponen utama (*pc*) data original. Nilai ini sering kali disebut sebagai *principal component score*. Vektor p_i adalah koefisien kombinasi linier antara variabel original dengan komponen utama (*pc*). Koefisien ini sering kali disebut sebagai *principal component loading vectors*.

Jika matriks U menyatakan hasil dekomposisi ortonormal matriks kovarian S menjadi matriks diagonal L mengikuti aturan aljabar matriks, maka hubungan ini dapat dituliskan sebagai berikut:

$$S = ULU^T \quad (3)$$

Matriks diagonal L beranggotakan elemen $l_1, l_2, \dots, l_k$, yang merupakan nilai *eigen* (*eigenvalue*) dari matriks kovarian S . Hubungan di atas sering digunakan untuk transformasi dasar PCA (Jackson, 1991).

Penyepakatan data merupakan aktivitas penting yang perlu dilakukan dalam industri proses yang berlangsung kontinyu (Venkatasubramanian, 2003). Pada dasarnya, penyepakatan data adalah pendekatan diagnosis kesalahan secara kuantitatif yang terfokus pada deteksi kesalahan sensor dan bias sensor. Tujuan utama aktivitas ini adalah menyepakatkan data pengukuran. Data pengukuran yang tersepakati merupakan data yang konsisten dan memiliki presisi lebih tinggi daripada data pengukuran yang diperoleh langsung.

Penyepakatan data biasanya terdiri atas tiga bagian, yaitu: identifikasi bias pada parameter, estimasi bias, dan memperbaiki pengukuran pada sensor (Venkatasubramanian, 2003). Lebih lanjut, penyepakatan data memberikan informasi tentang kemungkinan keberadaan kesalahan nyata, keakurasian hasil pengukuran, dan perambatan kesalahan pengukuran. Pada prinsipnya, penyepakatan data merupakan permasalahan optimasi multidimensional dan multivariabel baik linier maupun non-linier.

Aktivitas penyepakatan data dapat dilakukan pada kondisi tunak (*steady-state*) maupun tak tunak (*transient*). Penyepakatan data untuk keadaan tunak merupakan permasalahan penghilangan kesalahan (*error*) yang diberikan oleh sekumpulan data dari variabel-variabel sensor (Venkatasubramanian, 2003). Hal ini biasanya dilakukan dengan teknik estimasi *least squares* linier dan telah dibuktikan dapat memberikan hasil yang dapat dipercaya. Di lain sisi, penyepakatan data dinamik merupakan permasalahan penghilangan kesalahan dari variabel-variabel sensor di dalam suatu rentang waktu. Penyepakatan data dinamik merupakan permasalahan yang secara signifikan lebih sukar dibandingkan penyepakatan data untuk keadaan tunak. Ini disebabkan oleh kehadiran batasan/kendala diferensial dan efek yang nyata dari nonlinieritas pada strategi penyelesaiannya.

3. Metodologi

Penelitian mengenai aplikasi teknik *principal component analysis* (PCA) dalam penyepakatan data pada kompresor multi tahap ini dilaksanakan dengan menggunakan perangkat sebagai berikut: komputer personal sebanyak 1 buah (yang digunakan untuk pembuatan modul PCA, pembuatan model simulasi, pengolahan data pengukuran, penyusunan strategi pengolahan data pengukuran dengan PCA, penerapan teknik penyepakatan data pengukuran, dan simulasi untuk menguji strategi penerapan PCA yang telah disusun) dan perangkat lunak (aplikasi *Microsoft Office*, aplikasi pemrograman *Visual Basic* pada

Microsoft Excel, dan simulator proses Aspen-Hysys 2004). Penelitian ini diawali dengan pembuatan modul PCA dan model simulasi yang akan digunakan pada tahap pengolahan data pengukuran dengan PCA, penerapan teknik penyepakatan data pada data pengukuran, dan pengkajian penerapan PCA untuk penyepakatan data dinamik. Data yang akan diolah diperoleh dari model simulasi yang telah dibuat. Data tersebut dianggap sebagai data hasil pengukuran.

Pengolahan data pengukuran dilakukan untuk data yang diperoleh dari model simulasi sistem kompresi gas sintesis multi tahap. Model simulasi tersebut dibangun berbantuan simulator proses Aspen-Hysys 2004. Data yang akan diolah dengan menggunakan prosedur PCA diperoleh dari model simulasi yang dioperasikan secara dinamik. Alat ukur yang digunakan dianggap bekerja dengan baik dan selalu dikalibrasi. Himpunan data pertama tersusun atas 1000 data laju alir massa (kg/jam) semua aliran yang terdapat dalam model. Himpunan data kedua tersusun atas 1000 data untuk masing-masing variabel tekanan (kPa) dan temperatur (K) di setiap aliran masukan dan keluaran kompresor. Karena data yang diperoleh dari simulator Aspen-Hysys merupakan data yang tersepakati, maka dilakukan penambahan kesalahan acak dalam rentang simpangan baku alat ukurnya pada data dengan subrutin randomisasi berdasarkan distribusi normal. Data tersebut selanjutnya dianggap sebagai data hasil pengukuran pada kondisi operasi normal. Data pada kondisi operasi normal tersebut diuji dengan menggunakan statistika deskriptif untuk mengetahui distribusinya.

Prosedur pengolahan data dengan PCA pada umumnya digunakan untuk mendeteksi adanya kesalahan nyata di suatu himpunan data. Pengolahan data laju alir dengan PCA untuk penyepakatan data keadaan tunak diawali dengan penyusunan persamaan kendala berdasarkan kaidah neraca massa, sedangkan untuk data tekanan dan temperatur persamaan kendala disusun berdasarkan hubungan linier antara selisih tekanan (keluaran dan masukan) dan selisih temperatur (keluaran dan masukan). Konstanta untuk hubungan linier tersebut diperoleh dari rata-rata data pengukuran yang hanya mengandung kesalahan acak yang telah dikumpulkan sebelumnya. Konstanta yang diperoleh tersebut dapat digunakan dalam persamaan linier yang menghubungkan variabel tekanan dan temperatur. Persamaan linier tersebut dapat digunakan untuk menggantikan persamaan yang umum dipakai untuk sistem kompresi dalam menghubungkan variabel tekanan dan temperatur berdasarkan asumsi proses kompresi berlangsung secara isentropik atau politropik..

