On-line optimization of a crude unit heat exchanger network

Tore Lid *
Statoil Mongstad
N-5954 Mongstad

Stig Strand
Statoil Research Centre
Arkitekt Ebbellsvei 10
N-7005 Trondheim

Sigurd Skogestad
Department of Chemical
Engineering, NTNU
N-7491 Trondheim

November 30, 2000

Abstract

This paper describes modeling and on-line optimization of a crude unit heat exchanger network at the Statoil Mongstad refinery. The objective is to minimize the energy input in the gas fired heater by optimally distributing the cold crude oil in the heat exchanger network. The steady state mass and energy balance of the 20 heat exchangers in the network yields the process model. This model is fitted to the measured values using data reconciliation and unmeasured values like heat exchanger duty and heat transfer coefficients are computed. The fitted model is used to compute the optimal split fractions of crude in the network. This system has been implemented at the refinery and has resulted in a 2% reduction in energy consumption. In operational modes where the unit is constrained on energy input this gives a increased throughput and a significant contribution to the refinery profit.

Keywords

Reconciliation, Optimization, Crude unit, Heat exchanger network

Introduction

This paper describes the development of a real time optimization system including model development, data reconciliation and on-line optimization. The case studied is a heat exchanger network for pre-heat of feed in a crude oil distillation unit at the Statoil Mongstad Refinery. The system is implemented and is now running in closed loop at the refinery. The optimal operation is computed using a steady state model which before each run is fitted to the current operation point. Process measurements contain uncertainties as random errors and possibly gross errors. This may be a result of miscalibration or failure in the measuring device. This uncertainty is reduced when the current operation point is estimated using a larger number of measurements, than the number of unknowns in the process model, to compute a set of reconciled data. Model parameters are estimated simultaneously or computed from the reconciled data. The optimal operation is computed as the maximum of the objective subject to the process model, current process operation and model parameters. The optimal operation is finally implemented as setpoints in the process control system. A large number of methods for data reconciliation have been suggested. These include robust objective functions (Chen et al., 1998), statistical tests, analysis of measurement redundancy and variable observability (Crowe et al., 1983). However, most examples and case studies presented in the literature are based on simulated processes, and most papers consider the data reconciliation decoupled from the optimization. One noteworthy exception is (Chen et al., 1998) who present an application of data reconciliation to a Monsanto sulfuric acid plant, but the paper is somewhat limited on details on the specific approach they have taken. The objective of this paper is therefore to present an actual industrial implementation, where we provide details about the data reconciliation approach, model and optimization.

*Tore.Lid@Statoil.com

Data reconciliation

Data reconciliation is used to determine the current operation point. If measurement had no uncertainty the current operation point could be determined from $n - m$ independent measurements, where n is the number of variables and m the number of equations in the model. Since the measurements are uncertain and there are a surplus of measurements, compared to the number of unknown variables in the model, data reconcilia-

tion is used to reduce this uncertainty. The reconciled values minimizes some function of all measurement errors subject to the model equations. This is written as

$$\begin{aligned} \min_x \sum_{i=1}^{n_m} \psi(\epsilon_i/\sigma_i) \\ s.t. \qquad Ax = 0 \\ g(x) = 0 \end{aligned} \tag{1}$$

All variables are collected in the vector x of dimension $n \times 1$. The measurement errors $\epsilon_i = x_i - y_i$ where y_i is a measurement of the variable contained in x_i. All measurement errors is scaled by its standard deviation σ_i. The process model is separated into a set of linear equations, $Ax = 0$, and nonlinear equations, $g(x) = 0$, since most NLP solvers take linear and nonlinear equations as separate arguments. If the uncertainty in the measurements are normal distributed with zero mean the summed squared measurement error is used as objective function, ψ, in equation 1. However, in the case of nonzero measurement error mean, gross errors, this method gives a biased estimate of the process variables. There are several methods for reducing of the effect of gross errors. In (Crowe et al., 1983) and (Crowe, 1986) collective and individual statistical tests of the measurement errors are used to exclude measurements with gross errors. In (Chen et al., 1998),(Tjoa and Biegler, 1991) and (Johnston and Kramer, 1995) objective functions less sensitive to gross errors are used. In this work the *Combined Gaussian* objective function is selected due to its numerical robustness and promising "small example" results. In all robust objective functions the measurement error is scaled by its standard deviation. Normally this distribution is not known and the standard deviation has to be estimated from measured data or determined by a reasonable guess based on the actual measurement equipment installed and its measurement range. The Combined Gaussian function is based on a weighted sum of two Gaussian distributions, one distribution of the random errors and one of the gross errors. The combined Gaussian probability density function is written as

