Bayero Journal of Pure and Applied Sciences, 11(1): 459 - 464
ISSN 2006 - 6996

SELF-OPTIMIZING CONTROL METHODOLOGY FOR MIXED INTEGER PROGRAMMING PROBLEMS: A CASE STUDY OF REFINERY PRODUCTION SCHEDULING

Aminu A. Hamisu[a,b,*], Yi. Cao[b] and Adamu S. Girei[b]
[a]Department of Chemical Engineering Ahmadu Bello University, Zaria, Nigeria.
[b]Process Systems Engineering Group, Cranfield University, Bedfordshire, UK
[*]Corresponding author: alamin446@gmail.com

ABSTRACT
Problem formulation as mixed integer nonlinear programming (MINLP) is one of the most challenging task in refinery scheduling optimization. In most of the work reported in refinery scheduling, uncertainties from design point of view predominate. However, there is also a need to consider operational uncertainties (disturbances) as they affect the accuracy and robustness of the overall schedule. This study proposed a novel approach under self-optimizing control (SOC) framework to deal with multi-period refinery scheduling problems under uncertain conditions. The goal is to maintain global optimum by controlling the gradient of the cost function at zero via approximating necessary conditions of optimality (NCO) over the whole uncertain parameter space. A regression model for the plant expected revenue (profit) as a function of independent variables using optimal operation data was obtained and a feedback input (manipulated variable) was derived. The performance of the proposed approach was tested using case studies. The first case assumed a system with no disturbance with the base case model giving an optimal profit of $56,696,407 while the proposed approach yields $50,523,054, translating to 10.888 % loss. The percentage loss for the second, third and fourth cases with disturbances are 5.807 %, 4.409% and 7.898% respectively. The results obtained have shown that the idea presented was able to effectively deal with the situation at hand with percentage loss within a reasonable degree.
Keywords: Refinery scheduling, MINLP formulation, Operational uncertainty (disturbances), Necessary condition of optimality, Feedback control

INTRODUCTION

Most chemical process plants including refinery are operated in such a way that operators set decision variables as set points and with the aid of proportional integral derivative (PID) controllers the set points are kept at their desired values. To satisfy requirements set by environmental laws, operate within safety limits, survive market competition and to meet tighter quality specifications of products; plants must operate near optimal. Self-optimizing control (SOC) as a strategy helps to achieve the aforementioned objectives by selecting appropriate control variables (CVs) so that when they are maintained at their set point values, the overall plant operation is optimal or near optimal even in the presence of disturbances (Skogestad, 2000).

When a disturbance is introduced into a chemical plant, measurements are taken and control actions are implemented to compensate for the effects of the disturbance. In the past decades, several techniques for CV selections have been reported. Most of these techniques require process models to determine CVs offline and largely depend on the ability to linearize nonlinear models around their nominal operating points. This procedure is time-consuming which results in the plant operation being locally optimal and become impractical where no process model is available (Kariwala, 2007; Alstad and Skogestad, 2007).

Despite the benefits of applying SOC, there are challenges emanating from the CV selection. The selected CVs must be such that they give acceptable loss and therefore able to avoid any need to re-optimize set points when disturbances are introduced into the system. Difficulty arising from using model for SOC has been overcome by incorporating measurements in the optimization framework. Single measurements or combination of measurements may be used as CVs. Halvorsen et al. (2003) introduced methods for finding subset or combination of measurements as controlled variables. Controlling these subset or combination of measurement at constant set points implies operating the plant at desired economic condition. Overcoming challenges due to model linearization requires a global approach.

Necessary condition of optimality (NCO) is a viable complementary method that seeks to overcome the local shortcomings of the existing SOC methodologies (Jäschke and Skogestad, 2011). François et al. (2005) are of the view that measurements can be used to enforce NCO in the presence of disturbance variables (uncertain parameters) where the NCO are separated into active constraints and cost sensitivities (gradients). Owing to the fact that some NCO components are non-measurable online due to the presence of disturbance (s) in the objective function, Ye et al. (2012b) proposed that CVs can be selected in such a way that they approximate unmeasured NCO over the whole uncertain parameter space. The CVs can then be obtained through regression methods. The CV selection problem is therefore transformed into a regression problem and does not need a model to be a priori (Ye et al., 2012a). The difficulty using NCO lies in the inability to compute the gradient online.

