Optimization algorithms intended for self-tuning feedwater heater model

To cite this article: P Czop et al 2013 IOP Conf. Ser.: Mater. Sci. Eng. 46 012016

View the article online for updates and enhancements.
Optimization algorithms intended for self-tuning feedwater heater model

P Czop, T Barszcz and J Bednarz
AGH University of Science and Technology, Department of Robotics and Mechatronics
E-mail: pczop@agh.edu.pl.

Abstract. This work presents a self-tuning feedwater heater model. This work continues the work on first-principle gray-box methodology applied to diagnostics and condition assessment of power plant components. The objective of this work is to review and benchmark the optimization algorithms regarding the time required to achieve the best model fit to operational power plant data. The paper recommends the most effective algorithm to be used in the model adjustment process.

Introduction

Great advances in control and monitoring systems applied in thermal power plants have created a possibility to develop new efficient methodologies for diagnostics and condition assessment of power unit components. The state-of-the-art gray-box model diagnostic approach, proposed and developed by Barszcz and Czop in [1-2], requires efficient methods to adjust the first-principle models to operational data. One of the most demanding tasks is tuning of first-principle models which involve three categories of parameters: geometrical, physical and phenomenological (e.g. heat exchange coefficients). Geometrical and physical parameters are deduced from construction or operational documentation, while phenomenological ones are adjusted using optimisation routines. A feasibility study of optimisation routines applied in a tuning process of the first-principle is presented in this paper based on a feedwater heater installed in a 225MW coal-fired power block. Heaters installations are typically affected by fouling and corrosion phenomena which may have effect on heat transfer and fluid transportation process [1-2]. Feedwater heaters are typically designed as three-zone heat exchangers with a condensing section, desuperheater and integrated subcooler (Figure 1).

1 Corresponding author, Al. Mickiewicza 30, 30-059 Cracow, Poland.
The drainage system of the feedwater heater consists of a drain removal path from each heater. The normal drain flow path is cascaded to the next lower stage heater, and the alternate path is diverted to the condenser. When the turbine is loaded at a given rate, steam is allowed to enter the bank of heaters through extraction outlets [1-2].

2. Review the methods to update the model parameters

The process of heater model calibration consists of (i) selection of representative data sets, including a sufficiently broad operating range that the model proved could work correctly using a data fit measure and (ii) adjustment of model parameters to fit a model response to data. The model is represented as a set of non-linear state-space equations formulated in the continuous-time domain. The objective of the estimation is to minimize the error function between the measured signals and model responses by means of an iterative numerical technique. The function describing the error has to be a positive and decreasing function. The procedure for updating model parameters consists of two in-a-loop phases: (i) simulation of a model by solving differential equations numerically in Simulink [4], and (ii) numerical minimization in the parameter space with respect to an error-related criterion function using the Matlab Optimization Toolbox [4]. After each simulation of the model for fixed-length input signals, the simulated output data are sampled and the criterion function is re-evaluated to calculate a new set of model parameters. Detail discussion on model tuning algorithms and validation techniques is given in [3]. A general form of the first-principle model which represents power-block thermodynamic behaviour in the continuous domain is as follows:

\[
\frac{d}{dt} x(t) = f(t, x(t), u(t), w(t); \theta) \\
y(t) = h(t, x(t), u(t), v(t); \theta) \\
x(0) = x_0
\]  \hspace{1cm} (1)
where vector \( f(.) \) is a nonlinear, time-varying function of the state vector \( x(t) \) and the control vector \( u(t) \), while vector \( h(.) \) is a nonlinear measurement function; \( w(t) \) and \( v(t) \) are sequences of independent random variables and \( \theta \) denotes a vector of unknown parameters. In nonlinear systems, the state vectors and the measurements vectors are not Gaussian distributed. The predictor resulting from the model (1) takes the form

\[
\hat{y}(t | \theta) = g(t, Z^{-1}; \theta)
\]

while the prediction error equation has the form

\[
\epsilon(t, \theta) = y(t) - g(t, Z^{-1}; \theta)
\]

The sum of squared errors is used as an error criterion. This problem is known in numerical analysis as “the nonlinear least-square problem” [5].