Setelah persamaan kendala disusun, matriks penyeimbang dapat disusun berdasarkan persamaan kendala tersebut. Matriks tersebut dan matriks kovarian data digunakan untuk memperoleh matriks kovarian residual yang akan didekomposisi. Dekomposisi ortonormal dilakukan pada matriks kovarian residual sehingga diperoleh vektor komponen utama.

Komponen utama dan kontribusi residual persamaan kendala terhadap komponen utama dihitung untuk data pengukuran laju alir. Data pengukuran yang digunakan berupa data hasil rata-rata dari himpunan data yang mengandung kesalahan acak, data pengukuran yang diambil secara acak dari himpunan data yang mengandung kesalahan acak, dan data pengukuran yang mengandung kesalahan nyata. Selanjutnya dilakukan plot data untuk elemen-elemen vektor komponen utama dari matriks kovarian residual persamaan kendala dan kontribusi residual persamaan kendala terhadap masing-masing elemen vektor komponen utama. Selain itu, komponen utama dan kontribusi residual persamaan kendala terhadap komponen utama dihitung untuk data pengukuran tekanan dan temperatur. Data pengukuran yang digunakan berupa data pengukuran yang diambil secara acak dari himpunan data yang mengandung kesalahan acak dan data pengukuran yang mengandung kesalahan nyata. Elemen-elemen vektor komponen utama dan kontribusi residual persamaan kendala terhadap masing-masing elemen vektor komponen utama yang diperoleh selanjutnya diplot.

Prosedur PCA yang diterapkan pada data pengukuran terdiri atas beberapa langkah. Langkah-langkah teknik PCA tersebut, yaitu: penyusunan persamaan kendala, penyusunan matriks residual dari persamaan kendala menggunakan persamaan berikut:

$$e = Bx^+ \quad (4)$$

dimana e adalah residual dari batasan tereduksi untuk proses keadaan tunak linier, B adalah matriks penyeimbang, dan x^+ adalah nilai hasil pengukuran, penyusunan matriks kovarian dengan persamaan berikut:

$$\text{cov}(e) \equiv H_e = BSB^T \quad (5)$$

dimana H_e adalah matriks kovarian residual dan S adalah matriks kovarian data, dekomposisi matriks kovarian yang diperoleh menjadi matriks *eigenvalue* (L_e) dan *eigenvector* (U_e) yang memenuhi hubungan berikut:

$$L_e = U_e^T H_e U_e \quad (6)$$

penghitungan koefisien kombinasi linier (W_e) dari suatu persamaan kombinasi linier untuk residual

berdasarkan hubungan berikut:

$$W_e = U_e L_e^{1/2} \quad (7)$$

penghitungan komponen utama (y_e) dari residual berdasarkan persamaan berikut:

$$y_e = W_e^T e \quad e^* = W_e^T e \quad (8)$$

pemeriksaan kontribusi dari residual ke-j (g_j) terhadap komponen utama yang dicurigai dengan menggunakan persamaan:

$$g_j = w_{e,j} e_j \quad (9)$$

Setelah teknik PCA diterapkan pada data pengukuran, kemudian dilakukan penyepakatan data terhadap data pengukuran yang diperoleh dari model simulasi sistem kompresi gas sintesis multi tahap. Penyepakatan data dilakukan untuk memperoleh data pengukuran yang tersepakati dengan cara meminimumkan fungsi obyektif dan memenuhi persyaratan kendala yang disusun berdasarkan neraca massa maupun neraca energi (atau persamaan yang menghubungkan variabel tekanan dan temperatur). Peminimuman fungsi obyektif ini dilakukan dengan menggunakan *Solver* yang ada pada perangkat lunak *Ms. Excel* atau dengan algoritma *Statistical Data Fit* (SDF). Algoritma SDF dilakukan dengan menggunakan persamaan berikut.

$$\hat{x} = x^+ K f x_0 + F_x x^+ x_0 \quad (10)$$

$$K = R_x F_x^T F_x R_x F_x^T^{-1} \quad (11)$$

dimana $\hat{x}$ adalah nilai sejati pengukuran, x^+ adalah nilai hasil pengukuran, x_0 adalah nilai pengukuran pembuat nol fungsi atau persamaan kendala, $f(x_0)$ adalah fungsi yang akan dinolkan atau persamaan kendala, F_x adalah matriks penyeimbang atau matriks *Jacobian*, dan R_x adalah matriks kovarian data.

Penyepakatan data dilakukan baik untuk data pengukuran yang hanya mengandung kesalahan acak maupun data pengukuran yang mengandung kesalahan nyata. Hal ini dilakukan untuk mengetahui seberapa penting tahap deteksi dan eliminasi data yang mengandung kesalahan nyata dilakukan.

Pengkajian penerapan PCA untuk penyepakatan data dinamik diawali dengan pembangunan model simulasi untuk menggambarkan proses yang terjadi pada unit operasi kompresor multi tahap dengan menggunakan perangkat lunak *Aspen-Hysys* 2004. Data pengukuran laju alir massa, tekanan,

dan temperatur disiapkan untuk menguji penerapan PCA tersebut. Data pengukuran tersebut diperoleh dari model simulasi sistem kompresi gas sintesis multi tahap. Setiap data pengukuran yang diperoleh langsung dianalisis menggunakan PCA dengan melakukan pengamatan terhadap skor komponen-komponen utama dari residual persamaan kendala yang telah disusun. Selanjutnya, teknik penyepakatan data dilakukan apabila pada data pengukuran tersebut tidak mengandung kesalahan nyata.

4. Hasil dan Analisis

Sistem kompresi multi tahap yang dimodelkan dalam simulasi ini merupakan proses kompresi empat tahap untuk gas sintesis yang akan digunakan sebagai umpan konverter amonia pada pabrik pupuk. Aliran gas sintesis dikompresi dua tahap sebelum digabungkan dengan aliran gas (yang mengandung amonia) yang diumpanbalikkan dari kompresor tahap keempat dan konverter amonia. Pendinginan dilakukan di antara tahap kompresi yang berguna untuk meningkatkan efisiensi kompresor tahap selanjutnya. Terdapat 48 aliran dalam model simulasi ini yang akan digunakan sebagai sumber data pengukuran laju alir massa.

