Increase of coal burning efficiency via automatic mathematical modeling

Patrick Bangert

algorithmica technologies GmbH[1]

Germany

Abstract

The entire process of a coal power plant from coal delivery to electricity generation can be modeled using machine learning methods that generate a single set of equations to govern the entire plant. The basis for the model is a database of historical data from the data historian. No engineering or human input is required at all. This model is then used to compute the optimal operational point at any moment in time. The criterion for optimality is the overall equipment efficiency over the entire plant. It is found that this optimization can achieve more than 1% absolute improvement over current operations in a real coal power plant over the long term. This method requires no changes to be made to the plant and requires nominal human effort to implement. The achievements are based on operating changes that are now computed and not arrived at using human experience. Having a uniform operational strategy 24 hours a day combined with the fact that this model can incorporate the full complexity of the plant is the original element that lies at the basis of the 1% improvement.

1. Statement of the Problem

A coal power plant essentially works by creating steam from water by heating it via a coal furnace. This steam is passed through a turbine, which turns a generator that makes the electricity. See figure 1 for a diagram.

---

[1] Außer der Schleifmühle 67; 28203 Bremen; p.bangert@algorithmica-technologies.com

The plant has an efficiency that is approximately 40% depending on the design. However, the efficiency is not a constant but changes over time depending on how the plant is operated. While many smaller processes are automated using various technologies, the large scale processes are often controlled by human operators. Depending on the knowledge, experience and level of difficulty of any particular plant state, the decisions of the operators get the plant closer to the maximum possible efficiency.

With operators working in shifts, no one operator controls the plant over the long-term but usually in eight hour shifts. It can be observed that the efficiency oscillates in a rough eight hour pattern showing that human decision making has a significant influence on the efficiency. Not only are some operators better than others, it is not practical to extract and structure the experience and knowledge contained in the human brain of the best operator in such a fashion as to teach it to the others.

Furthermore, the plant outputs several thousand measurements at high cadence. An operator cannot possibly keep track of even the most important of these at all times. The degree of complexity is too great for the human mind to handle and the consequence is that suboptimal decisions are taken.

In this paper, a novel method is suggested to achieve the best possible, i.e. optimal, efficiency at any moment in time. This has achieved an efficiency increase of 1.1% absolute in a real coal power plant. Moreover, this efficiency increase is available uniformly over time effectively increasing the base output capability of the plant.
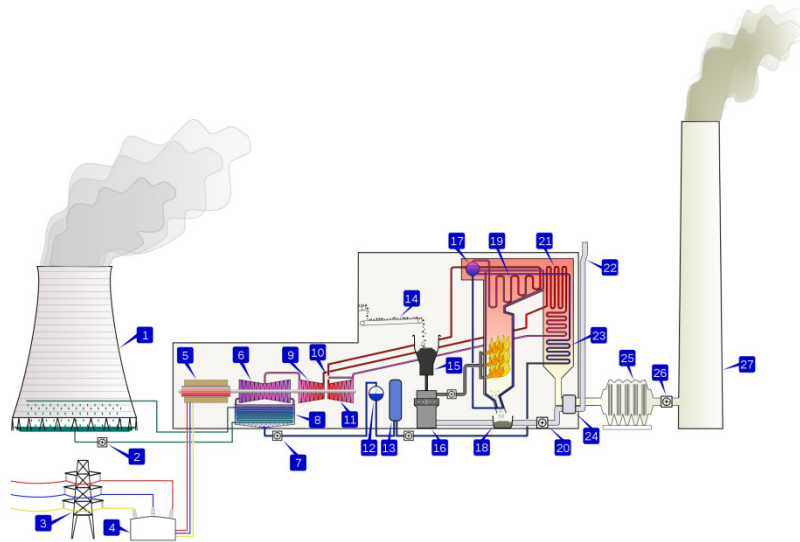
Fig. 1: The main constituents of a coal power plant.

2. Methodology

Sensor equipment is installed in all important parts of the plant and thus alerting the operator via the control system about the current state of the plant. The numerical values of all sensors can be arranged into a vector. Let us assume that we have a total of N measurements on and around the plant that we wish to look at. We may represent the state of the plant at time t by an N-dimensional vector, $\mathbf{x}^{(t)}$. Via the data historian, we may obtain a set of such vectors for past times. If we order this set with respect to time, then this set is called a time-series, $\mathbf{H} = (\mathbf{x}^{(-h)}, \mathbf{x}^{(1-h)}, \mathbf{x}^{(2-h)}, \dots , \mathbf{x}^{(0)})$ where time $t = 0$ is the current moment and time $t = -h$ is the most distant moment in the past that we wish to look at. Thus the time-series $\mathbf{H}$ is effectively a matrix with $h+1$ columns and N rows.

