

Prediksi *Slippage* menggunakan Metode Regresi

Berkah Hani

School of Engineering and Technology: Petroleum Engineering, Tanri Abeng University, Jakarta, Indonesia
berkah.hani@tau.ac.id

Diterima 31 Oktober 2018
Disetujui 21 Desember 2018

Abstract — Many pump slippage formula have been developed over the years. Most of these formulas over predict the pump slippage and a more accurate formula is needed. A new method for calculating slippage is introduced by now. Using an average velocity to determine slippage value and also the position of *plunger* against barrel (concentric and eccentric position) are studied in this new method formula. A formula which is has the lowest value will be used as method selection in this research. The formula will show that high or low pump slippage will make an impact to pump efficiency which is directly connected to production optimization. Slippage is important for lubricant around plunger inside the barrel. The accuracy slippage maximizes oil production. This paper explained several regression methods to predict slippage value. Several regression methods that is used for example Huber estimation, Ordinary Least Square, Theil-Sen Estimation, RANSAC estimation and Support Vector Machine-Regression. The result shows using SVR give the better result quantitatively compared to other linear regression algorithms.

Index Terms—*Slippage prediction, Huber Estimation, Linear Regression, Support Vector Regression, Theil-Sen Estimation, RANSAC-Estimation*

I. PENDAHULUAN

Slippage berpengaruh terhadap lubrikasi di sekitar *plunger*, efisiensi pompa yang dihasilkan sehingga memaksimalkan laju produksi minyak yang lebih optimal. Sayangnya perhitungan saat ini memberikan nilai hasil yang jauh melebihi dari hasil *slippage* yang ada pada data lapangan.

Perhitungan – perhitungan yang sebelumnya memperlihatkan hasil jauh melebihi harga *slippage actual* dilapangan. Hal ini menyebabkan efek yang besar pada desain peralatan pompa yang akan dipersiapkan pada sumur tertentu. Dan tentu saja pada bagian penting yaitu sistem lubrikasi yang kurang memadai pada saat terjadi gesekan antara *plunger* dan *barrel*. Hal ini dapat membahayakan jika terjadi pergesekan yang terlalu sering sehingga dapat menimbulkan percikan api. Hal inilah yang harus dihindari sedini mungkin [1], [2].

Beberapa penelitian yang mencoba menyelesaikan permasalahan perhitungan *slippage*: Robinson[1], Davis, Reekstin dan Stearn [2] dan Arco-Harbinson Fisher[3]. Metode-metode yang diajukan merupakan turunan sebelumnya dan hasilnya selalu *over* dari data yang ada. Dikarenakan hal inilah dikemukakan metode statistik yang dianggap dapat melakukan prediksi terhadap *slippage value*. Metode-metode regresi dipilih dikarenakan metode ini baik dalam melakukan prediksi nilai dan pengajuan beberapa metode yang *robust* terhadap *outlier* juga diharapkan memberikan hasil yang jauh lebih baik dibandingkan dengan pendekatan matematis seperti penelitian sebelumnya.

Artikel ilmiah ini dibagi menjadi 5 bagian dimana Bagian 1 adalah pendahuluan, Bagian 2 merupakan landasan teori. Bagian 3 merupakan eksperimen, Bagian 4 hasil eksperimen dan Bagian 5 kesimpulan.

II. LANDASAN TEORI

A. *Slippage*

Slippage adalah banyak cairan (dalam *barrel/day*) disekeliling *plunger* didalam *barrel* yang berguna untuk lubrikasi *plunger* tersebut dari gesekan yang terjadi antar plat yaitu *barrel* dan *plunger*

Slippage yang sesuai adalah keseimbangan antara lubrikasi pompa yang mampu memperpanjang umur pompa tersebut dengan efisiensi volume pompa. Efisiensi volume pompa dijelaskan sebagai banyaknya fluida yang terdisplesi per *stroke* dibagi dengan teori perhitungan *displacement* pompa.