Data yang akan diolah dengan menggunakan prosedur PCA diperoleh dari model simulasi sistem kompresi gas sintesis multi tahap yang dioperasikan secara dinamik. Himpunan data laju alir proses tersusun atas 1000 data laju alir massa (kg/jam) semua aliran yang terdapat dalam model selama waktu tertentu. Selain itu, juga disusun himpunan data yang terdiri atas 1000 data untuk masing-masing variabel tekanan (kPa) dan temperatur (K) di setiap aliran masukan dan keluaran kompresor. Kesalahan acak ditambahkan pada data yang diperoleh dari simulasi dengan asumsi simpangan baku data berturut-turut untuk laju alir massa, tekanan, dan temperatur adalah sebesar 1, 0,5, dan 0,5. Data yang telah terkumpul dan dirandomisasi ini dianggap sebagai data pada kondisi operasi normal. Data pada kondisi operasi normal tersebut diuji dengan menggunakan statistika deskriptif untuk mengetahui distribusinya. Berdasarkan pengujian, data yang diperoleh merupakan data yang terdistribusi secara normal.

Berdasarkan model simulasi sistem kompresi gas sintesis multi tahap yang disajikan pada Gambar 1 diperoleh 48 variabel laju alir massa. Persamaan kendala disusun berdasarkan neraca massa total untuk variabel-variabel tersebut. Berikut ini disajikan persamaan kendala yang disusun berdasarkan variabel-variabel laju alir massa:

$$pk_1 = f_{14} - f_{41} \quad (12)$$

$$pk_2 = f_1 + f_{41} - f_3 \quad (13)$$

$$pk_3 = f_3 - f_4 \quad (14)$$

$$pk_4 = f_4 - f_{42} - f_5 \quad (15)$$

$$pk_5 = f_{42} - f_{35} \quad (16)$$

$$pk_6 = f_5 - f_6 \quad (17)$$

$$pk_7 = f_{37} - f_{43} \quad (18)$$

$$pk_8 = f_{43} + f_6 - f_7 \quad (19)$$

$$pk_9 = f_7 - f_8 - f_9 \quad (20)$$

$$pk_{10} = f_9 - f_{44} \quad (21)$$

$$pk_{11} = f_8 - f_{10} \quad (22)$$

$$pk_{12} = f_{10} - f_{11} \quad (23)$$

$$pk_{13} = f_{11} - f_{12} \quad (24)$$

$$pk_{14} = f_{12} - f_{13} - f_{14} \quad (25)$$

$$pk_{15} = f_{14} - f_{45} \quad (26)$$

$$pk_{16} = f_{13} - f_{15} - f_{16} \quad (27)$$

$$pk_{17} = f_{15} - f_1 \quad (28)$$

$$pk_{18} = f_{16} + f_{30} - f_{25} \quad (29)$$

$$pk_{19} = f_{25} - f_{27} \quad (30)$$

$$pk_{20} = f_{27} - f_{17} \quad (31)$$

$$pk_{21} = f_{17} + f_{29} - f_{26} \quad (32)$$

$$pk_{22} = f_{26} - f_{32} \quad (33)$$

$$pk_{23} = f_{32} + f_{31} - f_{33} \quad (34)$$

$$pk_{24} = f_{33} - f_{18} \quad (35)$$

$$pk_{25} = f_{18} - f_{36} \quad (36)$$

$$pk_{26} = f_{36} - f_{19} - f_{20} \quad (37)$$

$$pk_{27} = f_{20} - f_{40} \quad (38)$$

$$pk_{28} = f_{19} - f_{48} - f_{21} \quad (39)$$

$$pk_{29} = f_{48} - f_{39} \quad (40)$$

$$pk_{30} = f_{21} - f_{22} \quad (41)$$

$$pk_{31} = f_{22} - f_{23} - f_{24} \quad (42)$$

$$pk_{32} = f_{24} - f_{31} \quad (43)$$

$$pk_{33} = f_{23} - f_{30} \quad (44)$$

$$pk_{34} = f_{38} - f_{46} \quad (45)$$

$$pk_{35} = f_{46} - f_2 - f_{28} \quad (46)$$

$$pk_{36} = f_{28} - f_{47} \quad (47)$$

$$pk_{37} = f_2 - f_{29} \quad (48)$$

Matriks penyeimbang disusun berdasar pada koefisien persamaan kendala yang telah disajikan. Matriks penyeimbang tersebut dan matriks kovarian data digunakan untuk memperoleh matriks kovarian residual yang akan didekomposisi. Dekomposisi ortonormal dilakukan pada matriks kovarian residual sehingga diperoleh vektor komponen utama dari residual persamaan kendala. Komponen utama dan kontribusi residual persamaan kendala terhadap komponen utama yang diperoleh untuk data pengukuran yang diambil secara acak dari himpunan data yang mengandung kesalahan acak dan data pengukuran yang mengandung kesalahan nyata secara berturut-turut disajikan dalam Gambar 2 sampai dengan Gambar 5.

Hasil yang diperoleh untuk data pengukuran yang diambil acak dari himpunan data yang mengandung kesalahan acak disajikan pada Gambar 2 dan Gambar 3. Berdasarkan hasil tersebut, skor komponen-komponen utama yang diperoleh tidak lebih besar dari 2 atau kurang dari -2, sedangkan nilai kontribusi residual tidak lebih besar dari 4 atau kurang dari -2. Skor komponen-komponen utama berada dalam rentang dua kali simpangan baku sampel (± 2). Jika digunakan batas pengendalian secara statistika sebesar ± 3 kali simpangan baku data untuk data yang berada pada kondisi operasi normal, maka dapat dinyatakan tidak terdapat kesalahan nyata pada data pengukuran tersebut.