$$f_i = \frac{1}{\sigma_i\sqrt{2\pi}} \left[(1-p) \exp\left(-\frac{1}{2}\frac{\epsilon_i^2}{\sigma_i^2} \right) + \frac{p}{b} \exp\left(-\frac{1}{2}\frac{\epsilon_i^2}{\sigma_i^2 b^2} \right) \right] \tag{2}$$

with the probability of a gross error in the measurements p and the ratio of the standard deviations of the gross errors to that of the random errors b. The objective function to be minimized is the negative logarithm of the probability density function, $\sum_{i=1}^{n_m} \log(1/f_i)$. The Combined Gaussian objective function is graphed in figure 1 with the least squares function for comparison. Compared to the least squares method the robust functions gives less penalty for measurement errors larger

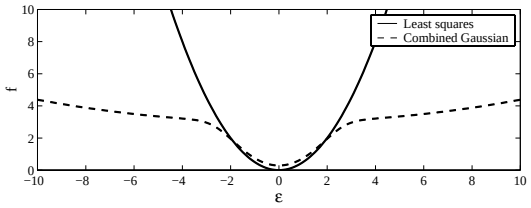

Figure 1: Combined Gaussian objective function. The standard deviation $\sigma = 1$, $p = 0.3$ and $b = 6$.

than 3σ. For the reconciled data this typically gives large measurement errors in few variables and small error in the other variables. At least intuitively this is what one would expect from the process measurements though it is difficult to verify. In equation 1 there is no limitation on the number of measurements and on which variable to measure. Before the reconciled variables are accepted some analysis has to be made to check if the unmeasured variables are observable. The measurements can also be classified as redundant or nonredundant measurements which can be used to evaluate the reconciled variables and decide if data reconciliation can be done if a specific measurement is out of service. Let x_r^* be a solution to the reconciliation problem in equation 1. The nonlinear constraints are linearized at the optimal solution x_r^* such that $g(x) \approx g(x_r^*) + G(x - x_r^*)$, where $G = \partial g(x)/\partial x|_{x=x_r^*}$. The linear and linearized constraints can now be written as

$$\hat{A}x - \hat{b} = 0 \tag{3}$$

where

$$\hat{A} = \begin{bmatrix} A \\ G \end{bmatrix} \qquad \hat{b} = \begin{bmatrix} 0 \\ g(x_r^*) - Gx_r^* \end{bmatrix}$$

The variables in x are separated into measured variables, y, and unmeasured variables, z. The matrix $\hat{A}$ is partitioned into $\hat{A}_1$ and $\hat{A}_2$ where $\hat{A}_1$ holds the columns of $\hat{A}$ corresponding to the measured variables and $\hat{A}_2$ the columns of $\hat{A}$ corresponding to the unmeasured variables. Equation 3 can now be written as

$$\hat{A}_1 y + \hat{A}_2 z = \hat{b} \tag{4}$$

To be able to compute the unmeasured variables, from the the measured variables, the matrix $\hat{A}_2$ must have full column rank. If the number of measurements $n_y < n-m$, where n is the number of variables and m the number of equations, the size of $\hat{A}_2$ is $m \times n_z$ where $n_z > m$ and the matrix $\hat{A}_2$ has rank less than n_z. This implies that equation 4 has no unique solution for z when y is known. A requirement is that the number of measurements $n_y \geq n-m$, which is obvious, and that $\hat{A}_2$ has full column rank. The measurements can also be separated into redundant and nonredundant measurements. If a

variables measurement is redundant it is possible to compute its value if its measurement is removed. This is not the case for a nonredundant measurement and removing this measurement causes $\hat{A}_2$ to be rank deficient. A simple test for redundancy is to check if $P^T\hat{A}_1$ has columns with only zero elements, where P is defined as a matrix that span the null space of $\hat{A}_2^T$. Any zero columns in corresponds to nonredundant measurements. Also note that for a nonredundant measurement i we always have that $y_i - y_{mi} = 0$ and that this measurement does not contribute directly in the calculations of the reconciled values.

Optimization

The typical process optimization problem has a linear objective function like product price times product flow which is to be maximized. For system simplicity the same process model and variable vector are used in both data reconciliation and process optimization. In the optimization problem some of the variable values are already known. These are typically disturbance variables and connects the data reconciliation with the optimization. The variable values are specified in the optimization problem as a set of linear constraints ($Rx = r$) where $r = Rx_r^*$. The matrix R has one nonzero element in each row, equal to one, corresponding to the element in x, which is set equal to its reconciled value. The optimization problem can now be written as

$$\begin{aligned} \min_x \; & -p^T x \\ s.t. \quad & Ax = 0 \\ & g(x) = 0 \\ & Rx = r \\ & x_{\min} \geq x \geq x_{\max} \end{aligned} \tag{5}$$

Inequality constraints in process optimization are typically bounds on singe variables. Inequality constraints on combinations of variables may be added in this formulation by introducing slack variables.