Pada saat *upstroke*, terjadi perbedaan tekanan yang besar di sekeliling *plunger*, yang membuat fluida bergerak diantara *plunger* dan *barrel*. Fluida yang mengalir di antara *plunger* dan *barrel* disebut *pump slippage* dan ini merupakan salah satu faktor yang mempengaruhi lubrikasi dan efisiensi pompa tersebut.

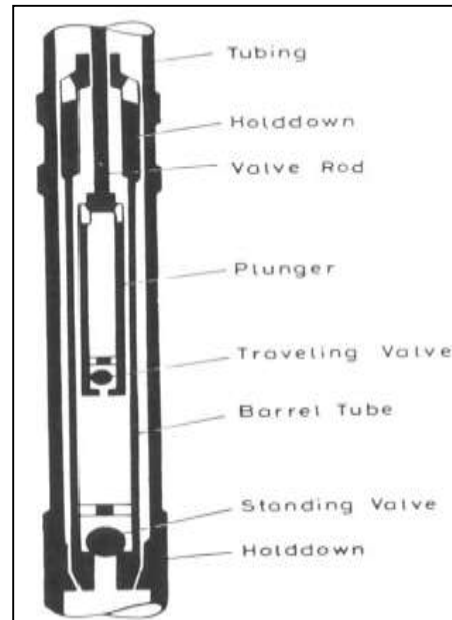
Ketepatan dari banyaknya lubrikasi pompa merupakan salah satu faktor yang dapat memperpanjang umur pompa. Faktor lain yang mempengaruhi efisiensi pompa adalah terjepitnya *plunger* yang berada didalam *barrel* dapat memperpendek dari panjang *stroke actual* pompa.

Penyebab terjepitnya *plunger* dikarenakan *clearance* antara *plunger* dan *barrel* tidak tepat. Jika *clearance* terlalu kecil maka kesempatan *plunger* untuk terjepit di dalam *barrel* lebih besar dan juga jika *clearance* terlalu besar, produktifitas dari sumur tersebut menjadi rendah. *Clearance* yang ada harus memiliki area yang cukup agar fluida yang mengalir sesuai dengan banyaknya kebutuhan *slippage* yang diperlukan untuk melubrikasi pompa tersebut. Pasir dan partikel lainnya harus dapat melewati jarak yang ada diantara *plunger* dan *barrel*.

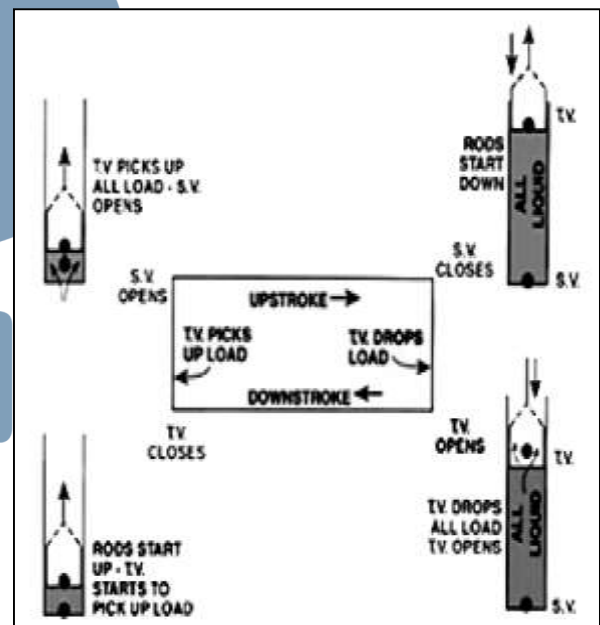
Standarisasi dari banyaknya *slippage* yang dibutuhkan yaitu berkisar antara 2% - 10% dari total laju produksi sumur per hari untuk melubrikasi *plunger* pada pompa. Maka dari itu diperlukan perhitungan untuk menghitung besar *clearance* yang dibutuhkan atau ukuran persentase *slippage* yang akan mempengaruhi ukuran *plunger* pompa [4].