Recently, a methodology was developed by Girei et al. (2014) that computes CVs as function of measurements from real plant or simulated data using finite difference approximation. Grema and Cao (2014) extended the methodology to dynamical systems where the gradient is approximated using Taylor series expansion. Their approach is not directly applicable for problems involving mixed integer programming with multiple time periods and therefore a new methodology has to be developed. Development of methodology to handle mixed integer problems is the motivation behind this paper.

Therefore in this study, a multi-period data driven approach involving mixed integer problems to determine CV as a function of measurements is presented. The methodology is then applied to refinery scheduling problem with uncertainties in crude oil composition.

1. Data Driven Self-optimizing Control for Scheduling

Although the methodology is developed to deal with mixed integer problems, the discussion here will be mainly on refinery production scheduling. Generally, scheduling is a static mixed integer optimization problem with uncertainties in model parameters. The problem can be formulated as:

$$\min_{\mathrm{u}} J(\mathbf{u}, \mathbf{d}) \quad (1)$$

subject to

$$\mathbf{g}(\mathbf{u}, \mathbf{d}) \leq 0 \qquad \mathbf{u}_b: (\mathbf{0}, \mathbf{1}) \quad (2)$$

Where J is an objective function to be minimized (cost or negative profit). As the problem is mixed integer, the control inputs are separated into continuous manipulated variables $\mathbf{u_c}$ and integer manipulated variables $\mathbf{u_b}$ with $\mathbf{u_c}, \mathbf{u_b} \in \mathbb{R}^{n_u}$. The integer manipulated variables range from 0 to 1 and $\mathbf{d} \in \mathbb{R}^{n_d}$ are the uncertain parameters or disturbances. $\mathbf{g}: \mathbb{R}^{n_u} X\ \mathbb{R}^{n_d} \rightarrow \mathbb{R}^{n_g}$ are the constraints to be satisfied, which are usually related to unit capacities, mass balance, inventory, and storage. The variables $\mathbf{u_b}$ here are typically chosen to control the continuous variables such as flowrates by either forcing one or more variables to be in between 0 and 1. For simplicity, a single manipulated variable that is varying with time periods is assumed. For t time periods and y measurements, the objective function in Equation 1 can be transformed into

$$J = \sum_t J_t(u_t, \boldsymbol{y_t}, \boldsymbol{d_t}) \qquad \forall\, t \quad (3)$$

Where J_t is the contribution of J in each time period or event point t. $u_t, \boldsymbol{y_t}$, and $\boldsymbol{d_t}$ are manipulated variables, measurements, and disturbances at time period t respectively. The scheduling horizon H is discretized into time periods t_n $(n = 1, 2, 3, \dots, N)$ of variable lengths LT_n. Variable time discretization is based on the fact that events or activities do not always happened at the time boundaries. i.e. for scheduling horizon of 10 days discretized into 10 time periods, some task can take less or more than 1 day to complete. To obtain CVs, the following procedures are followed.

i. Manipulated variables u are identified along with flow streams y that disturbances will have an impact upon. The flow streams are the measurements. The scheduling model is then solved to obtain solution vector

- $u_{0,t} = u_{0,1}, u_{0,2}, u_{0,3}, \dots, u_{0,N}$ for the manipulated variables,
- $\boldsymbol{y_{0,t}} = \boldsymbol{y_{0,1}}, \boldsymbol{y_{0,2}}, \boldsymbol{y_{0,3}}, \dots, \boldsymbol{y_{0,N}}$ for the measurements and
- J_0 as profit (or cost).