Observe that this matrix contains all the decisions of the operators and all the reactions of plant to these decisions. The knowledge and experience of the operators is thus plainly visible in the data. If the history is long and detailed enough, this information is effectively all one needs to know about this plant.

We recall the topic of control theory. Here we are faced with a black box that has input signals and output signals. The process that connects input to output is totally unknown and represented by the black box. Control theory now aims to discover the relationship between input and output by performing experiments. If we send such a signal in, then we observe such a signal as output. Given enough such data and some analysis, control theory provides

tools for creating a set of equations that govern the behavior of the black box. The resulting set of equations is called a mathematical model.

Note that the model does not allow us to 'understand' the process inside the black box. But it does allow us to compute the output of the black box given a sample input. Using the results of optimization theory, we can reverse this process and compute the input needed to achieve a given desired output.

Control theory is meant to be applied manually. For a process as complex as that of a power plant, this is impractical due to the amount of work that would be required. It is suggested to use machine learning [1] to develop the set of equations automatically. There are various techniques available to achieve this such as neural networks [2]. We opt for the technique of recurrent neural networks [3]. Here we must differentiate classificatory neural networks [2] from recurrent neural networks [3]. The first can tell the difference between a finite number of types of objects while the second can represent the evolution over time.

The advantages of using machine learning are that the model is produced within a very short time (usually days), that it is adaptive (i.e. it learns continuously as it experiences more data), that it can change to match new situations (the new data is learnt) and that the entire problem can be modeled (and not a simplified version as in the manual approach). Thus, this method is economical.

In the state vector that describes the plant, there are elements of three different types. First, there are measurements that can be directly controlled by the operator. An example is the amount of coal per hour being put into a particular mill. We call these **controllable**, $\mathbf{x}_c^{(t)}$. Second, there are measurements that cannot be controlled at all by the operators and thus represent a state of the world. An example is the outside air temperature. We call these **uncontrollable**, $\mathbf{x}_u^{(t)}$. Third, there are measurements that are indirectly controlled via the controllable measurements. An example is a vibration in the turbine. We call these **semi-controllable**, $\mathbf{x}_s^{(t)}$.

Uncontrollable measurements provide boundary conditions for the problem and so we really have a set of models depending on the boundary conditions. This poses no problem for machine learning and is simply included in the model of the black box that is the plant. The

only requirement is that it must be clearly defined which measurements belong into which of the three possible groups. Once this is known, the learning may begin.

What we obtain is a function $f(\mathbf{x}_c^{(t)} ; \mathbf{x}_u^{(t)}) = \mathbf{x}_s^{(t)}$. In words, this means that we have a function with the controllable measurements as variables, the uncontrollable measurements as given parameters and the semi-controllable measurements are functional outputs. The plant efficiency is, of course, among the semi-controllable outputs of the function $f(\ldots)$.

With this model and given a particular boundary condition $\mathbf{x}_u^{(t)}$, we may compute the reaction of the plant $\mathbf{x}_s^{(t)}$ to any particular operator decision $\mathbf{x}_c^{(t)}$. This is effectively a plant simulation. Such a system may be used for training and practice of the operators.

More interestingly, we ask whether the function may be inverted, i.e. whether the function $f^{-1}(\mathbf{x}_s^{(t)} ; \mathbf{x}_u^{(t)}) = \mathbf{x}_c^{(t)}$ can be obtained. Generally, it is not possible to invert functions directly. However, we do not require a closed form solution of this problem but only a numerical solution. This may be achieved using the theory of numerical methods [4].

In particular, we are not necessarily interested in general inversion but rather in a very special form of inversion, namely optimization. Given particular boundary conditions, we would wish to know what input variables lead to the optimal state of the plant. The optimum state is defined by some merit function $g(\mathbf{x}_s^{(t)} ; \mathbf{x}_u^{(t)})$. The simplest such merit function is the plant efficiency but we may also take into account market prices and other business features to define what we believe to be the optimum.

Thus we ask, what is $\mathbf{x}_c^{(t)}$ such that $g(\mathbf{x}_s^{(t)} ; \mathbf{x}_u^{(t)})$ achieves a global maximum where the relationship between the variable vector and the merit function is contained in the inverted model $f^{-1}(\mathbf{x}_s^{(t)} ; \mathbf{x}_u^{(t)}) = \mathbf{x}_c^{(t)}$. This is a classic optimization problem. As the functions are only known numerically and they are highly non-linear and time-dependent, this is a complicated optimization problem requiring state-of-the-art treatment but such problems can be solved.