Kesalahan nyata berupa bias diberikan pada data laju alir aliran 8 dan 33 dalam simulasi untuk sistem kompresi gas sintesis multi tahap. Hasil yang diperoleh untuk data pengukuran yang mengandung kesalahan nyata disajikan pada Gambar 4 dan Gambar 5. Berdasarkan hasil tersebut, hampir semua komponen utama memiliki skor > 3 dan < -3 , hanya komponen utama 16, 19, 25, dan 37 yang berada dalam rentang ± 3 . Kontribusi residual yang besar terhadap skor komponen-komponen utama tersebut diberikan oleh residual persamaan kendala 9, 11, 23, dan 24. Laju alir massa yang berkaitan dengan persamaan kendala tersebut adalah laju alir aliran 7, 8, 9, 10, 18, 31, 32, dan 33. Berdasarkan persamaan kendala lain yang tidak memberikan nilai residual yang besar dapat dilakukan analisis terhadap laju alir yang dicurigai mengandung kesalahan nyata. Laju alir aliran 7 terdapat juga dalam persamaan kendala 8 yang memberikan kontribusi residual tidak besar sehingga laju alir aliran 7 tidak dicurigai. Laju alir aliran 9 dan 10 juga tidak dicurigai karena variabel tersebut terdapat juga dalam persamaan kendala 10 dan 12. Laju alir aliran 18 tidak dicurigai berdasarkan persamaan kendala 25, sedangkan laju alir aliran 31 dan 32 tidak dicurigai karena terdapat juga pada persamaan 32 dan 22. Untuk itu, dapat dinyatakan bahwa laju alir yang dicurigai dan dideteksi mengandung kesalahan nyata adalah laju alir aliran 8 dan 33.

Pengolahan data dengan menggunakan PCA juga dilakukan terhadap data pengukuran tekanan dan temperatur untuk model simulasi ini. Persamaan kendala disusun untuk variabel tekanan dan temperatur aliran masukan dan aliran keluaran masing-masing tahap kompresi pada model simulasi tersebut.

Persamaan kendala disusun untuk variabel tekanan dan temperatur dengan asumsi selisih tekanan keluaran dan tekanan masukan memiliki hubungan linier dengan selisih temperatur keluaran dan temperatur masukan suatu unit

Gambar 1. Model Simulasi Sistem Kompresi Gas Sintesis Multi tahap

Gambar 2. Komponen Utama dari Matriks Kovarian Residual Himpunan Data yang Mengandung Kesalahan Acak untuk Data Pengukuran yang Diambil secara Acak

Gambar 3. Kontribusi dari Residual Persamaan Kendala terhadap Komponen Utama untuk Data Pengukuran yang Diambil secara Acak dari Himpunan Data yang Mengandung Kesalahan Acak

Gambar 4. Komponen Utama dari Matriks Kovarian Residual untuk Data Pengukuran yang Mengandung Kesalahan Nyata

Gambar 5. Kontribusi dari Residual Persamaan Kendala terhadap Komponen Utama untuk Data Pengukuran yang Mengandung Kesalahan Nyata

kompresi. Pemodelan persamaan linier yang demikian berdasar pada pemodelan yang telah dilakukan oleh Rotem, dkk. (1999) untuk variabel tekanan dan temperatur. Pemodelan tersebut berbeda dengan persamaan yang umum digunakan untuk kasus kompresi berdasarkan teori dasar termodinamika. Perubahan tekanan umumnya tidak berhubungan secara linier dengan perubahan temperatur. Akan tetapi, pemodelan persamaan linier yang digunakan ini tidak menghasilkan kesalahan yang signifikan terhadap ketidaklinieran yang ada sehingga kesalahan tersebut dapat diabaikan.

Pemilihan persamaan linier dalam pemodelan dilakukan karena pengolahan data dengan PCA merupakan analisis yang berdasar pada hubungan/korelasi linier antarvariabel. Jika persamaan linier dapat diterapkan dengan baik pada pemodelan variabel tekanan dan temperatur, maka pemodelan dengan persamaan tak linier tidak terlalu diperlukan. Selain itu, penyelesaian persamaan linier memiliki beban komputasi yang lebih kecil daripada persamaan tak linier. Salah satu alasannya adalah persamaan linier tidak memerlukan tahap linierisasi pada pengolahan data dengan PCA. Persamaan kendala yang disusun adalah sebagai berikut:

$$pk_1 = T_4 - T_3 - k_1(P_4 - P_3) \quad (49)$$

$$pk_2 = T_{10} - T_8 - k_2(P_{10} - P_8) \quad (50)$$

$$pk_3 = T_{17} - T_{27} - k_3(P_{17} - P_{27}) \quad (51)$$

$$pk_4 = T_{18} - T_{33} - k_4(P_{18} - P_{33}) \quad (52)$$

Prosedur PCA untuk penyepakan data keadaan tunak diterapkan pada data yang telah dihimpun berdasarkan persamaan kendala di atas. Komponen utama dan kontribusi residual persamaan kendala terhadap komponen utama yang diperoleh untuk data pengukuran yang diambil secara acak dari himpunan data yang mengandung kesalahan acak dan data pengukuran yang mengandung kesalahan nyata secara berturut-turut disajikan dalam Gambar 6 sampai dengan Gambar 9 berikut.

Gambar 6. Komponen Utama dari Matriks Kovarian Residual Himpunan Data yang Mengandung Kesalahan Acak untuk Data Pengukuran Tekanan dan Temperatur yang Diambil secara Acak

Gambar 7. Kontribusi Residual Persamaan Kendala terhadap Komponen Utama untuk Data Pengukuran Tekanan dan Temperatur yang Mengandung Kesalahan Acak

Gambar 8. Komponen Utama dari Matriks Kovarian Residual untuk Data Pengukuran Tekanan dan Temperatur yang Mengandung Kesalahan Nyata

Gambar 9. Kontribusi Residual Persamaan Kendala terhadap Komponen Utama untuk Data Pengukuran Tekanan dan Temperatur yang Mengandung Kesalahan Nyata

Hasil yang diperoleh untuk data pengukuran yang diambil acak dari himpunan data yang mengandung kesalahan acak disajikan pada Gambar 6 dan 7. Berdasarkan hasil tersebut, komponen utama yang diperoleh tidak lebih besar dari 1,5 atau kurang dari -1,5, sedangkan nilai kontribusi residual tidak lebih besar dari 1 atau kurang dari -1. Skor komponen-komponen utama berada dalam rentang tiga kali simpangan baku sampel ($\pm 1,5$). Jika digunakan batas pengendalian secara statistika sebesar ± 3 kali simpangan baku data untuk data yang berada pada kondisi operasi normal, maka dapat dinyatakan tidak terdapat kesalahan nyata pada data pengukuran tersebut. Data pengukuran tersebut hanya mengandung kesalahan acak.