Three methods of minimizing the error function are feasible as presented in Table 1. These are (i) direct search, (ii) first-order and (iii) second-order methods.

| Method                        | Algorithm                                                                 |
|-------------------------------|---------------------------------------------------------------------------|
| Direct search methods         | Simplex                                                                   |
|                               | Pattern search                                                           |
| First-order (gradient)        | Gradient (steepest-descent)                                              |
| Second-order (gradient and Hessian) | Gauss-Newton                  |
|                               | Quasi-Newton                                                              |
|                               | Trust-regions                                                            |
|                               | Damped Gauss-Newton (Adaptive Gauss-Newton)                              |
|                               | Levenberg-Marquardt                                                     |

Direct search methods use only the value of the function to find the minimum. The first-order method uses the information provided by the first derivative (gradient) of the error function, while the second-order method uses both, information regarding the first and the second order derivatives (gradient and Hessian form) of the error function. First- and second-order methods update the estimates of the error function iteratively according to the general scheme

\[
\hat{\theta}^{(k+1)} = \hat{\theta}^{(k)} + \alpha K \cdot f^{(k)}
\]

where \( \hat{\theta}^{(k)} \) denotes the k-th iteration point in the search, \( f^{(k)} \) is a search direction based on the information about \( V(\theta) \) acquired at previous iterations, and the sequence of positive scalars \( \alpha \) determine velocity in which the value of \( V(\theta) \) is decreased. In general, the function \( f^{(k)} \) has the form

\[
f^{(i)} = -\left[V'(\hat{\theta}^{(i)})\right]^{-1}V'(\hat{\theta}^{(i)})^T
\]

The first and second derivatives are obtained as follows.
\[ V_N'(\theta) = -\frac{1}{N} \sum_{i=1}^{N} \epsilon(t, \theta) \psi(t, \theta) \]  
\[ V_N''(\theta) = \frac{1}{N} \sum_{i=1}^{N} \psi(t, \theta) \psi'(t, \theta) - \frac{1}{N} \sum_{i=1}^{N} \epsilon'(t, \theta) \epsilon(t, \theta) \]  
where \( \psi'(t, \theta) \) denotes the d-by-d Hessian matrix of \( \epsilon(t, \theta) \), \( d = \text{dim}(\theta) \) [5]. A general family of iterative search routines is described by the scheme

\[ \hat{\theta}_N^{(k+1)} = \hat{\theta}_N^{(k)} + \mu_N^{(k+1)} \left[R_N^{(k)} \right]^{-1} V_N'(\hat{\theta}_N^{(k)}, Z_N) \]  

where \( R_N^{(k)} \) is a d-by-d matrix, which modifies the search direction (changes the search direction according to user preferences), while the step size \( \mu_N^{(k+1)} \), is chosen so that it is satisfied in the inequality

\[ V_N(\hat{\theta}_N^{(k+1)}, Z_N) < V_N(\hat{\theta}_N^{(k)}, Z_N) \]  

It is often necessary to use variable step length to improve the convergence of the algorithm; a sequence of scalars \( \mu_N^{(k+1)} \) is used to control the step length. If the term \( \mu_N^{(k+1)} \neq 1 \), equation (8) describes the so-called damped Gauss-Newton algorithm.

3. The feedwater heater model

The family of feedwater heater models was extensively discussed in [1]. It has four phenomenological parameters, which are the heat exchange coefficients [1]. Two parameters describe the heat exchange process between steam-condensate volumes and feedwater volumes, while the remaining two parameters describe the heat exchange process between steam-condensate volumes and the metal housing [1]. The two latter parameters have to be neglected due to lack of measurements of the temperature of the metal housing. This simplification neglects a bypass energy flow through the metal housing. This simplification is allowed since the heat exchange process through the heater wall is significantly smaller compared to heat exchange through metal tubes with feedwater. This allows the model simplifications to be verified experimentally. The control system could not be directly reconstructed in the simulation, due to its complexity and limited relevance to the functionality required in the model (e.g. trip logic). Hence, the module maintaining a constant level of the condensate inside the heater was simplified using a PID controller model. The parameters of the first-principle model are updated based on operational data. The total length of each sequence corresponds to 12 hours and includes daily fluctuation in the produced electrical energy form 225MW to 160MW.

Performance of the procedure of adjusting model parameters is evaluated by analysing the value of Pearson’s product-moment correlation coefficient between the measured and simulated values of the N-sample-long output signals (Figure 2 bottom graph). For each output signal, the correlation coefficient is computed separately; in the particular case of the model discussed in this paper, the output signals were the temperature of the condensate and the temperature of the feedwater. The convergence trajectory plots were also investigated to ensure a stable trend towards constant values of the parameters, which correspond to convergence towards the minimum of the criterion function, within less than 10 iterations (Figure 2 top and mid graph).
Figure 2. Exemplary fitting accuracy of models to operational data as a function of time (legend: FW – feedwater, CO – condensate).

