ROBUST PARAMETER ESTIMATION IN ROD PUMPED SYSTEMS

Vladimir Pechenkin and Biplav Chapagain
DV8 Energy Inc.

INTRODUCTION

Modern controllers are required to estimate various parameters from field data to provide effective diagnostics and control of sucker rod pumping installations. In some cases, however, the data are not only corrupted by noise but also contain *outliers* that are in gross disagreement with the postulated model. If included, outliers can distort the fitting process so dramatically that the fitted parameters become arbitrary. In such circumstances, the deployment of robust estimation methods is essential. This paper discusses the application of *random sampling consensus* (RANSAC) to rod pump systems. The method can identify the outliers even when they constitute up to 50% of data.

RANDOM SAMPLING CONSENSUS

This approach proposed in [1] is well described in [2], which we follow. Given that a significant portion of the data may be outliers, RANSAC is the opposite to conventional smoothing techniques. Instead of using as much data as feasible to obtain an initial solution and then attempting to identify outliers, the smallest possible subset of the data to estimate the parameters is selected. For example, two points are randomly picked for a straight line. This process is repeated enough times on different subsets to ensure that at least one of the subsets contains only good points with high probability *p*, say, 95% or 99%. The best solution is the one that maximizes the number of points that are "in agreement" with the model. Once the outliers are identified and removed, the remaining set of *inliers* is combined to generate the final solution.

Ideally, every possible subset of the data would be considered. If this is computationally infeasible, an important question is how many times random sampling must be repeated. Ref. [2] gives a simple formula that relates the number of subsets *m* to the fraction of contaminated data *ε*: $p = 1 - (1 - (1 - \varepsilon)^n)^m$. Here n is the number of features in each sample. If *ε* is unknown, as is usually the case, an educated worst-case estimate must be made in order to determine *m*. The criterion that distinguishes inliers from outliers will be elaborated on in the following section.

MODEL EXAMPLE

To illustrate the method, consider a straight green line $y = ax + b$ in Figure 1, where $a = 1.5$ and $b = 10$. Suppose that we have obtained $N = 200$ measurement points but 50% of them are gross outliers. Those are represented by red circles randomly scattered all over the plot. The remaining 100 measurements are "good" in a sense that they are slightly perturbed by Gaussian noise only. Those are depicted as blue triangles. The estimation algorithm knows neither the number of the outliers nor the standard deviation of the noise *σ*. The worst-case scenario $\varepsilon = 0.5$ is assumed. RANSAC selects a pair of points $m = 17$ times to ensure that at least one selection contains two inliers with probability $p = 0.99$. For each pair a set of *N* Euclidean distances d_i from every point to the calculated line is recorded. Point i is classified as an inlier if $d_i < 1.96\sigma$ and as an outlier otherwise. The standard deviation is estimated from the *median* of the residuals: $\sigma = 1.4826\sqrt{med_i(d_i)}$ [2]. The black line in Figure 1 represents the result of a particular RANSAC run. 95 points were classified as inliers and the final estimation of the coefficients of the straight line was made by using only those points.

APPLICATION TO ROD PUMPING: TRAVELING VALVE CHECK

The problem that motivated this work is the estimation of the plunger leakage rate from the traveling valve check. This method proposed in [3] involves stopping the pumpjack during the upstroke when the weight of the fluid column is carried by the traveling valve. The surface $F_s(t)$ load is measured and recorded as a function of time. The obtained curve is then processed to extract the pump leakage rate from the polished rod load loss rate. The original work [3] attempted to estimate the leakage by picking just three samples and then passing a parabola through them. The first point is always at time $t = 0$ corresponding to the maximum load $F_s(0)$. The other two are selected arbitrarily, however. Further research [4] proposed a transformation of the load curve with the intention to pass a straight line through all the samples via linear regression. In both cases though a noisy $F_s(t)$ signal presented a significant challenge.

In this work we first refine the model of [3,4] rewriting the equations that relate the traveling valve leakage rate Q to the axial load at the pump $F_p(t)$. As shown in [4], Q may be expressed in two different ways.

$$Q = -\frac{A_p}{K}\frac{dF_p(t)}{dt} = C_x \frac{F_p}{A_p}$$

Here A_p is the plunger area, *K* is the spring constant and C_x characterizes the leakage rate and is the quantity to be estimated. Integration of this differential equation yields the following result:

$$F_p(t) = F_p(0)\exp\left(-\frac{KC_x t}{A_p}\right)$$

The next step is to relate the load on pump to the load at the polished rod. When dynamic effects are neglected, the relationship between $F_s(t)$ and $F_p(t)$ is given by [3,4]: $F_s(t) = F_p(t) + W_r + F_{up}$. Here W_r is the buoyant rod weight and F_{up} is the static mechanical friction on the upstroke. This equation holds for any time instant including $t = 0$. Rearranging the terms gives the following two equations:

$$F_s(0) - W_r - F_{up} = F_p(0)$$
$$F_s(t) - W_r - F_{up} = F_p(0)\exp\left(-\frac{KC_x t}{A_p}\right)$$

Eliminating $F_p(0)$ and taking the logarithm of both sides then yields

$$\ln\frac{F_s(0) - W_r - F_{up}}{F_s(t) - W_r - F_{up}} = \frac{KC_x t}{A_p}$$

In general, W_r is supposed to be known and F_{up} may be calculated. If the test is carried on long enough though, the traveling valve eventually opens and the fluid load on pump vanishes. Denote this case symbolically as $t = \infty$ and $F_s(\infty)$ as F_∞, then $F_\infty = W_r + F_{up}$ and

$$\ln(F_s(0) - F_\infty) - \ln(F_s(t) - F_\infty) = KC_x t/A_p$$

This is an equation of a straight line that may be rewritten in a more canonical form as

$$\ln\left(\frac{F_s(t)}{F_\infty} - 1\right) = -\frac{KC_x t}{A_p} + \ln\left(\frac{F_s(0)}{F_\infty} - 1\right)$$

RANSAC can then be used to estimate C_x from the samples of $F_s(t)$. Since the number of parameters is $n = 1$, it may be feasible to go through all the samples and select the best one.