The solution from the first simulation run represents the nominal schedule.

ii. The manipulated variables are then slightly but randomly perturbed for the whole time periods or event points to have

$$u_{i,t} = u_{i,1}, u_{i,2}, u_{i,3}, \dots, u_{i,N}\ , \qquad i = 1, 2, 3, \dots, I \quad (4)$$

and the scheduling model is simulated for i trajectories to obtain measurements

$$\boldsymbol{y_{i,t}} = \boldsymbol{y_{i,1}}, \boldsymbol{y_{i,2}}, \boldsymbol{y_{i,3}}, \dots, \boldsymbol{y_{i,N}} \quad (5)$$

and (cost or negative profit) .

y is a vector because at each time period there may be more than one measurement $(m = 1, 2, 3, \dots, M)$ with the manipulated variable included as one of the measurements.

iii. The gradient (change in objective function with respect to manipulated variables) is the CV and approximated using Taylor series expansion.

The detail derivation is presented in the Appendix section.

Case study

The refinery scheduling model and operational data in Hamisu (2015) are adopted here to show the applicability of the proposed solution approach. However; here the binary variable $XC_{cr,t}$ is allowed to take any value between 0 and 1. The optimal profit J_0 in this case was obtained as $56,696,407. The optimal values of the variables $XC_{cr,t}$ are used as parameters in the SOC model. The optimal parameter values of crude 2 (manipulated variables) are then perturbed slightly but randomly around their nominal operating points to form sequence of solutions to be used for regression analysis. 20x9 (180) data points were generated in accordance with Equation 11 (see Appendix) from which 8 regression coefficients are obtained. The regression coefficients determined are presented in Table 1.

Table 1: Parameter values from regression

Coefficient	Value (x 10^4)
θ_0	6.7996
θ_1	0.1396
θ_2	0.4154
θ_3	0.0430
θ_4	-0.2164
θ_5	-0.1722
θ_6	-0.1252
θ_7	-5.9679

This gives the optimal feedback control law:

$$u_{fb,t} = (1/5.9679)[6.7996 + 0.1396y_{1,t} + 0.4154y_{2,t} + 0.0430y_{3,t} - 0.2164y_{4,t} - 0.1722y_{5,t} - 0.1252y_{6,t}] \quad \textbf{(6)}$$

The measurements $y1 - y6$ are straight run (SR) fuel gas, SR gasoline, SR naphtha, SR distillate, SR gas oil, and SR residuum streams of crude 1 respectively. Cases with and without disturbances are considered to illustrate the capability of the proposed approach.

Case 1

This first case considered a problem with no disturbance introduced into the system. The optimal objective value of the base case model with nominal values of the manipulated variables is compared with the objective value obtained after feedback implementation. The optimal profit for the base case model was obtained as $56,696,407. Implementing Equation 17 for the base case model gives an optimal profit of $50,523,054. The loss was computed in accordance with Equation 16 to obtain a value of 10.888%. Scenarios in the next case will better illustrate the advantage of the proposed SOC methodology.

Case 2

Scenario A

This scenario considered a change in composition of crude 1 by 5% for the whole scheduling horizon. The measurements representing cut fractions from crude 1 are taken and the optimal manipulated variable is computed and implemented in the SOC model to obtain optimal profit $J(u_{fb}, d)$ of $53,403,869. Compared with the optimal value of $56,696,407, a loss of 5.807 % was obtained. The production levels of the cut fractions using SOC are compared with the actual amounts produced at the nominal operating conditions. These are shown in Figures 1 to 6.