4. The benchmark of the model tuning methods

The validation tests were conducted using a few minimization algorithms representing three categories of methods as summarized in Table 1, i.e. direct search, first-order, and second-order. The sum of squared errors was used as the error criterion to evaluate the fit of the model to operational data. A direct search approach [5], which does not use numerical or analytic gradients, is represented by two algorithms: (i) pattern search involving different sub-routines to perform a search and the simplex search method. Both methods are recommended for optimization tasks that contain discontinuities. The pattern search method has the following choices: no search step, Latin hypercube, a pattern search
(GPS or MADS methods), or genetic algorithm [6]. The simplex search method uses the Nelder-Mead algorithm [7]. The first-order and the second-order gradient methods are represented by the gradient descent and nonlinear least squares optimization algorithms from the Optimization Matlab Toolbox [4]. Second-order methods are more effective compared to the first-order ones when the second-order information is obtainable in an analytical way. The calculation of second-order information, using numerical differentiation, is computationally expensive. The second-order method is available in two versions: (i) a subspace trust-region [8-9], and (ii) Levenberg-Marquardt method. A subspace trust-region method approximates only a certain region (the so-called trust region) of the objective function with a model function (often a quadratic) as opposed to the entire function as with the Newton–Raphson optimization algorithm. When an adequate model of the objective function is found within the trust region, then the region is expanded. If the approximation is poor then the region is contracted. Table 2 shows the evaluation results of different algorithms, for the stopping criterion chosen as the exceeded number of iterations or approached stable value for the parameters, i.e. the difference between the current and the previous value is zero.

| Method              | Algorithm      | Matlab function | Condensate temperature fit [-] | Feedwater temperature fit [-] | T [min] |
|---------------------|----------------|-----------------|-------------------------------|-------------------------------|---------|
| direct search       | Nelder-Mead    | fminsearch      | 0.865                         | 0.603                         | 35.1    |
| first-order         | Steepest-descent| fmincon         | 0.874                         | 0.514                         | 235.8   |
| second-order        | Levenberg-Marquardt | lsqnonlin     | 0.897                         | 0.615                         | 48.2    |
| second-order        | Trust-region   | lsqnonlin       | 0.894                         | 0.632                         | 27.9    |

The results that provide the best trade-off between data fit to operational data and calculation time were achieved for direct methods, while the worst results were for the first-order method. Nevertheless, if iteration limits for algorithms are replaced by the parameter tolerance value bounds, the convergence is faster for second-order methods. The pattern search algorithms provide step-like convergence showing a few significant variations in the values of parameters before they approach the stable value. This variation requires many additional iterations compared to gradient methods, thus the stable value is not approached as fast as in the gradient methods. Moreover, the pattern search function in Matlab ignores the initial values of model parameters (Latin hypercube) by replacing the guess values with pre-calculated values which differ significantly. In turn, the iteration process took much more time than was required for second-order algorithms.

5. Conclusions

The objective of this work is to review and benchmark the optimization algorithms regarding the time required to achieve the best model fit to operational power plant data. The model was a feedwater heater first-principle model developed in [1]. Table 1 compares different optimisation methods used for adjusting the model to operational data. The second-order trust-region method was shown to be the most optimal for the problem of updating parameters of the heater model developed in [1-2].

Acknowledgements

The author gratefully acknowledges the financial support of the research project N504 493239, funded by the Polish Ministry of Science (MNiI).

Nomenclature
x(t) - state vector used in a general representation of a first-principle model
u(t) - control vector used in a general representation of a first-principle model
y(t) - output vector used in a general representation of a first-principle model
w(t), v(t) - the sequences of independent random variables
θ - the unknown parameter vector used in a general representation of a first-principle model

Abbreviations

FP - First-Principle
DD - Data-Driven
GPS - Generalized Pattern Search
ODE - Ordinary Differential Equations
MADS - Mesh Adaptive Direct Search

References

[1] Barszcz T Czop P 2011 A Feedwater Heater Model Intended for Model-Based Diagnostics of Power Plant Installations, *Applied Thermal Engineering* vol 31 (8-9) pp 1357-1367
[2] Barszcz T and Czop P 2012 Presentation of a Virtual Power Plant environment and its application with combined first-principle and data-driven models intended for the diagnostics of a power plant - part I *Simulation Transactions of The Society for Modeling and Simulation International, Simulation* vol 88(2) pp139-166
[3] Barszcz T Czop P 2011 Estimation of feedwater heater parameters based on a grey-box approach. *International Journal of Applied Mathematics and Computer Science* vol 21 (4), pp 703-715
[4] MATHWORKS 2007 Inc., Matlab/Simulink (Natick US)
[5] Ljung L 1999 *System identification - Theory for the User* (Prentice-Hall, USA)
[6] Torczon VJ 1997 On the Convergence of Pattern Search Algorithms *SIAM Journal on Optimization* vol 7(1), pp 1–25
[7] Lagarias JC Reeds JA Wright MH and Wright PE 1998 Convergence Properties of the Nelder-Mead Simplex Method in Low Dimensions, *SIAM Journal of Optimization* vol 9(1) pp. 112-147
[8] Coleman TF and Li Y 1996 An Interior, Trust Region Approach for Nonlinear Minimization Subject to Bounds *SIAM Journal on Optimization* vol 6(2) pp 418–445
[9] Coleman TF and Li Y 1994 On the Convergence of Reflective Newton Methods for Large-Scale Nonlinear Minimization Subject to Bounds *Mathematical Programming* vol 67(2) pp 189-224