B. Cara Kerja Plunger Pump

Plunger pump terdiri dari beberapa bagian utama yaitu; *barrel*, *plunger*, *traveling valve* dan *standing valve*. Gambar 1 [2] menggambarkan beberapa tipe *plunger pump*. *Traveling valve* dihubungkan ke *plunger*, jadi pada saat *plunger* bergerak ke atas, *traveling valve* akan menutup. Fluida tidak akan mengalir kembali melalui *plunger*. Saat bersamaan, *standing valve* akan terbuka dan fluida mengalir ke dalam *barrel* yang berada dibawah *plunger*. Dan ketika *plunger* mencapai titik tertinggi dari siklus *up stroke*, maka secara otomatis *plunger* kembali bergerak kebawah untuk memulai proses *down stroke*. Pada proses ini, *traveling valve* akan terbuka dan *standing valve* otomatis tertutup, mendesak fluida masuk ke dalam dan membuka *traveling valve* dan terus bergerak keluar *plunger*. Saat *plunger* mencapai titik terendah, maka proses *down stroke* berakhir dan memulai proses *up stroke*. Begitu seterusnya sebagai suatu sistem *up stroke* dan *down stroke*. Hal ini dapat kita lihat pada gambar 2 [2].



Gambar 1. Diagram *plunger pump*.



Gambar 2 *Plunger pump cycle*

C. Regresi

Linear Regresi memiliki fungsi yang dapat dilihat pada persamaan (1): [5]

$$\hat{y} = b_0 + b_1X_1 + b_2X_2 + \dots + b_nX_n \quad (1)$$

Dimana \hat{y} adalah hasil prediksi dari target, b_0 adalah intersep, b_1 - b_n adalah koefisien regresi dan X adalah fitur yang ada

Metode-metode regresi yang digunakan adalah:

1. Ordinary Least Square (OLS)

Metode OLS merupakan metode sederhana yang digunakan dalam pembuatan koefisien regresi, sayangnya metode ini tidak terlalu robust terhadap *outlier*. Perumusan untuk OLS dapat dilihat pada persamaan 2:[5]

$$b = (X_i^T \cdot X_i)^{-1} X_i^T Y \quad (2)$$

Dimana Y adalah target.

2. Huber Estimation

Huber estimation memberikan *linear loss* kepada data sehingga dapat diketahui apakah data tersebut adalah *outlier* atau bukan. Berbeda dengan *Theil-Sen* dan *RANSAC* yang mengabaikan *outlier*, *Huber estimation* menganggap bahwa *outlier* memiliki arti dalam pembuatan model akan tetapi metode ini memberikan *weight* yang lebih sedikit terhadap data yang dianggap sebagai *outlier* [6]. *Loss Function* untuk *Huber Estimation* dapat dilihat pada persamaan (3) dan (4):[7]

$$\min_{w, \sigma} \sum_{i=1}^n \left(\sigma + H_\epsilon \left(\frac{X_i w - y_i}{\sigma} \right) \sigma \right) + \alpha \|w\|_2^2 \quad (3)$$

Dimana:

$$H_\epsilon(z) = \begin{cases} z^2, & \text{if } |z| < \epsilon, \\ 2\epsilon|z| - \epsilon^2, & \text{otherwise} \end{cases} \quad (4)$$

Dimana nilai w dan σ adalah parameter yang perlu dioptimalkan

3. Theil-Sen Estimation

Theil-Sen estimation merupakan salah satu metode yang *robust* terhadap *outlier*, metode ini didapatkan dengan cara menghitung *slopes* dan *intercept* dari *sub-population* untuk semua kombinasi populasi [8]. *Theil-Sen Estimation* sering disebut dengan *generalized-median-estimator*[9]. Fungsi *Theil-Sen Estimator* dapat dilihat pada persamaan (5):[8]

$$b = \text{Med}\{Y_i - Y_j / X_i - X_j : X_i \neq X_j, 1 \leq i < j \leq n\} \quad (5)$$

Dimana $\text{Med}\{B_j; j \in J\}$ melambangkan median dari $\{B_j; j \in J\}$

4. Random Sample Consensus (RANSAC)

Random Sample Consensus estimation adalah salah satu algoritma metode robust regression selain *Huber* ataupun *Theilsen*. Metode ini digunakan untuk mengestimasi parameter dengan akurasi yang tinggi walaupun memiliki banyaknya *outlier* pada dataset [10].