Kesalahan nyata berupa bias diberikan pada data pengukuran tekanan P_8 dan P_{10} serta temperatur T_{10} dan T_{33} dalam simulasi. Hasil yang diperoleh untuk data pengukuran yang mengandung kesalahan nyata diberikan pada Gambar 8 dan 9. Berdasarkan hasil tersebut, terdapat komponen utama yang skornya di luar rentang tiga kali simpangan baku ($\pm 1,5$) sebesar 5,25 untuk komponen utama 3. Prosedur PCA dapat mendeteksi adanya kesalahan nyata pada data pengukuran tersebut. Residual persamaan kendala kedua dan keempat memberikan kontribusi yang signifikan terhadap komponen utama. Data pengukuran yang terlibat pada persamaan kendala tersebut adalah T_8 , T_{10} , P_8 , P_{10} , T_{18} , T_{33} , P_{18} , dan P_{33} . Pengamatan lebih lanjut secara individual terhadap data pengukuran tekanan dan temperatur yang terkontaminasi oleh kesalahan nyata dapat dilakukan pada unit operasi yang berhubungan dengan aliran tersebut. Hal ini dapat dilakukan karena data pengukuran tersebut terdeteksi mengandung kesalahan nyata berdasarkan pengujian komponen utamanya.

Penyepakan data dilakukan terhadap data pengukuran yang diperoleh dari model simulasi sistem kompresi gas sintesis multi tahap. Penyepakan data dilakukan untuk memperoleh

data pengukuran yang tersepakati dengan cara meminimumkan fungsi obyektif. Peminimuman fungsi obyektif ini dilakukan dengan menggunakan algoritma *Statistical Data Fit* (SDF) untuk penyepakan data pada data laju alir massa. Algoritma SDF dipakai karena peminimuman fungsi obyektif dengan menggunakan *Solver* pada perangkat lunak *Ms. Excel* kurang memuaskan untuk kasus data laju alir. Penyepakan data untuk data tekanan dan temperatur dilakukan dengan cara peminimuman fungsi obyektif dan memenuhi persyaratan kendala yang disusun. Peminimuman fungsi obyektif ini dilakukan dengan menggunakan *Solver* yang ada pada perangkat lunak *Ms. Excel*. Berikut ini disajikan tabel yang memuat hasil penyepakan data yang dilakukan terhadap data laju alir massa, untuk data pengukuran yang diambil acak dari himpunan data yang mengandung kesalahan acak dan data yang mengandung kesalahan nyata.

Tabel 1. Hasil Penyepakan Data Pengukuran Laju Alir Massa yang Diambil Acak dari Himpunan Data yang Mengandung Kesalahan Acak

No	Var.	x^*	3σ	$\bar{x}$	F. Obyektif	
					PK	SDF
1	f_{14}	60413.45	3	60413.89	1	0.00E+00
2	f_{11}	60414.71	3	60413.89	1	0.00E+00
3	F_1	15383.27	3	15382.09	2	0.00E+00
4	F_1	75796.11	3	75795.97	3	0.00E+00
5	F_4	75796.17	3	75795.97	4	0.00E+00
6	f_{12}	172.53	3	172.83	5	1.00E-12
7	f_{14}	172.4	3	172.83	6	0.00E+00
8	F_8	75622.73	3	75623.14	7	0.00E+00
9	F_8	75622.4	3	75623.14	8	0.00E+00
10	f_{17}	1252.23	3	1252.48	9	-4.00E-12
11	f_{14}	1252.31	3	1252.48	10	0.00E+00
12	F_7	76875.49	3	76875.62	11	0.00E+00
13	F_4	76875.63	3	76875.62	12	0.00E+00
14	F_9	0	3.00E-13	0	13	0.00E+00
15	f_{10}	76875.95	3	76875.62	14	-8.00E-13
16	f_{11}	76877.53	3	76875.62	15	0.00E+00
17	f_{12}	76873.11	3	76875.62	16	0.00E+00
18	f_{11}	76876.45	3	76875.62	17	0.00E+00
19	f_{14}	0	3.00E-13	0	18	0.00E+00
20	f_{14}	0	3.00E-13	0	19	0.00E+00
21	f_{15}	0	3.00E-13	0	20	0.00E+00
22	f_{14}	15381.28	3	15382.09	21	0.00E+00
23	f_{10}	61494.4	3	61493.53	22	0.00E+00
24	f_{10}	120912.8	3	120912.6	23	0.00E+00
25	f_{14}	182406	3	182406.2	24	0.00E+00
26	f_{17}	182406	3	182406.2	25	0.00E+00
27	f_{12}	182405.4	3	182406.2	26	-4.00E-11
28	f_{10}	202582.5	3	202583.5	27	0.00E+00
29	f_{10}	202583.3	3	202583.5	28	0.00E+00
30	F_2	197334.7	3	197334.8	29	0.00E+00
31	f_{24}	197335.9	3	197334.8	30	0.00E+00
32	f_{18}	5248.79	3	5248.67	31	0.00E+00
33	f_{17}	5249.79	3	5248.67	32	0.00E+00
34	f_{10}	379741.6	3	379741	33	0.00E+00
35	f_{12}	379740.9	3	379741	34	0.00E+00
36	f_{11}	483859	3	483861	35	-7.00E-12
37	f_{11}	863601.6	3	863602	36	0.00E+00
38	f_{18}	863602.3	3	863602	37	0.00E+00
39	f_{10}	863602.4	3	863602		
40	f_{10}	863602	3	863602		
41	f_{10}	0	4.00E-13	0		
42	f_{10}	0	4.00E-13	0		
43	f_{18}	258827.9	3	258828.4		
44	f_{19}	258828.3	3	258828.4		
45	f_{11}	604773.7	3	604773.6		
46	f_{12}	604774.4	3	604773.6		
47	f_{13}	120912.1	3	120912.6		
48	f_{14}	483861.8	3	483861		

Penyepakatan data dapat dilakukan dengan baik untuk data pengukuran yang hanya mengandung kesalahan acak. Data pengukuran yang tersepakati memberikan nilai residual nol untuk tiap persamaan kendala. Hal ini berarti data laju alir massa yang tersepakati mengikuti kaidah neraca massa. Selain itu, setiap data yang tersepakati berada dalam rentang ± 3 kali simpangan baku sehingga data tersebut dapat diterima sebagai data pengukuran yang benar.

