

## The Role of Predictive Models in Energy Efficiency Optimization of Complex Industrial Plants

Karel Mařík, Jiri Rojíček, Petr Stluka

Honeywell Prague Laboratory, V Parku 2326/18, 148 00 Prague 4, Czech Republic  
e-mail: {Karel.Marik; Jiri.Rojicek; Petr.Stluka} @Honeywell.com

Energy management of complex industrial plants covers the supply side (utility plant) as well as the demand side (process plant). The energy optimization concept itself can be characterized as a closed-loop optimization in the former case, while in the process plant the focus is rather on continuous improvement of energy efficiency. Energy demand modeling and forecasting is an important aspect that touches both sides. The better models and predictions imply the better decisions can be made and the larger cost reductions achieved. The paper provides an industrial research perspective on the area of plant-wide energy optimization, while the primary focus is on the process of building predictive models of various types, and on their subsequent exploitation under different optimization scenarios.

### 1. Introduction

Current environmental, legislative, and economic conditions require plants in energy intensive industries to pay special attention to monitoring and optimization of energy efficiency and carbon emissions. In oil refineries and other big industrial complexes like petrochemicals, chemicals, pharmaceuticals, or paper-making plants, the utility plant is responsible for the major supply of energy – primarily steam and power – to the process plant. The energy can either be generated in own facilities, or purchased from local distribution companies. Frequently, the utilities have a contract allowing them to sell excessive amounts of energy back into the electricity grid and take advantage of variable tariffs. Depending on local conditions, the industrial utility plant may also serve as a source of heating for neighboring residential areas. The typical energy flows are illustrated in *Figure 1*.

Although both industrial utility and process plants are tightly interconnected, their operational and business objectives are different. In the utility plants, the generation of energy directly is the primary business objective, which is also consistently addressed throughout the facility by adopting hierarchical solutions for closed-loop real-time optimization of individual pieces of equipment (boilers, turbines) or their groups (several steam boilers connected to the same header). In contrast to that, the process plants are primarily driven by the objective to produce appropriate mix of products to meet orders coming from the downstream industries. Energy contained in the consumed utilities is the second largest operating cost – after the cost of raw materials – and the

general desire is to reduce this cost as much as possible, but never in a way that could threaten timely delivery of products. This is why in the current operating practice the production optimization is handled independently of the energy efficiency optimization.

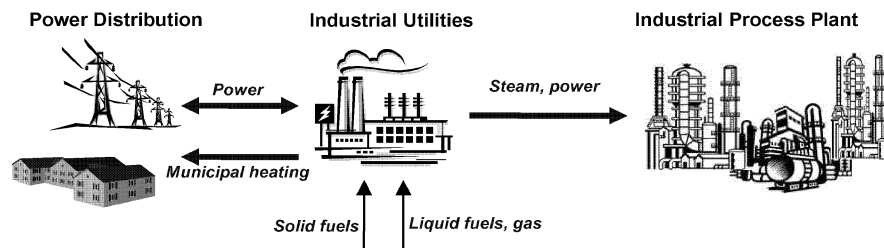


Figure 1: Major flows of energy in industrial utility and process plants

## 2. Utility Plant Energy Management

Energy management in the industrial utility plants includes optimization applications organized in several layers (Havlena, 2007).

- The basic level is focused on *real-time optimization* of individual pieces of equipment – basically the pressure control related devices like boilers, letdown valves and vents, but also other types of more complex equipment including turbo generators or condensing turbines. Multivariable predictive control techniques fit very well into this area as described e.g. by Findejs (2008).
- The second level applications deal with the problem of *optimal allocation of load* between several pieces of equipment running in parallel. This task is usually executed in real-time to ensure fast response to dynamically changing conditions and requirements coming from the process plant.
- Lastly, the third level applications optimize operation of the utility plant over significantly longer periods of time – ranging from hours to days – taking into account multiple possible configurations of the utility plant that can be selected for meeting the energy demand requirements. Flexible starts and stops of some pieces of equipment are assumed, which are mathematically translated into MILP type of *optimal scheduling* problem.

The concrete implementations may differ so that the second level applications can either be bundled with the real-time optimization components into one solution package, or the load allocation problem can be solved as a subtask of the multi-period MILP optimizer. Solution details and practical results achieved when optimizing operation of a CHP plant using the second approach were described by Schindler (2004) or more recently by Mařík et al. (2008).

Both approaches also differ in the way how they deal with varying energy demands. In the closed loop approach this variable usually is considered as the disturbance variable that comes from the outside of the utility plant and in the relatively short time frame from 5 to 60 minutes they cannot be efficiently predicted.