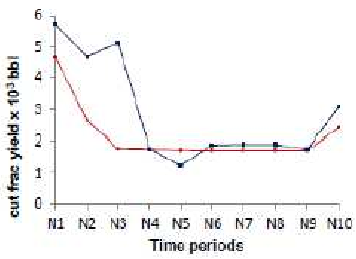

Figure 1: Production levels for SR fuel gas at different time periods

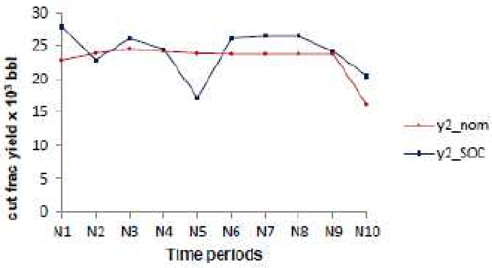

Figure 2: Production levels for SR gasoline at different time periods

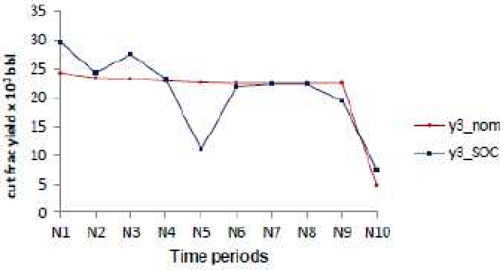

Figure 3: Production levels for SR naphtha at different time periods

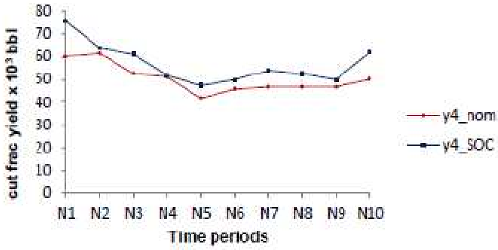

Figure 4: Production levels for SR distillate at different time periods

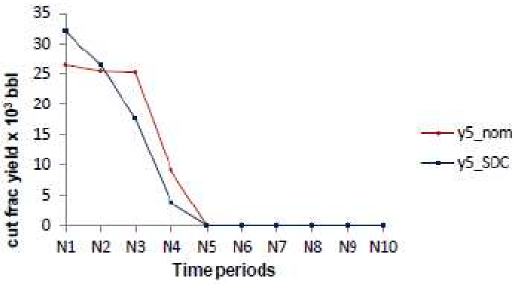

Figure 5: Production levels for SR gas oil at different time periods

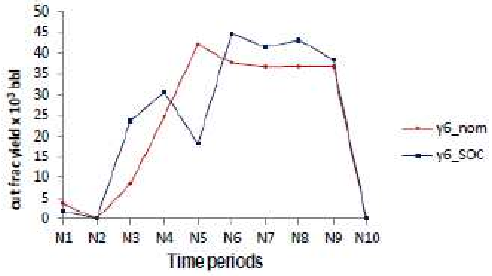

Figure 6: Production levels for SR residuum at different time periods

With the exception of SR fuel gas and SR distillate streams from the figures above, other cut fractions have their production levels approaching minimal values at the end of the scheduling horizon. Deviations from true optimal values due to feedback control for the cut fractions is not unconnected with the fact that refinery production has so many constraints to be satisfied. The amount produced depends on the CDU processing capabilities. Setting the upper limit for the crudes to be processed in CDU is an industrial practice that cannot be ignored. Mandating the plant to operate within this limit will therefore have an impact on the maximum permissible tuning that will allow the feedback implementation to restore the plant profit. Considering this limitation, the loss of 5.807% obtained is still a reasonable value.

Scenario B

In this scenario, a 3% change in composition is considered with the other information exactly same as in Scenario A. The profit due to feedback implementation strategy was obtained to be $54,196,473 against $56,696,407 for the true optimal value. This translates to 4.409% loss which is better compared to Scenario A. This improvement is due the magnitude of the disturbance being smaller in this scenario. Again, based on the reasons mentioned in the preceding scenario, this loss value of 4.409% is still within the acceptable range.

Scenario C

Here, there is an increase in composition by 9% on the first period then a drop of 4% is recorded on the fifth period. This scenario is more common in refineries where fluctuation do occur from one time period to another. The profit recorded due to SOC is $52,218,634. Comparing with the true optimal value gives a loss of 7.898%. For the first two time periods, the u_{fb} values are -0.0919 and -0.0444 respectively. These cannot be implemented because $\boldsymbol{u_b}$ must be between 0 and 1. Saturation is therefore applied here forcing the two values to be zero.