Tabel 2. Hasil Penyepakatan Data Pengukuran Laju Alir Massa yang Mengandung Kesalahan Nyata

No	Var.	$\bar{x}$	3σ	$\hat{x}$	F. Obyektif	
1	f_{14}	60414.36	3	60415.01	PK	SDF
2	f_{41}	60415	3	60415.01	1	0.00E+00
3	f_1	15382.48	3	15378.49	2	0.00E+00
4	f_4	75796.91	3	75793.5	3	0.00E+00
5	f_4	75795.19	3	75793.5	4	0.00E+00
6	f_{42}	172.89	3	174.53	5	5.00E-12
7	f_{15}	171.73	3	174.53	6	0.00E+00
8	f_4	75621.13	3	75618.96	7	0.00E+00
9	f_6	75621.56	3	75618.96	8	0.00E+00
10	f_{17}	1251.74	3	1246.64	9	2.00E-13
11	f_{15}	1250.76	3	1246.64	10	0.00E+00
12	f_7	76876.14	3	76865.6	11	0.00E+00
13	f_8	76800	3	76865.6	12	0.00E+00
14	f_9	0	3.00E-13	0	13	0.00E+00
15	f_{10}	76874.55	3	76865.6	14	2.00E-12
16	f_{11}	76874.54	3	76865.6	15	0.00E+00
17	f_{12}	76874.29	3	76865.6	16	0.00E+00
18	f_{13}	76875.68	3	76865.6	17	0.00E+00
19	f_{14}	0	3.00E-13	0	18	0.00E+00
20	f_{14}	0	3.00E-13	0	19	0.00E+00
21	f_{15}	0	3.00E-13	0	20	0.00E+00
22	f_{15}	15383.05	3	15378.49	21	0.00E+00
23	f_{16}	61492.97	3	61487.11	22	0.00E+00
24	f_{10}	120914.4	3	120915.3	23	0.00E+00
25	f_{15}	182406.2	3	182402.5	24	0.00E+00
26	f_{17}	182405.7	3	182402.5	25	0.00E+00
27	f_{17}	182406.4	3	182402.5	26	-4.00E-11
28	f_{18}	202585.5	3	202583	27	0.00E+00
29	f_{16}	202584.2	3	202583	28	0.00E+00
30	f_7	197334.9	3	197332.1	29	0.00E+00
31	f_{10}	197333.2	3	197332.1	30	0.00E+00
32	f_{18}	5248.84	3	5250.96	31	0.00E+00
33	f_{17}	5249.45	3	5250.96	32	0.00E+00
34	f_{16}	379741.1	3	379734.5	33	0.00E+00
35	f_{12}	379741	3	379734.5	34	0.00E+00
36	f_{11}	483862	3	483850.7	35	1.00E-11
37	f_{13}	863500	3	863585.2	36	0.00E+00
38	f_{18}	863599.8	3	863585.2	37	0.00E+00
39	f_{16}	863602.8	3	863585.2		
40	f_{19}	863601.8	3	863585.2		
41	f_{10}	0	4.00E-13	0		
42	f_{10}	0	4.00E-13	0		
43	f_{18}	258827	3	258819.2		
44	f_{19}	258827	3	258819.2		
45	f_{11}	604774.2	3	604766		
46	f_{12}	604773.2	3	604766		
47	f_{13}	120913.2	3	120915.3		
48	f_{14}	483860.1	3	483850.7		

Kesalahan nyata pada data laju alir 8 dan laju alir 33 menyebabkan hasil dari sebagian besar data yang tersepakati berada di luar rentang simpangan baku datanya. Data pengukuran awal yang tidak diberi bias merupakan data yang dekat dengan nilai sejati pengukuran dan memiliki beda dari nilai sejati tidak lebih dari simpangan bakunya. Dengan demikian, data yang tersepakati tersebut tidak tepat jika diterima sebagai data pengukuran yang benar sehingga dapat dinyatakan

penyepakatan data tidak dapat memberikan hasil yang baik jika terdapat kesalahan nyata pada data pengukuran. Keberadaan kesalahan nyata pada data pengukuran perlu dideteksi dan dilokalisir dengan cepat sehingga dapat dilakukan tindakan khusus terhadap data tersebut untuk mencegah terjadinya aksi pengendalian dan persepsi yang salah pada proses yang berlangsung.

Penyepakatan data juga dilakukan untuk data pengukuran tekanan dan temperatur. Penerapan penyepakatan data dilakukan untuk data tekanan dan temperatur yang diambil acak dari himpunan data yang mengandung kesalahan acak (Tabel 3) dan data pengukuran yang mengandung kesalahan nyata (Tabel 4).

Penyepakatan data dapat diterapkan dengan baik pada data pengukuran yang hanya mengandung kesalahan acak. Data tersepakati yang diperoleh dapat diterima sebagai data yang benar karena berada dalam rentang nilai distribusi z pada tingkat keyakinan 95% dikali simpangan bakunya dan data tersebut memberikan nilai residual nol untuk semua persamaan kendala yang disusun.

Penyepakatan data pengukuran tekanan dan temperatur yang mengandung kesalahan nyata memberikan hasil yang tidak dapat diterima sebagai data yang benar. Data tersepakati yang dihasilkan memiliki nilai residual yang tidak sama dengan nol atau mendekati nol pada persamaan kendala ke-2 dan ke-4. Data P_8 , P_{10} , T_{10} , dan T_{33} yang mengandung kesalahan nyata tidak dapat memenuhi persyaratan persamaan kendala ke-2 dan ke-4. Data tersepakati yang dihasilkan variabel-variabel tersebut dalam rentang nilai distribusi z pada tingkat keyakinan 95% dikali simpangan bakunya tidak dapat menghasilkan persamaan kendala yang sama dengan atau dekat dengan nol. Berdasarkan hasil tersebut, penyepakatan data tidak dapat dilakukan dengan baik jika terdapat kesalahan nyata pada data pengukuran.

