Increase of Overall Combined-Heat-and-Power (CHP) Efficiency via Mathematical Modeling

P. Bangert[1)], J.-A. Czernitzky[2)]

[1)] algorithmica technologies GmbH, Bremen

[2)] Vattenfall Europe Wärme AG, Berlin

Abstract

The entire process of a combined-heat-and-power (CHP) coal-fired power plant from coal delivery to electricity and heat generation can be modeled using machine learning methods that generate a single set of equations that describe the entire plant. The basis for the model is the historical data from the data historian. No engineering or human input is required. This model is used to compute the *optimal operational point* at any moment in time, taking into account the current status of the plant as well as relevant external and internal factors. The criterion for optimality is the overall equipment efficiency over the entire plant. 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 computed. No human judgment is involved, thus the model excludes a potential bias or human error. 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 novel element that lies at the heart of the improvement possibility in the range of one percent. This has been demonstrated in the Reuter-West CHP plant of Vattenfall in Berlin, Germany.

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.

The plant has an efficiency that depends 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. The power plant Reuter-West is largely automated also in these parts. Thus, the maximum possible efficiency of the plant depends largely on the decisions of the operators, defined by the knowledge and experience of the operator as well as the level of difficulty of any particular plant state. However, the employment of continuous and uniform knowledge and experience for the plant operation is not realistically possible as no one operator controls the plant over the long-term but usually only over an eight-hour shift. Observation results show oscillations of parameters in a rough eight-hour pattern which supports the argument that a fluctuation in the knowledge and experience of human operators may lead to a fluctuation in the decision making and thus a varying influence on the operation of the plant. While some operators may be better than others, it is often not fully practical and/or possible to extract and structure the experience and knowledge of the best operators in such a fashion as to teach it to the others.

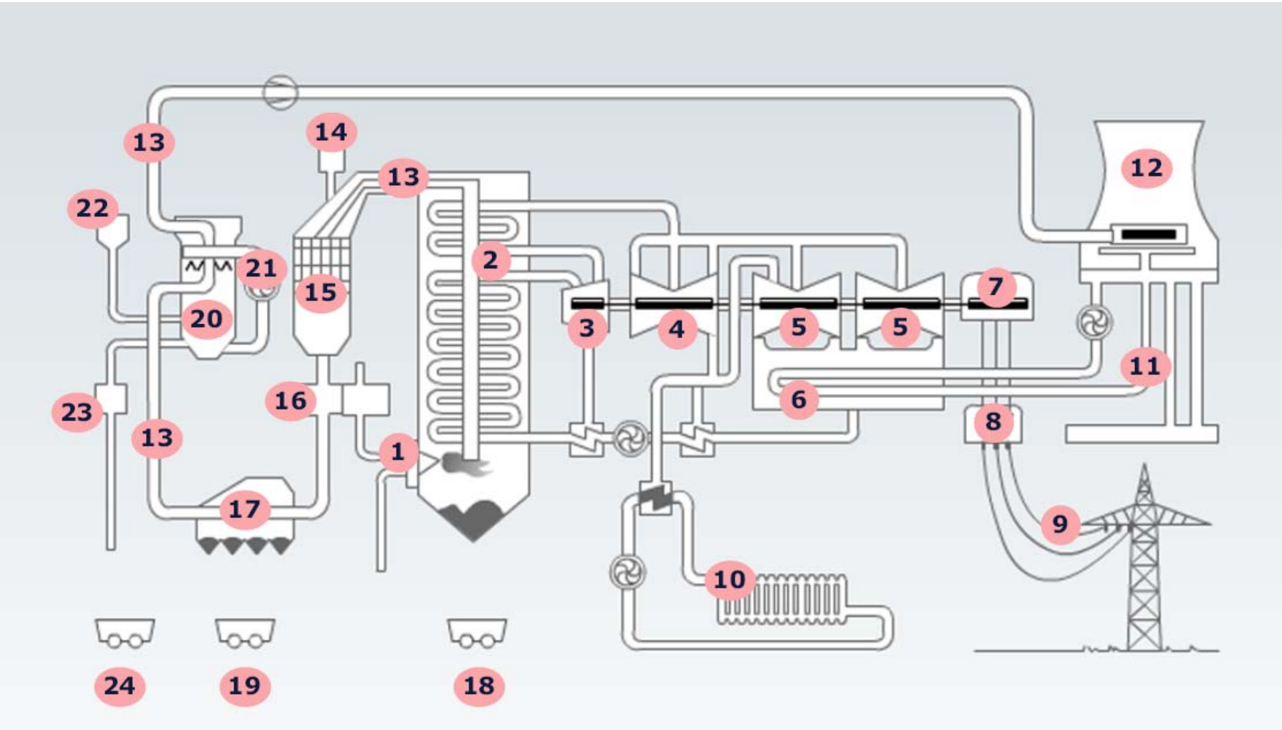

Figure 1. The diagram describes the CHP process in overview including the major steps: (1) Entry-point of the air, (2) boiler with water and steam, (3) high-pressure turbine, (4) mid-pressure turbine, (5) low-pressure turbine, (6) condenser, (7) generator, (8) transformer, (9) feed into power grid, (10) district heating, (11) cooling water source, (12) cooling tower, (13) flue gas, (14) Ammonia addition (15) denitrification of flue gas, (16) air pre-heater, (17) dust filter, (18) ash end-product, (19) filtered ash end-product, (20) desulfurization of flue gas, (21) wash cycle, (22) chalk addition, (23) cement/gypsum removal, (24) cement/gypsum end-product.