The loss value in this scenario is greater than those obtained in Scenarios A and B due to fact that fluctuation is more intense here with an abrupt change by 9% at the initial stage and a sudden drop by 4% on the fifth period.

In summary, the loss values are greater than 1% in all scenarios because the refinery plant is a complex with a multiple number of units interconnected and hence the units cannot be treated independently. Even though decisions from scheduling are implemented on a day-to-day basis, the schedule generation does not follow the same pattern. All schedules decisions no matter the length of the horizon have to be obtained at the same time instance for implementation at later dates.

CONCLUSION

One of the challenges in refinery scheduling is the generation of schedules with consideration to uncertainty in process parameters. Fluctuation in product demand, change in crude oil composition, and other uncertainties do manifest during execution of the schedules. In the presence of these uncertainties, the schedule generated under deterministic conditions may become infeasible, suboptimal or difficult to implement. This paper presents an efficient methodology under self-optimizing control framework to deal with disturbances for process plants with optimization model posed as mixed integer nonlinear programming formulation. The methodology computes CVs as function of measurements from simulated data using Taylor series expansion. The procedure went beyond addressing feasibility issues due to the influence of uncertain parameters but also ensure optimal or near optimal operation is maintained. The methodology was applied to refinery scheduling problem with uncertainties in crude oil compositions to come up with a feedback control law that compensates for the effect of the uncertainties. Through case studies, the idea presented was able to effectively deal with the situation at hand with percentage loss within a reasonable degree.

REFERENCES

Alstad, V. & Skogestad, S., 2007. Null space method for selecting optimal measurement combinations as controlled variables. *Industrial and Engineering Chemistry Research*, 46(3), pp.846–853.

François, G., Srinivasan, B. & Bonvin, D., 2005. Use of measurements for enforcing the necessary conditions of optimality in the presence of constraints and uncertainty. *Journal of Process Control*, 15(6), pp.701–712.

Girei, A. et al., 2014. Data-driven self-optimizing control. In *24th European Symposium on Computer Aided Process Engineering*. pp. 649–654.

Grema, A.S. & Cao, Y., 2014. Optimal feedback control for reservoir waterflooding. In *20th International Conference on Automation and Computing*. pp. 20–25.

Halvorsen, I.J. et al., 2003. Optimal Selection of Controlled Variables. *Industrial & Engineering Chemistry Research*, 42(14), pp.3273–3284.

Hamisu, A. A., Kabantiok, S. & Wang, M., 2013. Refinery scheduling of crude oil unloading with tank inventory management. *Computers & Chemical Engineering*, 55, pp.134–147.

Hamisu, A. A., Mohammad, A., Bilal, S., Abubakar, U. & Sha'aban, Y. A., (2016) "Model Development and Optimization of Refinery Production Scheduling", Techno Science Africana Journal Vol 13 (2), pp. 1-13.

Jäschke, J. & Skogestad, S., 2011. NCO tracking and self-optimizing control in the context of real-time optimization. *Journal of Process Control*, 21(10), pp.1407–1416.

Kariwala, V., 2007. Optimal measurement combination for local self-optimizing control. *Industrial and Engineering Chemistry Research*, 46(11), pp.3629–3634.

Kariwala, V., Cao, Y. & Janardhanan, S., 2008. Local self-optimizing control with average loss minimization. *Industrial and Engineering Chemistry Research*, 47(4), pp.1150–1158.

Skogestad, S., 2000. Plantwide control: The search for the self-optimizing control structure. *Journal of Process Control*, 10(5), pp.487–507.

Ye, L. et al., 2012a. A data-driven approach for selecting controlled variables. *Proceedings of 8th Intl. Symposium on …*, (2008), pp.904–909.