A case study

In the crude unit the crude feed is separated into suitable components for production of propane, butane, gasoline, jet fuel, diesel and fuel oil. The crude is preheated in a heat exchanger network where heat is recovered from the hot products and circulating refluxes. As shown in figure 2 the cold crude (DCR) is separated into seven parallel streams (A-G) and heated by the hot products. The flow in each pass and BSR heat exchanger bypasses provides the degrees of freedom necessary for optimization. The optimization objective is to save energy and to recover as much heat as possible. The heater is the main energy

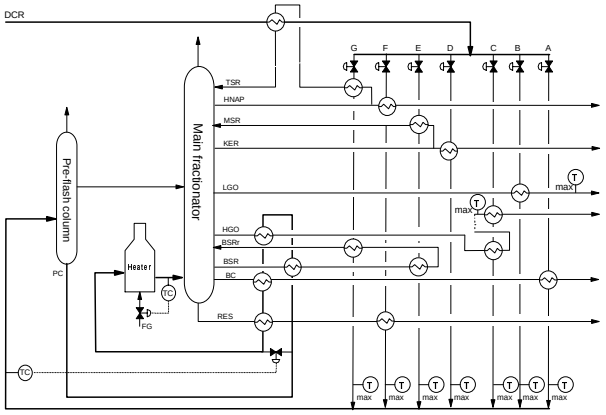

Figure 2: Simplified crude unit overview

input in the process and heater outlet temperature is held constant. The minimum energy is then achieved by maximizing of the heater inlet temperature. Both distillation columns have feed conditions independent on the heat exchanger network operation. The inlet temperatures of both columns are assumed to have perfect temperature control. The feed flow and composition are then independent of operation of the heat exchanger network. With this simplification a model of the distillation columns is not needed and a mass and energy balance of the heat exchanger network is a sufficiently detailed model for optimization. The optimal solution must be within several process operating constraints. The total crude flow or throughput is to be unchanged. At each crude pass outlet there is a maximum temperature constraint to avoid flashing. On main column LGO and HGO products, exiting the heat exchangers, there is a maximum temperature limit as the products are fed to the LGO and HGO driers (the driers are not drawn in figure 2). The preflash column inlet temperature is to be unchanged. Some of the heat exchangers are also included in the bottom circulating reflux (BSR) and the total duty in BSR is to be unchanged.

The process model

The heat exchanger network can be viewed as a set of nodes or unit operations connected by arcs or in this case pipes. A set of balance equations, mass and energy balance, describes the internals of each node. Variables for the arcs or pipes are fluid temperature and mass flow. The nodes in this network are stream mix nodes, stream split nodes and heat exchanger nodes. This selection of variables makes all nodes independent of other variables than those included in the input and output arcs. Heat exchanger nodes also have some internal variables like heat transfer coefficient and duty. This variable selection makes the model structure a simple and surveyable

and it makes it practical possible to compute analytical derivatives of the nonlinear model equations. This reduces the numerical computational load in solving the model. The following describes the simplified balance equations for each type of node.

Mixing of streams

In a node where n streams are mixed into one outlet stream the mass and energy balance equations can be written as

$$F_{\text{out}} - \sum_{i=1}^{n} F_{\text{in}_i} = 0 \tag{6}$$

$$F_{\text{out}} h(T_{\text{out}}) - \sum_{i=1}^{n} F_{\text{in}_i} h(T_{\text{in}_i}) = 0 \tag{7}$$

where $h(T)$ is the specific enthalpy of the fluid. The mass balance results in one linear equation and the energy balance in one nonlinear equation.

Splitting of streams

In a node where one inlet stream are separated into n outlet streams the mass and energy balance equations can be written as

$$F_{\text{in}} - \sum_{i=1}^{n} F_{\text{out}_i} = 0 \tag{8}$$

$$T_{\text{in}} - T_{\text{out}_i} = 0 \qquad \forall\, i = 1...n \tag{9}$$

The mass balance results in one linear equation and the energy balance results in n linear equations.