Tabel 3. Hasil Penyepakatan Data Pengukuran Tekanan dan Temperatur dari Himpunan Data yang Mengandung Kesalahan Acak

No	Var.	$\bar{X}$	σ	$\hat{x}$	F. Obyektif	
1	P_1	3197.09	0.50	3197.07	PK	Optimasi
2	P_4	5906.84	0.50	5906.85	1	0E+00
3	P_8	5873.24	0.50	5873.23	2	0E+00
4	P_{10}	9953.78	0.50	9953.80	3	0E+00
5	P_{17}	9851.31	0.50	9851.31	4	6E-15
6	P_{17}	16881.95	0.50	16881.95	F. Oby.	9.60
7	P_{13}	16797.20	0.50	16797.20	Prob.	95%
8	P_{18}	17200.82	0.50	17200.83	Dist. Z	1.96
9	T_1	302.36	0.50	302.94		
10	T_4	382.27	0.50	381.70		
11	T_8	315.87	0.50	316.56		
12	T_{10}	387.05	0.50	386.37		
13	T_{17}	297.92	0.50	298.12		
14	T_{17}	364.83	0.50	364.63		
15	T_{13}	318.68	0.50	319.27		
16	T_{18}	322.77	0.50	322.17		

Data yang mengandung kesalahan nyata perlu dideteksi dan dieliminasi atau diperlakukan khusus terlebih dahulu sebelum proses penyepakan data dilakukan.

Tabel 4. Hasil Penyepakan Data Pengukuran Tekanan dan Temperatur yang Mengandung Kesalahan Nyata

No	Var.	$\bar{x}$	σ	$\hat{\bar{x}}$	F. Obyektif	
1	P ₁	3197.20	0.50	3197.19	PK	Optimasi
2	P ₄	5906.50	0.50	5906.51	1	-1E-10
3	P ₈	5850.00	0.50	5850.98	2	-3.48
4	P ₁₀	9900.00	0.50	9899.02	3	0.00
5	P ₂₇	9851.25	0.50	9851.25	4	3.22
6	P ₁₇	16882.04	0.50	16882.04	F.Oby.	31.96
7	P ₃₃	16797.69	0.50	16796.71	Prob.	95%
8	P ₁₈	17200.41	0.50	17201.39	Dist. Z	1.96
9	T ₁	302.92	0.50	303.30		
10	T ₄	382.43	0.50	382.05		
11	T ₈	316.19	0.50	315.21		
12	T ₁₀	380.00	0.50	380.98		
13	T ₂₇	296.89	0.50	296.79		
14	T ₁₇	363.20	0.50	363.30		
15	T ₃₃	315.00	0.50	315.98		
16	T ₁₈	323.08	0.50	322.10		

Selanjutnya, penerapan prosedur PCA untuk penyepakan data dinamik dilakukan pada data pengukuran yang diperoleh selama waktu tertentu secara langsung. Penyepakan data

dinamik merupakan proses untuk mengeliminasi kesalahan dari variabel-variabel sensor di dalam suatu rentang waktu. Tahap ini diawali dengan pembangunan model simulasi untuk menggambarkan proses yang terjadi pada unit operasi kompresor multi tahap berbantuan perangkat lunak Aspen-Hysys 2004.

Data pengukuran laju alir massa, tekanan, dan temperatur disiapkan untuk menguji penerapan PCA tersebut. Persamaan kendala yang disusun berdasarkan laju alir massa, tekanan, dan temperatur pada pengolahan data dengan PCA untuk model simulasi sistem kompresi gas sintesis multi tahap digabungkan untuk memperoleh matriks penyeimbang yang baru. Persamaan kendala yang didapat 41 buah tersusun dari 37 persamaan kendala berdasarkan neraca massa total untuk variabel laju alir dan 4 persamaan kendala berdasarkan hubungan antara variabel tekanan dan temperatur. Kesalahan nyata berupa bias diberikan pada variabel f_{33} , f_{105} , f_{27} , f_{18} , T_4 , T_8 , T_{17} , dan T_{33} . Salah satu hasil penerapan prosedur PCA pada data pengukuran yang diperoleh untuk suatu waktu tertentu dengan simulasi kesalahan nyata disajikan dalam Gambar 10 dan Gambar 11 berikut.

Gambar 10. Komponen Utama dari Matriks Kovarian Residual untuk Data Pengukuran Laju Alir Massa, Tekanan, dan Temperatur

Gambar 11. Kontribusi dari Residual Persamaan Kendala terhadap Komponen Utama untuk Data Pengukuran Laju Alir Massa, Tekanan, dan Temperatur

Hasil yang diperoleh untuk data laju alir massa, tekanan, dan temperatur disajikan pada Gambar 10 dan 11. Simpangan baku alat ukur yang digunakan pada data diasumsikan sebesar 5 untuk laju alir massa dan 0,5 untuk tekanan dan temperatur. Berdasarkan hasil tersebut, hampir semua komponen utama memiliki skor lebih besar dari 5 dan kurang dari -5. Kontribusi residual yang besar terhadap skor komponen-komponen utama tersebut diberikan oleh residual persamaan kendala 2, 3, 11, 12, 19, 20, 24, dan 25. Laju alir massa yang berkaitan dengan persamaan kendala tersebut adalah laju alir aliran 1, 3, 4, 8, 10, 11, 17, 18, 25, 27, 33, 36, dan 41. Laju alir aliran 3, 10, 18, dan 27 merupakan variabel-variabel yang terdapat dalam persamaan kendala yang memberikan kontribusi residual besar terhadap komponen-komponen utama. Variabel-variabel tersebut tidak terdapat pada persamaan kendala lain yang tidak memberikan kontribusi yang besar. Untuk itu, dapat dinyatakan bahwa laju alir yang dicurigai dan dideteksi mengandung kesalahan nyata adalah laju alir aliran 3, 10, 18 dan 27. Variabel-variabel T_1 , T_8 , T_{17} , dan T_{33} juga merupakan data yang mengandung kesalahan nyata berupa bias. Akan tetapi, kontribusi residual persamaan kendala yang berkaitan dengan variabel-variabel tersebut tidak terlalu signifikan dibandingkan kontribusi residual dari persamaan kendala 2, 3, 11, 12, 19, 20, 24, dan 25. Hal ini berkaitan dengan besarnya bias yang mungkin terjadi pada data pengukuran tersebut. Data laju alir memiliki bias yang lebih besar dibandingkan bias data temperatur untuk signifikansi deviasi data yang sama. Jika ditinjau dari residual persamaan kendala yang berkaitan dengan data pengukuran temperatur tersebut, maka variabel-variabel tersebut perlu dicurigai sebagai data yang mengandung kesalahan nyata. Peninjauan secara individual terhadap data pengukuran tersebut dapat dilakukan.