DYNAMOMETER CARD ANALYSIS

Surface and downhole dynamometer cards are used routinely for diagnostics and control. Reference [5] presents several interesting cases where the slope of the surface card is utilized to estimate the spring constant *K* and diagnose various important conditions. Those include downhole rod string sticking, unanchored tubing as well as a loss of synchronization between the load and the position signals. The cards are often “noisy” for a variety of reasons and therefore extracting useful information from them may be challenging. Application of RANSAC following the steps of the model example may thus be quite helpful.

Downhole cards can be utilized to estimate the fluid load on pump, which requires the calculation of the fluid load lines. Reference [6] discusses a statistical approach to such computations, which is similar in spirit to RANSAC. Since the load lines are horizontal, they are again characterized by just one parameter. The potential benefit of RANSAC in this case is an estimate of the standard deviation, which determines a set of points to be used to obtain the final solution.

CONCLUSION

RANSAC is a powerful robust estimation algorithm that originated in the areas of computer vision and statistics. We believe that it has potential in rod pump applications as well, providing more accurate results and in some cases opening a door to working with noisy data that cannot be meaningfully processed with conventional tools.

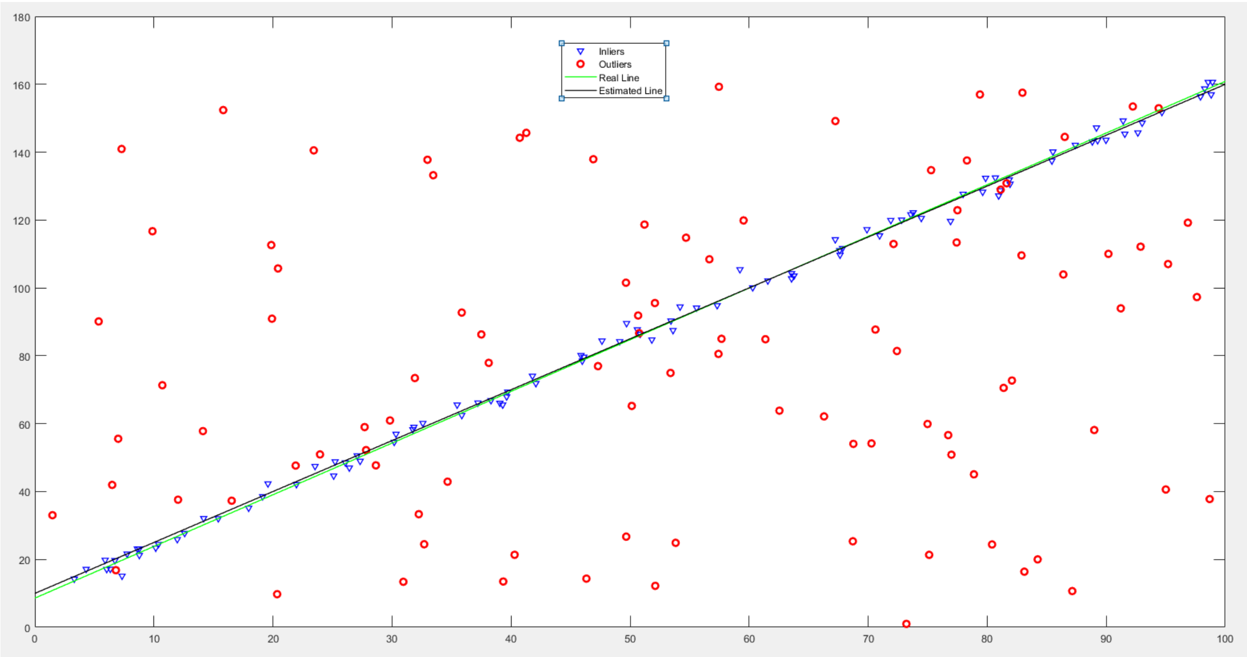

Figure 1 – RANSAC estimation of a straight line in the presence of 50% of gross outliers.

REFERENCES

[1] M.A. Fischler and R.C. Bolles, “Random Sample Consensus: A Paradigm for Model Fitting with Application to Image Analysis and Automated Cartography,” Communication of the Association for Computer Machinery, vol. 24(6), pp. 381–395, Jun. 1981.

[2] P.H.S. Torr and D.W. Murray, “The development and Comparison of Robust Methods for Estimating the Fundamental Matrix,” International Journal of Computer Vision, vol. 24(3), pp. 271–300, Sep. 1997.

[3] K.B. Nolen and S.G. Gibbs, “Quantitative Determination of Rod-Pump Leakage with Dynamometer Techniques,” SPE Production Engineering, vol. 5(3), pp. 225–230, Aug. 1990.

[4] T.M. Mills, “Method for Measuring Leakage Rate and Inferring Production Rate of an Oilfield Downhole Pump,” US Patent 8,849,594, Sep. 2014.

[5] O.L. Rowlan, C. Haskins, C.A. Taylor and K. Skinner, “Examples of Forces not Accounted for by the Wave Equation,” Southwestern Petroleum Short Course, 2018.

[6] V.M. Pons, “Fluid Load Line Calculation and Concavity Test for Downhole Pump Card,” US Patent 9,810,212, Nov. 2017.