Heat exchanger

For a heat exchanger node hot and cold side mass and energy balance and heat transfer is written as

$$F_{\text{c}_{\text{in}}} - F_{\text{c}_{\text{out}}} = 0 \tag{10}$$

$$F_{\text{h}_{\text{in}}} - F_{\text{h}_{\text{out}}} = 0 \tag{11}$$

$$Q + F_{\text{c}_{\text{in}}}(h(T_{\text{c}_{\text{in}}}) - h(T_{\text{c}_{\text{out}}})) = 0 \tag{12}$$

$$Q - F_{\text{h}_{\text{in}}}(h(T_{\text{h}_{\text{in}}}) - h(T_{\text{h}_{\text{out}}})) = 0 \tag{13}$$

$$Q - \varepsilon C_{\min}(T_{\text{h}_{\text{in}}} - T_{\text{c}_{\text{in}}}) = 0 \tag{14}$$

where the mass balance results in two linear equations (10 ,11) and the energy balance results in two nonlinear equations (12, 13). The heat transfer is described by equation 14. The heat exchangers in this unit is of multiple tube and multiple shell pass type and the ε-Ntu method (Mills, 1995) is used for calculation of the heat transfer. In equation 14 ε is the efficiency and $C_{\min}$ is the minimum capacity. $C_{\min}$ is calculated as

$$C_{\min} = \min(C_{\text{c}}, C_{\text{h}}) \tag{15}$$

$$C_{\text{c}} = F_{\text{c}_{\text{in}}} \text{C}_{\text{p}_{\text{c}}} \approx F_{\text{c}_{\text{in}}} \frac{h\left(T_{\text{c}_{\text{out}}}\right) - h\left(T_{\text{c}_{\text{in}}}\right)}{T_{\text{c}_{\text{out}}} - T_{\text{c}_{\text{in}}}} \tag{16}$$

$$C_{\text{h}} = F_{\text{h}_{\text{in}}} \text{C}_{\text{p}_{\text{h}}} \approx F_{\text{h}_{\text{in}}} \frac{h\left(T_{\text{h}_{\text{out}}}\right) - h\left(T_{\text{h}_{\text{in}}}\right)}{T_{\text{h}_{\text{out}}} - T_{\text{h}_{\text{in}}}} \tag{17}$$

The efficiency, ε, is a function of the number of transfer units, Ntu, and the capacity ratio, R_C. R_C and Ntu is calculated as

$$R_C = \frac{C_{\min}}{C_{\max}} \qquad Ntu = \frac{UA}{C_{\min}} \tag{18}$$

where $C_{\max} = \max(C_{\text{c}}, C_{\text{h}})$. The efficiency ε equals ε_1 for heat exchangers with single shell pass ($n = 1$) and a even number of tube passes. ε equals ε_2 for heat exchangers with even number of tube passes and n shell passes. ε_1 and ε_2 is calculated as

$$\varepsilon_1 = 2\left\{1 + R_C + \sqrt{1+R_C^2}\frac{1+\exp\left(-\frac{Ntu}{n}\left(\sqrt{1+R_C^2}\right)\right)}{1-\exp\left(-\frac{Ntu}{n}\left(\sqrt{1+R_C^2}\right)\right)}\right\}^{-1} \tag{19}$$

$$\varepsilon_2 = \left[\left(\frac{1-\varepsilon_1 R_C}{1-\varepsilon_1}\right)^n - 1\right]\left[\left(\frac{1-\varepsilon_1 R_C}{1-\varepsilon_1}\right)^n - R_C\right]^{-1} \tag{20}$$

When the equations for $C_{\min}$ and ϵ is substituted into equation 14 each heat exchanger is described by two linear and tree nonlinear equations.

Model summary

There are totally 85 streams and 20 heat exchangers in the heat exchanger network. There are 9 stream mixes and 7 stream splits. The variables are 85 flows and 85 temperatures from the streams, 20 heat exchanger duties, 20 heat transfer coefficients and adds up to totally 210 variables. From the heat exchangers we have 40 linear and 60 nonlinear equations. From the stream mixing nodes we have 9 linear and 9 nonlinear equations and from the split nodes we have 29 linear equations. Coefficients for linear equations are collected in the matrix A where each equation occupy one row. The equation coefficients are placed in the column corresponding to its variable position in x. The nonlinear equation residues are collected in the residual vector $g(x)$. The model is now in the preferred form

$$Ax = 0 \tag{21}$$

$$g(x) = 0 \tag{22}$$

where A is a 78×210 matrix with the linear equation coefficients and $g(x)$ is a 1×69 vector of nonlinear equation residues.