Furthermore, the plant outputs several thousand measurements at high cadence. At such frequency an operator cannot possibly keep track of even the most important of these at all times. This intensity combined with the high degree of complexity of the outputs presents an overwhelming challenge to the human mind to handle and the consequence is that suboptimal decisions are made.

In this paper, a novel method is suggested to achieve the best possible, i.e. optimal, efficiency at any moment in time, taking into account all outputs produced as well as their complex interconnections. This method yields a computed efficiency increase in the range of one percent. Moreover, this efficiency increase is available uniformly over time effectively increasing the base output capability of the plant or reducing the CO2 emission of the plant per megawatt.

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, $\mathbf{H} = (\mathbf{x}^{(-h)}, \mathbf{x}^{(1-h)}, \mathbf{x}^{(2-h)}, \ldots, \mathbf{x}^{(0)})$ 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 the 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 all one needs to know about this plant in order to model it.

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 is represented by the black box. Control theory now aims to discover the relationship between input and output by performing experiments. If we send a particular signal in, then we observe another signal coming out. Given enough such data and some analysis, control theory provides tools for creating a set of (differential-) equations that govern the behavior of the black box. The resulting set of equations is called a mathematical model. A crucial element is the time evolution of the process, i.e. an action at a certain time will have some effects immediately, some effects over a short-term and other effects over the long-term – this time dependency must be contained in the model for realistic results.

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 necessary mathematical methods that allow recurrent neural networks to be trained efficiently for large datasets coming from real industrial facilities have been invented only in 2005 and thus these methods can only be applied now.

The advantages of using machine learning over a human engineered model are (1) that the model is produced within a very short time (usually days), (2) that it is adaptive (i.e. it learns continuously as it experiences more data), (3) that it can change to match new situations (the new data is learnt) and (4) that the entire problem can be modeled (and not a simplified version as in the manual approach). Thus, (5) 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 outside the plant. 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 as functional outputs. The plant efficiency is, of course, among the semi-controllable outputs of the function f(…).

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 a single measurement point but we may get complex such as the plant efficiency and even 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 long 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 the particular plant considered here, Reuter-West in Berlin, eight months and nearly 2000 measurement locations were selected as the history that was recorded at one value each every minute; yielding approximately 0.7 billion individual data points. 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.

The power plant is largely automated and so we considered, for test purposes, only the district heating portion of the plant to be under the influence of the optimization program. The controllable variables would then be the flow rate, temperature and pressure of the of the district heating water at various stages during the production.

The boundary conditions or uncontrollable parameters are provided by the coal quality, the temperature, pressure and humidity of the outside air, the amount of power demanded from the plant, the temperature demanded for the district heating water in the district and the temperature of the cooling water at various points during the production.

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. This is a significant gain in coal purchase but mainly a reduction of the CO_2 emissions that save valuable emission certificates.

In the analysis, about 800 different operational conditions (in the eight month history) were identified that the operators would have to react to. This is not practical for the human operator. The model is 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.

Specifically, an example situation may be that the outside air temperature changes during the day due to the sun rising. It could then be efficient to lower the pressure of district heating water by 0.3 bars. The program would make this suggestion and after the change is effected, the efficiency increase can be observed.

5 Conclusion

The main benefits of the current approach are: (1) processes *all* measured parameters from the plant in *real-time*, (2) encompasses all *interactions* between these parameters and their time evolution, (3) provides a *uniform* and *sustainable* operational strategy 24 hours per day and (4) achieves the *optimal* operational point and thus smoothes out variations in human operations.

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.

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 a variety 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.

6 References

[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

Increase of Overall Combined-Heat-and-Power (CHP) Efficiency via Mathematical Modeling

Dr. Patrick Bangert, CEO
algorithmica technologies GmbH
Außer der Schleifmühle 67, 28203 Bremen, Germany; Tel: +49 (0) 421 337-4646
p.bangert@algorithmica-technologies.com
www.algorithmica-technologies.com

Jörg-A. Czernitzky, Power Plant Group Director Berlin
Vattenfall Europe Wärme AG
Puschkinallee 52, 12435 Berlin, Germany; Tel: +49 (0) 30 267-15525
joerg-andreas.czernitzky@vattenfall.de
www.vattenfall.de

Proceedings of the VGB PowerTech e.V. Conference "Fachtagung Dampferzeuger, Industrie- und Heizkraftwerke 2010" (conference on steam generation, industrial power and heat plants 2010), 31.08.–2.9.2010 in Hannover, Germany.