Ye, L. et al., 2012b. Approximating necessary conditions of optimality as controlled variables. *Industrial & Engineering Chemistry ….*

Appendix

Considering that the objective function is changing with respects to multiple manipulated variables corresponding to different time periods, the following equation holds true:

$$\begin{aligned} J_1 - J_0 &= CV_{1,1}\Delta u_{1,1} + CV_{1,2}\Delta u_{1,2} + CV_{1,3}\Delta u_{1,3} + \dots + CV_{1,N}\Delta u_{1,N} \\ J_2 - J_0 &= CV_{2,1}\Delta u_{2,1} + CV_{2,2}\Delta u_{2,2} + CV_{2,3}\Delta u_{2,3} + \dots + CV_{2,N}\Delta u_{2,N} \\ J_3 - J_0 &= CV_{3,1}\Delta u_{3,1} + CV_{3,2}\Delta u_{3,2} + CV_{3,3}\Delta u_{3,3} + \dots + CV_{3,N}\Delta u_{3,N} \\ &\vdots \\ J_I - J_0 &= CV_{I,1}\Delta u_{I,1} + CV_{I,2}\Delta u_{I,2} + CV_{I,3}\Delta u_{I,3} + \dots + CV_{I,N}\Delta u_{I,N} \end{aligned} \quad \textbf{(7)}$$

Where,

$$\begin{aligned} \Delta u_{1,1} &= (u_1 - u_0)_1,\ \Delta u_{1,2} = (u_1 - u_0)_2,\ \dots,\ \Delta u_{1,N} = (u_1 - u_0)_N \\ \Delta u_{2,1} &= (u_2 - u_0)_1,\ \Delta u_{2,2} = (u_2 - u_0)_2,\ \dots,\ \Delta u_{2,N} = (u_2 - u_0)_N \\ \Delta u_{3,1} &= (u_3 - u_0)_1,\ \Delta u_{3,2} = (u_3 - u_0)_2,\ \dots,\ \Delta u_{3,N} = (u_3 - u_0)_N \\ &\vdots \\ \Delta u_{I,1} &= (u_I - u_0)_1,\ \Delta u_{I,2} = (u_I - u_0)_2,\ \dots,\ \Delta u_{I,N} = (u_I - u_0)_N \end{aligned} \quad \textbf{(8)}$$

The control variables

$$\begin{bmatrix} CV_{1,1}, & CV_{1,2}, & CV_{1,3}, & \dots & CV_{1,N} \\ CV_{2,1}, & CV_{2,2}, & CV_{2,3}, & \dots & CV_{2,N} \\ CV_{3,1}, & CV_{3,2}, & CV_{3,3}, & \dots & CV_{3,N} \\ & & \vdots & & \\ CV_{I,1}, & CV_{I,2}, & CV_{I,3}, & \dots & CV_{I,N} \end{bmatrix}$$

are non-measurable online and therefore can be replaced with measurement functions. Thus,

$$\begin{aligned} CV_{1,1} &= (\theta_0 + \theta_1 y_{1,1} + \theta_2 y_{1,2} + \theta_3 y_{1,3} + \dots + \theta_M y_{1,M})_1 \\ CV_{1,2} &= (\theta_0 + \theta_1 y_{1,1} + \theta_2 y_{1,2} + \theta_3 y_{1,3} + \dots + \theta_M y_{1,M})_2 \\ CV_{2,1} &= (\theta_0 + \theta_1 y_{2,1} + \theta_2 y_{2,2} + \theta_3 y_{2,3} + \dots + \theta_M y_{2,M})_1 \\ &\vdots \\ CV_{I-1,N} &= (\theta_0 + \theta_1 y_{I-1,1} + \theta_2 y_{I-1,2} + \theta_3 y_{I-1,3} + \dots + \theta_M y_{I-1,M})_N \\ CV_{I,N-1} &= (\theta_0 + \theta_1 y_{I,1} + \theta_2 y_{I,2} + \theta_3 y_{I,3} + \dots + \theta_M y_{I,M})_{N-1} \\ CV_{I,N} &= (\theta_0 + \theta_1 y_{I,1} + \theta_2 y_{I,2} + \theta_3 y_{I,3} + \dots + \theta_M y_{I,M})_N \end{aligned} \quad \textbf{(9)}$$