On the other hand, the utility plant scheduling application cannot generate reliable schedules without having a reasonable energy demand predictions for the given time interval. Generally, those demand predictions can be combinations of two different types of demand: (a) *internal demand* of the process plant that is dictated by production objectives and needs in individual technological processes; (b) *external demand* of local districts and residential areas that is driven by behavioral patterns of their inhabitants.

### **2.1 External Demand Modeling**

External residential energy demand is affected mainly by weather conditions, calendar-based variables, and seasonal effects. (Beran et al., 2006) *Ambient temperature* usually has the key impact, while the other environmental factors like *humidity*, *wind speed*, *cloud cover*, or *sun irradiation* can sometimes be used for better interpretation and finer modeling of the demand data. Calendar-based variables can efficiently help with capturing the behavioral patterns. These variables include *time of day*, which is defined on closed interval  $[0;1]$  where 0 corresponds to 0:00 and 1 to the midnight, and also categorical variables like *day of week*, *holiday* and *special day*, which cause clustering of similar days into coherent groups.

Modeling of the residential demand has been area of active research in the recent decades and the various techniques adopted to solve this problem could be categorized as artificial intelligence-based methods – including neural networks, expert systems, or support vector machines - and statistical methods, represented by e.g. similar day method, exponential smoothing and time series regression methods. (Weron, 2006)

### **2.2 Internal Demand Modeling**

Internal demand of the process plant is primarily given by the production plan or schedule that determines what products to make, by when, and how much. In a theory, there should be a good matching between such high-level production schedules and amounts of energy consumed in the process plant. However, there typically are more or less significant deviations caused by changing process conditions, operating modes, specific product grades, feedstock properties (e.g. type of crude oil being processed), or environmental conditions like ambient temperature. Techniques for developing suitable predictive models are discussed in the paragraph 3.2.

For purposes of utility plant optimization, it is necessary to translate schedules for individual process units into corresponding unit-level energy demands, and aggregate them consequently into one figure representing the whole process plant.

## **3. Process Energy Management**

Due to significant complexity of the overall site optimization, a hierarchical top-down approach is adopted nowadays as the standard, which includes production planning, scheduling, and real-time optimization layers. Outputs of corporate planning tools are fed into plant-wide production planners and schedulers that generate targets for individual processing units. Those are further projected into real-time optimizers (multi-unit or single-unit) in the form of limits. Given the overall complexity and primary focus on production, the currently used production planning and scheduling tools perform just material balance calculations without paying much attention to the energy

intensity of the developed schedule. Utility costs are normally considered only at the process unit level.

### 3.1 Energy Management at Process Unit Level

Process unit operators have responsibility for meeting production targets, formulated in terms of feed and product quantities and qualities, while minimizing operating costs on given unit. This is where the energy demand modeling comes into play.

Predictive models of internal energy demand – covering electricity, steam, and fuel – are used for estimation of the future amounts of energy needed by major process units like atmospheric and vacuum distillation, or fluid catalytic cracking. The energy use is determined primarily based on the currently executed production schedule and its parameters, which can be:

- *Discrete variables* – modes, regimes, campaigns, types of raw materials, etc.
- *Continuous variables* – typically production volumes and qualities
- *Time variables* – start times and end times of planned modes and campaigns

The energy use is also influenced by changing key process variables and environmental conditions like ambient temperature, both typically not captured in the schedule.

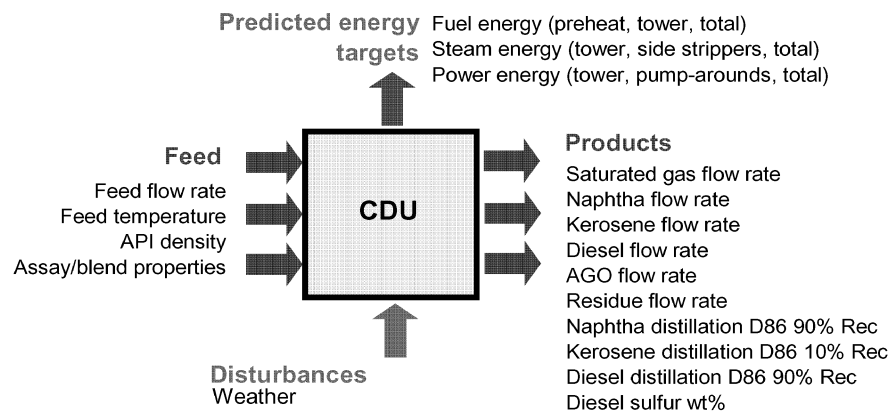


Figure 2: Variables used for energy target modeling on refinery crude distillation unit

Predictive models can be used for two different purposes: (1) generate estimates of expected energy consumption, which is consequently used as the operating target for unit operators; (2) advise operators on possible operational improvements in situations when the unit is deviating significantly from the expected energy use. Different level of detail is needed for each.

$$\text{Energy targets} = f(\text{Production targets}, \text{Disturbances}) \quad (1)$$

Energy target models are not intended to provide precise numbers, and one of their important outcomes is the estimated variance of energy consumption, which indicates how consistently is the unit operated and controlled.

$$\text{Energy} = f(\text{Production targets}, \text{Action variables}, \text{Disturbances}) \quad (2)$$

Models used for advising operators must be based on more detailed modeling, requiring to capture relationships between production quantities, qualities, energy consumption and action variables.