Algoritma *RANSAC* memiliki dua tahapan yang penting yaitu[11]:

1. Pada tahapan pertama, sebuah sampel subset dari data dipilih secara acak dari input dataset.

Pencarian model dan dilakukan pada subset sampel.

2. Pada tahapan kedua, algoritma mengecek elemen dari semua dataset yang berpengaruh terhadap pengestimasi model yang dilakukan pada tahapan pertama. Data yang tidak masuk di dalam *fit* dalam model akan dianggap sebagai *outlier*

5. Support Vector Machine - Regression (SVR)

Support Vector Machine - Regression merupakan salah satu metode yang bekerja pada bidang linear dan non-linear dimana dibagi menjadi *linear*, *polynomial* dan *radial function*. Metode *SVR* mengambil konsep yang sama dengan *Support Vector Machine (SVM)*. Pada kasus regresi, margin of tolerance (epsilon) ditetapkan dalam aproksimasi ke *SVM*. Algoritma *SVR* lebih rumit dibandingkan dengan *SVM* akan tetapi secara sederhana *SVR* tetap melakukan *minimize error*, dan *individualizing the hyperplane*.

Fungsi Non-Linear *SVR* dapat dilihat pada persamaan (6) dan *Kernel polynomial* dan *radial* dapat dilihat pada persamaan (7) dan persamaan (8) secara berurutan [12].

$$\hat{y} = b_0 + \sum_{i=1}^N b_i K(X_i, X) \quad (6)$$

$$K(X_i, X_j) = (X_i \cdot X_j)^d \quad (7)$$

$$K(X_i, X_j) = \exp\left(-\frac{\|X_i - X_j\|^2}{2\sigma^2}\right) \quad (8)$$

Dimana K adalah kernel, d adalah pangkat polinom.

D. Evaluasi

Beberapa metode pengukuran evaluasi diajukan untuk menentukan keabsahan penelitian yang akan dilakukan. Metode pengukuran evaluasi yang diajukan adalah *mean square error*, *coefficient determination*, dan *explained variance score*.

Coefficient determination adalah proporsi dari variance dalam variabel *dependent* terhadap variabel *independent*. Persamaan *coefficient determination* dapat dilihat pada persamaan (9) [5]:

$$r^2 = 1 - \frac{SS_{res}}{SS_{total}} \quad (9)$$

Dimana r^2 adalah *coefficient determination*, SS_{res} dan SS_{total} adalah *sum of square residual* dan total. Rumusan *sum of square* dapat dilihat pada persamaan (10) dan (11) [5]:

$$SS_{tot} = \sum_i (y_i - \hat{y})^2 \quad (10)$$

$$SS_{res} = \sum_i (y_i - f_i)^2 \quad (11)$$

Dimana f_i adalah prediksi dari y dan \hat{y} adalah *mean of target*.

Mean square error adalah estimator yang menghitung rata-rata dari *squares of error*. Rumusan mean square error dapat dilihat pada persamaan (12) [5].

$$MSE = \frac{1}{n} \sum_1^n (Y_i - \hat{Y}_i) \quad (12)$$

Explained variance score adalah sebuah nilai untuk mengukur variance model terhadap data yang diberikan [5].

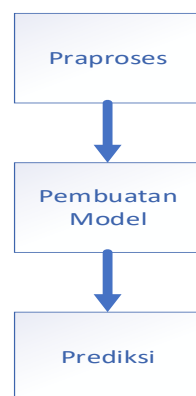
$$EVS = 1 - \left(\frac{Var(\hat{Y} - Y)}{Var(y)} \right) \quad (13)$$

Dimana Var adalah variance.

