Data Reconciliation for Blast Furnace

Sam Mathew
Consultant
Gyan Data Private Limited
Chennai-600113 India
sam@gyandata.com

About the Client

Tata Steel Limited (TSL) (formerly TISCO) is a multinational steel making company with its largest manufacturing plant located at Jamshedpur, Jharkhand. Its in-house R&D division at Jamshedpur deals with design, control and optimization of the steel making process and is renowned for its engineering capabilities towards process optimization.

Motivation

Blast-furnaces require an accurate estimation of the hot-metal temperature and its production which can only be inferred from other measurements. Hot-metal temperature estimation is especially important to avoid any "chilling" of the furnace which can freeze several upstream and downstream processes in steel-making. The design principles for optimization require that all measurements be consistent, satisfying heat and mass balances. Real time measurements often have random or gross errors (sensor biases) that impact the accuracy of data. Data reconciliation and gross error detection are twin statistical techniques that have been developed to systematically obtain consistent estimates.

Problem

The R&D team at TSL faced a problem with addressing the discrepancy of the consumption of raw-materials (iron ore and coal) and the corresponding steel production. The chemical analyses of ore, fuel, steel and slag are performed at different frequencies. Compounding matters further, the process measurements of flow-rate, flue-gas composition and temperatures are obtained at yet another different set of frequencies.

The raw-material in the form of sinter, pellet and ore is dumped in batches that spends approximately 6-8 hours within the furnace. The challenge therefore was to find consistent estimates of these measurements obtained at different frequencies and get the corresponding estimates of hot-metal production and temperature. Gyan Data offered to apply data reconciliation with gross-error detection after harmonizing the data across consistent time-stamps. TSL provided the necessary mass and energy balance equations that were required to be satisfied at any given period.

The difficulty was then to express these highly non-linear equations, involving flow, composition and enthalpy balances as constraint equations in an optimizer. The resulting problem contained around 100 constraint equations involving more than 250 variables, with over 40% of them being measured.

Solution

The blast furnace has time-series data obtained from the sensors and flow measurements at the furnace, or through chemical analyses done on the streams going in and out of it. Each of these are obtained at different time-stamps and therefore require to be harominzed and/or averaged before any attempts are to be made at reconciling them.

The overall task for the project was therefore divided into two parts :

1. Data preparation : Harmonize the data across same time-stamps, followed by obtaining periodic averages (e.g. daily average or shift average)
2. Data reconciliation : Feeding the data at each time-stamp to the reconciliation module in order to obtain consistent estimates of the input as well as unmeasured quantities (soft-sensors)

As per the requirment, the task was performed using two high-level programming languages as interfaces : Python and MATLab. The time-harmonization was done using available modules in Python while in MATLab its internal time-series functions were used. This was followed by obtaining averaged values across the chosen periods. (e.g. daily averaged) . The burden-side inputs on different kinds of iron-bearing material (pellet, sinter and ore) are dumped into the furnace in batches. This value needed to be summed up for specific periods and averaged, in order to obtain a per-hour flow rate.

The final model of variables, balance equations as constraints and the cost function to be minimized were written in terms of a Mathematical Programming Language (MPL). The MPL chosen in this case is the open-source package, CasADi [1, 3], which offered flexibility to express the equations in terms of symbolic variables and through expression graphs, efficiently calculating various constraints, even evaluating the necessary gradients and hessians of the objective and the constraints. The evaluations were made computationally efficient for the order of this problem using graph-coloring algorithms [2]. The interior-point method and gradient based open-source solver IPOPT [4] was chosen as the optimizer.

Additionally, the reconciled estimates were further processed to evaluate any gross-errors (biases) in order to detect any faulty measurements. If present, such sensors' measurements were dropped and a fresh data reconiciliation was performed to ensure that they do not adversely affect the estimates of other measurements.

References

[1] Andersson, J. *A General-Purpose Software Framework for Dynamic Optimization*. PhD thesis, Arenberg Doctoral School, KU Leuven, Department of Electrical Engineering (ESAT/SCD) and Optimization in Engineering Center, Kasteelpark Arenberg 10, 3001-Heverlee, Belgium, October 2013.

[2] Andersson, J. A. E., Gillis, J., and Diehl, M. CasADi 3.3 User Guide. http://casadi.sourceforge.net/users_guide/casadi-users_guide.pdf, Feb. 2018.

[3] Andersson, J. A. E., Gillis, J., Horn, G., Rawlings, J. B., and Diehl, M. CasADi – A software framework for nonlinear optimization and optimal control. *Mathematical Programming Computation* (In Press, 2018).

[4] Wächter, A., and Biegler, L. T. On the implementation of an interior-point filter line-search algorithm for large-scale nonlinear programming. *Mathematical programming 106*, 1 (2006), 25–57.