Selanjutnya, penyepakatan data dilakukan untuk data pengukuran laju alir, tekanan, dan temperatur yang mengandung kesalahan acak. Penyepakatan data tidak dilakukan untuk data pengukuran yang dianalisis menggunakan prosedur PCA sebelumnya karena data tersebut mengandung kesalahan nyata.

Penyepakatan data dapat dilakukan dengan baik untuk data pengukuran yang hanya mengandung kesalahan acak. Data pengukuran yang tersepakati memberikan nilai residual nol atau mendekati nol untuk tiap persamaan kendala. Hal ini berarti data laju alir massa yang tersepakati mengikuti kaidah neraca massa, sedangkan data tekanan dan temperatur yang tersepakati sesuai dengan persamaan yang menghubungkan antara tekanan dan temperatur tersebut. Selain itu, setiap data yang tersepakati berada dalam rentang ± 3 kali

simpangan baku sehingga data tersebut dapat diterima sebagai data pengukuran yang benar.

5. Kesimpulan

Dari penelitian yang telah dilakukan dapat diambil beberapa kesimpulan.

Prosedur PCA dapat diterapkan untuk mendeteksi dan mengidentifikasi kesalahan nyata pada data pengukuran laju alir, tekanan, dan temperatur.

Teknik penyepakatan data dapat diterapkan dengan baik pada unit operasi kompresor multi tahap jika kesalahan nyata tidak ada pada data hasil pengukuran.

Aplikasi teknik PCA pada penyepakatan data keadaan tunak dapat dilakukan untuk mendeteksi letak kesalahan nyata data pengukuran sehingga eliminasi atau koreksi terhadap data yang mengandung kesalahan nyata tersebut dapat dilakukan dengan tepat.

Teknik PCA dapat mendeteksi dan melokalisasi kesalahan dengan cepat sehingga memungkinkan untuk diterapkan pada penyepakatan data dinamik. Akan tetapi, aplikasi teknik PCA ini perlu diawali dengan pengumpulan pengetahuan proses yang lengkap sehingga tingkat keyakinan terhadap keputusan penerimaan/penolakan suatu data lebih tinggi.

Ucapan Terima Kasih

Penulis mengucapkan terima kasih kepada Gandi Rahmawan Setyadi atas saran dan diskusi selama penelitian ini berlangsung.

Daftar Notasi

$\bar{x}$	= nilai hasil pengukuran
$\hat{x}$	= nilai sejati pengukuran
X	= matriks data pengukuran
$\bar{X}$	= matriks nilai data sejati
E	= residual
t_i	= skor komponen utama
p_i	= vektor <i>loading</i> dari komponen utama
T	= transposisi dari matriks
S	= matriks kovarian
L	= matriks diagonal (<i>eigenvalue</i>)
U	= vektor nilai eigen (<i>eigenvector</i>)
B	= matriks penyeimbang
H_e	= matriks kovarian residual
K	= matriks K pada algoritma SDF
x_0	= variabel pembuat nol suatu fungsi
$f(x_0)$	= fungsi yang akan dinolkan/persamaan kendala
F_x	= matriks <i>Jacobian</i> /matriks penyeimbang
R_x	= matriks kovarian data
pk_i	= persamaan kendala ke- i
PK	= persamaan kendala

σ	= simpangan baku data
L_e	= matriks <i>eigenvalue</i> residual
U_e	= matriks <i>eigenvector</i> residual
W_e	= koefisien kombinasi linier residual
y_e	= komponen utama dari residual
g_j	= kontribusi dari residual ke-j
f_i	= laju alir massa aliran I
T_i	= temperatur aliran I
P_i	= tekanan aliran I
k_i	= konstanta pada persamaan energi unit kompresor tahap I

Daftar Pustaka

- [1] Amand, Th., Heyen, G., dan Kalitventzeff, B. (2001), Plant monitoring and fault detection: Synergy between data reconciliation and principal component analysis, *Computers and Chemical Engineering* **25**, 501-507.
- [2] Jackson, J. E. (1991), *A User's Guide To Principal Components*, John Wiley & Sons, New York, 4-104.
- [3] Klančar, G. (1998), Fault Detection and Isolation by means of Principal Component Analysis. University of Ljubljana, Slovenia.
- [4] Madron, F. dan Sharp, D. (1992), *Process Plant Performance: Measurement and Data Processing for Optimization and Retrofits*, Ellis Horwood, New York, 66-119.
- [5] Narasimhan, S., dan Shah, S. L. (2004), Model identification and error covariance matrix estimation from noisy data using PCA, *Proceedings of 7th International Symposium on Advanced Control of Chemical Processes*, Hong Kong, 567-572.
- [6] Rotem, Y., Wachs, A., dan Lewin, D.R. (1999), Ethylene Compressor Monitoring using Model-Based PCA, Technion I.I.T., Haifa 32000, Israel.
- [7] Shah, S., Miller, R., Takada, H., Morinaga, K., dan Satou, T. (1998), Modelling and Control of A Tubular Reactor: A PCA-Based Approach, Kurashiki, Okayama 712, Jepang.
- [8] Tong, H. dan Crowe, C. M. (1996), Detecting persistent gross errors by sequential analysis of principal components, *Computers and Chemical Engineering* **20**, Suppl., S733-S738.
- [9] Venkatasubramanian, V., Rengaswamy, R., Kavuri, S. N., dan Yin, K. (2003), A review of process fault detection and diagnosis Part III: Process history based methods, *Computers and Chemical Engineering* **27**, 327-346.