III. EKSPERIMEN

Metodologi penelitian ini dapat digambarkan dengan bagan seperti pada **Gambar 1**. Tahapan yang dilakukan pada penelitian ini adalah preprocessing data dan pembuatan model dengan menggunakan beberapa metode regresi seperti *Huber Estimation*, *Theil-Sen Estimation*, *Ordinary Least Square*, *Support Vector Regression* dan *Random Sample Consensus (RANSAC) estimation* dan selanjutnya akan dilakukan prediksi

Data merupakan data yang diambil di lapangan, data diambil di Kabupaten Sarolangun, Provinsi Jambi. Data diambil dengan sebuah alat yaitu sonolog. Sonolog adalah sebuah alat yang digunakan pada sumur-sumur minyak untuk mengetahui ketinggian fluida dengan peralatan *acoustic* dengan bantuan peralatan penghasil gelombang/ *acoustic pulse generator*.



Gambar 3. Skema Eksperimen

A. Tahapan Preprocessing

Pada tahapan ini data yang didapat akan melalui proses pemilihan fitur yang dimana fitur yang terdiri atas:

1. Well
2. Barrel, Size OD
3. Barrel, Size ID
4. Plunger Size OD
5. Plunger Clearance
6. Plunger length (inch)
7. Plunger length (feet)
8. API
9. Sg Air
10. Sg Gas
11. BHP After Perfo
12. BHP Current
13. PIP Current
14. Qtest
15. Vdynamic

Dari fitur-fitur diatas yang digunakan adalah *Plunger Clearance* dan *Vdynamic*. Fitur ini dipilih disebabkan *Plunger clearance* dan *Vdynamic* memiliki keterikatan dalam menentukan Slippage.

Selanjutnya adalah menghapus data yang dianggap memiliki *outlier* dalam hal ini data-data yang memiliki fitur bernilai kosong akan dihapus. Selanjutnya proses normalisasi yang dilakukan pada fitur maupun nilai *slippage* nya, hal ini dilakukan untuk menormalisasikan nilai menjadi antara 0 sampai dengan 1.

B. Pembuatan Model

Pada proses pembuatan model akan dilakukan perbandingan beberapa metode regresi dan juga metode *Support Vector Regression*.

Pembuatan model dilakukan untuk menentukan sebuah model matematika yang tepat digunakan dalam memprediksi hasil *slippage*. Metode yang digunakan adalah *Huber Estimation*, *Theil-Sen Estimation*, *Ordinary Least Square*, *Support Vector Regression* dan *Random Sample Consensus (RANSAC) estimation*. Dimana semua metode tersebut merupakan metode yang mengabaikan *outlier* data yang diambil.

C. Pengukuran Evaluasi

Pada penelitian ini ada beberapa kondisi yang akan digunakan, salah satunya proses pelatihan dan pengujian akan menggunakan konsep 80% data untuk proses pelatihan dan 20 % data akan digunakan untuk proses pengujian. Proses ini akan dilakukan secara acak.

Pengukuran evaluasi yang dilakukan adalah menggunakan *mean square error*, *coefficient determination*, dan *explained variance score*.

IV. HASIL EKSPERIMEN

Pada penelitian ini akan menggunakan beberapa metode regresi dan perhitungan evaluasi. Tabel 1 – Tabel 3 memberikan hasil dari evaluasi menggunakan MSE, r^2 , dan *explained variance score*.

Tabel 1. Hasil MSE berbagai metode regresi

| Metode | MSE |
|------------|--------------|
| OLS | 0.032 |
| Huber | 0.031 |
| Theil-Sen | 0.031 |
| RANSAC | 0.04 |
| SVR | 0.029 |

Tabel 2. Hasil r^2 berbagai metode regresi

| Metode | r^2 |
|------------|---------------|
| OLS | -0.224 |
| Huber | -0.185 |
| Theil-Sen | -0.195 |
| RANSAC | -0.555 |
| SVR | -0.131 |

Tabel 3. Hasil *explained variance score* berbagai metode regresi

| Metode | EVS |
|------------|---------------|
| OLS | -0.214 |
| Huber | -0.151 |
| Theil-Sen | -0.185 |
| RANSAC | -0.509 |
| SVR | -0.001 |

Berdasarkan hasil evaluasi yang didapatkan pendekatan dengan konsep linear seperti metode regresi linear (OLS, Huber, Theil-Sen, dan RANSAC) memberikan hasil yang tidak memuaskan dibandingkan dengan menggunakan pendekatan non-linear seperti SVR. Hal ini dapat dilihat pada Tabel 1- Tabel 3 bahwa metode non-linear seperti SVR memberikan nilai MSE yang paling rendah sehingga tingkat kesalahan yang dihasilkan lebih sedikit dibandingkan dengan metode linear yang diusulkan dan memiliki nilai r^2 dan *Explained variance score* tertinggi karena mendekati dengan probabilitas 1.