On-line data reconciliation

Data is sampled from the process as one hour averages and reconciled using the Combined Gaussian objective function. Standard deviations for measurements are selected to be 1°C for temperature measurements and 2% of the maximum measuring range for flow measurements. The Combined Gaussian parameters p and b are set to 0.3 and 6. To avoid numerical difficulties in the model equations, like reversed flows, appropriate variable bounds

are added to the data reconciliation problem in equation 1. There are 88 measurements in the process, which is a surplus of 25 compared to the number of unknown in the process model. The described analysis shows that all unmeasured variables are observable and that all measurements are redundant. As a example figure 3 shows

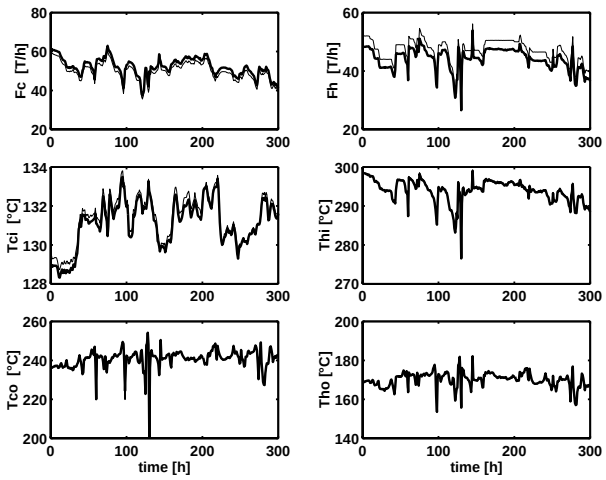

Figure 3: Measured (thin lines) and reconciled values (thick lines) for one of the heat exchangers

measured and reconciled values for 300 successive samples of one hour averages. The imbalance in the data is most likely caused by a gross error in the flow measurement of the hot stream F_{h}. The average error is 3.1 T/h and is fairly constant in all samples.

On-line optimization

In the optimization problem the number of equality constraints are increased to 205 which leaves 5 degrees of freedom. These degrees of freedom corresponds to the flow trough each of the seven passes in the hot train minus two since the total flow and BSR duty is set equal to the reconciled value. As a example measured data is reconciled and optimum operation computed. Compared to current operation the pass flow (A-G) is changed by [0.0,-9.2,-0.1,+9.0,+0.1,+1.0,-0.8]%. In addition bypass flows of heat exchangers in the BSR is changed such that more heat is added in each pass while keeping the total duty constant. This increases the heater pre-heat duty by 2.3MW. Compared to the heater duty of $\approx$ 100MW this gives a 2% reduction of energy requirement. Constraints on pass G outlet temperature and LGO drier inlet temperature is active at optimal operation. Optimal operation is implemented as flow ratio setpoints in the MPC controller. Both the data reconciliation and optimization problems is solved using a software package for constrained optimization problems (NPSOL from Stanford University). This system runs on a DEC-Alpha computer and the average solution time is 3 minutes.

Conclusion

A process model describing the mass and energy balance is developed and used for data reconciliation and optimization. The model is fitted to the measured values and optimal feed split fractions are computed and implemented in the control system once an hour. The reconciled values provides valuable information about the current condition of the measurement equipment and of the condition of the heat exchangers. Comparison of reconciled values and measured values have detected several flow measurements with pour performance and also a temperature measurement that was found to be installed in the wrong pipe. The evolution of heat transfer coefficients during operation is also used to detect fouling and schedule cleaning of the heat exchangers. The model is sufficiently detailed for optimization purposes and the predicted optimal heater inlet temperature is achieved in the process.

Acknowledgement

The authors acknowledge the support of the process control staff at the Statoil Mongstad Refinery and the Statoil Research Centre for all help in implementing this system as a Septic MPC and RTO application.

References

Chen, X., R. W. Pike, T. A. Hertwig, and J. R. Hopper, "Optimal implementation of on-line optimization," *European Symposium on Computer Aided Process Engineering*, pages 435–442 (1998).

Crowe, C. M., "Reconciliation of process flow rates by matrix projection, Part II: Nonlinear case," *AIChE Journal*, **32**(4), 881–888 (1986).

Crowe, C. M., Y. A. Garcia Campos, and A. Hrymak, "Reconciliation of process flow rates by matrix projection, Part I: Linear case," *AIChE Journal*, **29**(6), 881–888 (1983).

Johnston, L. P. M. and M. A. Kramer, "Maximum likelihood data rectification: Steady state systems," *AIChE Journal*, **41**(11), 2415–2426 (1995).

Mills, A. F., *HEAT AND MASS TRANSFER*, The Richard D. Irwin series in heat transfer. RICARD D. IRWIN, INC, first edition (1995).

Tjoa, I. B. and L. T. Biegler, "Simultaneous strategies for data reconciliation and gross error detection of nonlinear systems," *Computers & Chemical Engineering*, **15**(10), 679–690 (1991).