Substituting Equation 8 into Equation 6 gives

$$\begin{aligned} J_1 - J_0 &= \theta_0 (x_0)_1 + \theta_1 (x_1)_1 + \theta_2 (x_2)_1 + \dots + \theta_M (x_M)_1 \\ J_2 - J_0 &= \theta_0 (x_0)_2 + \theta_1 (x_1)_2 + \theta_2 (x_2)_2 + \dots + \theta_M (x_M)_2 \\ J_3 - J_0 &= \theta_0 (x_0)_3 + \theta_1 (x_1)_3 + \theta_2 (x_2)_3 + \dots + \theta_M (x_M)_3 \\ &\vdots \\ J_I - J_0 &= \theta_0 (x_0)_I + \theta_1 (x_1)_I + \theta_2 (x_2)_I + \dots + \theta_M (x_M)_I \end{aligned} \quad \textbf{(10)}$$

Where,

$$\begin{aligned} (x_0)_1 &= (\Delta u_{1,1} + \Delta u_{1,2} + \Delta u_{1,3} + \dots + \Delta u_{1,N})_1 \\ (x_1)_1 &= (y_{1,1}\Delta u_{1,1} + y_{1,2}\Delta u_{1,2} + y_{1,3}\Delta u_{1,3} + \dots + y_{1,N}\Delta u_{1,N})_1 \\ (x_2)_3 &= (y_{2,1}\Delta u_{2,1} + y_{2,2}\Delta u_{2,2} + y_{2,3}\Delta u_{2,3} + \dots + y_{2,N}\Delta u_{2,N})_3 \\ &\vdots \\ (x_M)_I &= (y_{M,1}\Delta u_{I,1} + y_{M,2}\Delta u_{I,2} + y_{M,3}\Delta u_{I,3} + \dots + y_{M,N}\Delta u_{I,N})_I \end{aligned} \quad \textbf{(11)}$$

Equation 9 can be re-arranged to

$$\Delta J_i = \boldsymbol{x}_i \boldsymbol{\theta} \quad \textbf{(12)}$$

Equation 11 can then be re-written as

$$\boldsymbol{Y} = \boldsymbol{X}\boldsymbol{\theta} \quad \textbf{(13)}$$

Implying that ΔJ_i is represented by vector $\boldsymbol{Y}$ and $\boldsymbol{x}_i$ by vector $\boldsymbol{X}$.

Using regression $\boldsymbol{\theta}$ can be determined.

$$\hat{\boldsymbol{\theta}} = (\boldsymbol{X}^T\boldsymbol{X})^{-1}\boldsymbol{X}^T\boldsymbol{Y} \quad \textbf{(14)}$$

Using the rule of thumb, the data points for regression should be at least ten times the number of coefficients to be estimated.

By analogy, Equation 8 can be represented in condensed form as

$$CV = \theta_0 + \theta_1 y_1 + \theta_2 y_2 + \theta_3 y_3 + \dots + \theta_M y_M \quad \textbf{(15)}$$

Controlling gradient at zero implies the LHS of Equation 14 equal to zero. It is important to note that $y_{M,t}$ is the manipulated variable as mentioned in step 2 of the solution procedure.

Therefore at each time period, the optimal feedback control input is obtained as:

$$u_{fb,t} = -\frac{1}{\theta_M}[\theta_0 + \theta_1 y_1 + \theta_2 y_2 + \theta_3 y_3 + \dots + \theta_{M-1} y_{M-1}] \quad \textbf{(16)}$$

Some values of the integer manipulated variables may be slightly below 0 or above 1. In such a case the variables are said to be 'saturated' and a constraint has to be imposed, forcing the saturated variables to be equal to their corresponding nearest value (0 or 1).

Implementing this feedback strategy in a close loop fashion will incur loss. The loss can be computed as

$$L = \frac{J_0 - J(\mathrm{u}_{fb}, \mathrm{d})}{J_0} \times 100 \quad \textbf{(17)}$$

Where J_0 is the true optimal J, while $J(\mathrm{u}_{fb}, \mathrm{d})$ is the objective function corresponding to implementing Equation 15 to maintain the CV at zero.