V. SIMPULAN DAN PENELITIAN SELANJUTNYA

Berdasarkan penelitian ini dapat dinyatakan bahwa metode regresi merupakan salah satu metode yang baik digunakan dalam memprediksi *slippage value*. Hasil terbaik didapatkan oleh menggunakan SVR

berdasarkan evaluasi hasil. Dengan adanya prediksi yang akurat pada perhitungan *slippage* diharapkan dapat memberikan gambaran lubrikasi yang tepat pada penggunaan pompa *plunger* yang dipakai. Hal ini tentu akan memberikan hasil optimasi produksi minyak yang baik tanpa ada gangguan seperti *sticking* alat jika persentase *slippage* yang ada *overpredict* atau terjadinya gesekan plat antar *plunger* dan *barrel* apabila persentase *slippage* yang ada kurang dari seharusnya. Pada penelitian selanjutnya akan dicoba menggunakan konsep *ensemble* terhadap nilai yang didapatkan.

UCAPAN TERIMA KASIH

Penulis ingin mengucapkan terimakasih untuk Felix Indra Kurniadi dari Universitas Tanri Abeng untuk memberikan masukan dan penjelasan mengenai komputasi dari setiap metode regresi

DAFTAR PUSTAKA

- [1] B. H. Robinson, "Economics of Pumping," *Drilling and Production Practice*. American Petroleum Institute, New York, New York, p. 6, 1935.
- [2] R. . Reekstin, "Minimizing Slippage in Subsurface Pumps," *Pet. Eng.*, vol. 32, pp. B23–B26, 1960.
- [3] R. Kyle Chambliss, L. Heinze, and R. M Sweazy, *PLUNGER LEAKAGE AND VISCOUS DRAG FOR BEAM PUMP SYSTEMS*. 2001.
- [4] B. Guo, *Petroleum Production Engineering, A Computer-Assisted Approach*, 1st Editio. Gulf Professional Publishing, 2007.
- [5] P. J. Rousseeuw and A. M. Leroy, *Robust Regression and Outlier Detection*. New York, NY, USA: John Wiley & Sons, Inc., 1987.
- [6] Scikit-learn, "Huber Regression." [Online]. Available: http://scikit-learn.org/stable/modules/generated/sklearn.linear_model.HuberRegressor.html#sklearn.linear_model.HuberRegressor. [Accessed: 30-Oct-2018].
- [7] P. J. Huber, J. Wiley, and W. InterScience, *Robust statistics*. Wiley New York, 1981.
- [8] X. Dang, H. Peng, X. Wang, and H. Zhang, "Theil-Sen Estimators in a Multiple Linear Regression Model." 2009.
- [9] Scikit-learn, "Theil-Sen estimator: generalized-median-based estimator." [Online]. Available: http://scikit-learn.org/stable/modules/linear_model.html#theil-sen-regression. [Accessed: 30-Oct-2018].
- [10] M. A. Fischler and R. C. Bolles, "Random Sample Consensus: A Paradigm for Model Fitting with Applications to Image Analysis and Automated Cartography," *Commun. ACM*, vol. 24, no. 6, pp. 381–395, Jun. 1981.
- [11] Wikipedia, "Random sample consensus." [Online]. Available: https://en.wikipedia.org/wiki/Random_sample_consensus. [Accessed: 30-Oct-2018].
- [12] S. Sayad, "Support Vector Machine - Regression." [Online]. Available: http://www.saedsayad.com/support_vector_machine_reg.htm. [Accessed: 31-Oct-2018].