3. Theoretical limitations

Of course, whatever methods we choose, they cannot have arbitrary accuracy or stability. Thus, every $\mathbf{x}^{(t)}$ has an inherent measurement induced uncertainty $\Delta\mathbf{x}^{(t)}$ attached to it. This means that the true value of the state vector is somewhere in the range $[\mathbf{x}^{(t)} - \Delta\mathbf{x}^{(t)}, \mathbf{x}^{(t)} + \Delta\mathbf{x}^{(t)}]$.

Please note that no measurements made in the real world are ever completely precise. There are random and structured errors associated with the measurement process, also physical sensors drift with age and environmental effects. All of these must be taken into account to determine a reasonable measurement uncertainty $\Delta\mathbf{x}^{(t)}$.

A further limitation is the length of the history. The history must contain a record of the variations that are to be expected in the future so that these variations, correlations and other structures may be included in the model. It is thus desirable that the history be as large as possible and also the time unit (governing the frequency of measurements) be as small as possible. Together these two define a history that contains the maximum available knowledge about the system.

Our efforts are thus limited by three fundamental factors: (1) The number and identity of the measurements made, (2) the length, frequency and variability of recorded history and (3) the inherent accuracy of a measurement itself. Together these three factors will determine whether a reliable and stable model can be found.

4. Application

Initially, the machine learning algorithm was provided with no data. Then the points measured were presented to the algorithm one by one, starting with the first measured point $\mathbf{x}^{(-h)}$. Slowly, the model learned more and more about the system and the quality of its representation improved. Once even the last measured point $\mathbf{x}^{(0)}$ was presented to the algorithm, it was found that the model correctly represents the system.

In a particular power plant, the history covered nine months of data extracted once per minute for nearly 2000 measurements. After modeling, the accuracy of the function deviated from the real measured output by less than 0.1%. This indicates that the machine learning method is actually capable of finding a good model and also that the recurrent neural network is a good way of representing the model.

Controllable variables include settings on the furnace such as the coal loading into the furnace and the provision of air for the furnace. Also some operational settings (set-points) of the turbine and the generator were available. Boundary conditions were provided by coal quality,

the power output of the plant and various weather effects such as the air temperature and humidity and also the cooling water temperature.

The model was then inverted for optimization of plant efficiency. The computation was done for the entire history available and it was found that the optimal point deviated from the actually achieved points by 1.1% efficiency in absolute terms. Moving from, say 40%, to 41.1% efficiency is a significant gain for any power plant.

In the analysis nearly 1000 different operational conditions (in the nine month history) were identified that the operators would have to react to. This is not practical. The model is now capable of determining the current state of the plant, computing the optimal reaction to these conditions and communicating this optimal reaction to the operators. The operators then implement this suggestion and the plant efficiency is monitored. It is found that 1.1% efficiency increase can be achieved uniformly over the long term.

The model can provide this help continuously. As the plant changes, these changes are reflected in the data and the model learns this information continuously. Thus, the model is always current and can always deliver the optimal state.

In daily operations, this means that the operators are given advice whenever the model computes that the optimal point is different from the current point. The operators then have the responsibility to implement the decision or to veto it.

For those parts of the power plant that are already automated, the model is valuable also. Automation generally functions by humans programming a certain response curve into the controller. This curve is obtained by experience and is generally not optimal. The model can provide an optimal response curve. Based on this, the programming of the automation can be changed and the efficiency increases. The model is thus advantageous for both manual and automated parts of the power plant.

Effectively the model represents a virtual power plant that acts identically to the real one. The virtual one can thus act as a proxy on that we can dry run all sorts of strategies and then port these to the real power plant only if they are good. That is the basic principle of the approach. The novelty here is that we have demonstrated on a real power plant, that it is possible to

generate a representative and correct model based on machine learning of historical process data. This model is more accurate, all encompassing, more detailed, more robust and more applicable to the real power plant than any human engineered model possibly could be.

[1]   Bishop, C.M.: Pattern Recognition and Machine Learning. Heidelberg: Springer 2006

[2]   Rosenblatt, F.: The perceptron: a probabilistic model for information storage and organization in the brain. Psychological Reviews 65, 1958, 386-408

[3]   Mandic, D., Chambers, J.: Recurrent Neural Networks for Prediction: Learning Algorithms, Architectures and Stability. Hoboken: Wiley 2001

[4]   Press, W.H., Teukolsky, S.A., Vetterling, W.T., Flannery, B.P.: Numerical Recipes. Cambridge: Cambridge University Press 2010