### 3.2 Modeling Approaches

The most popular modeling techniques used in the process industries include first principle models represented by various commercially available flowsheeting tools, and then also statistical data-driven modeling techniques, which build models by fitting relevant process data.

The *first principle models* have the advantage of being based around fundamental understanding of the physics and chemistry of given process, and if accurate, can be used within a wide range of operating conditions. The downside is that these types of model are relatively costly to build and maintain, and sometimes also not easy to tune. The *data-driven models* can be derived by some form of statistical regression or non-linear fitting of data using various black-box modeling techniques. The data-driven models are more easily to build, they are also more robust and better dealing with natural variations in process measurements, but on the other hand, they are not useful outside the data set used to generate them.

When creating the high-level energy target models for a process unit, the first choice usually is a simpler statistical model like the ordinary linear regression, which allows to calculate energy consumption based on actual production parameters. This approach works well until it is possible to effectively select input variables for the regression, using e.g. the domain knowledge about given process unit. When this is impossible, it is necessary to apply more sophisticated techniques for multi-dimensional problems:

- *Dimensionality reduction* – transformation of original variables into lower-dimensional set of new variables using e.g. Principal Component Analysis.
- *Multivariate statistical techniques* – e.g. linear or non-linear Partial Least Squares methods
- *Statistical learning techniques* – neural networks, regression and classification trees, ensemble methods like bagging or boosting.

Creating efficient detailed energy optimization (i.e. operator advisory) models is significantly more challenging because the set of considered variables is extended by the action variables. In this case the first principle models may represent viable alternative to the data-driven models. But also now it is necessary to consider overall effort needed for model setup and maintenance. An interesting strategy may be to use the first principle model for systematic evaluation of many possible process states. Running the model with changing input variables may help to generate a lookup table that can be subsequently used as a fast method for optimization purposes.

Another approach may be represented by the local modeling techniques (Atkeson et al., 1997), which are based on the idea of using simple models around actual process states instead of complex global models. The local model is usually generated on-the-fly and it attempts to fit the training data only in the region around the given operating point. The local model itself cannot be used to find the global optimum values of action variables but it can advice about directions for possible improvement and this strategy can be applied recursively.

A serious practical problem, which is common to any real life applications of modeling in the process industries, is the frequent occurrence of data quality issues, including missing values and measurement errors (outliers). Modeling methods must be robust enough to be able to cope with the imperfect data.

The missing data points can be generally treated in three ways: they can be imputed, ignored or this problem can be eliminated by the choice of a suitable modeling technique. Simple linear imputation is the least appropriate as it can spoil subsequent modeling and optimization processes. In spite of that, ignoring the missing data points can be a better method for handling such problems. However this method is reasonable only if there is only a small portion of records with missing values. Lastly, a quite efficient approach can be to choose a modeling method that is specifically designed for dealing with the missing data, like e.g. the classification and decision trees do.

#### **4. Conclusion**

The paper summarized energy optimization concepts used in the industrial utility plants and process plants, as well as related types of predictive models, which include:

- *Models for external energy demand of residential areas* represent a well studied research field. The prediction accuracy can be very good because a relatively small number of influencing variables needs to be considered, and the daily consumption profiles are usually smooth.
- *Models for internal energy demand of individual process units* represent much bigger practical challenge because specifics of every unit must be considered, always starting with a different set of variables, which is considerably larger than that used for external demand modeling. The high-level models can be used for setting energy targets based on existing production plan, while the detailed models may be used to advise operators to take appropriate actions aiming at improved energy efficiency.

#### **References**

- Atkeson Ch.G., Moore A.W. and Schaal S., 1997, Artificial Intelligence Review, Volume 11 , Issue 1-5, pp. 11-73
- Beran, Z, Mařík K. And Stluka P., 2006, Data-Centric Demand Forecasting for Utilities, Computer-Aided Chemical Engineering, Volume 21, Elsevier, Amsterdam
- Findejs J., Havlena V. and Pachner D., 2008, Multivariable predictive circulating fluidized bed combustor control, AT&P Journal PLUS2
- Havlena V., Pekař J., 2007, Advanced Process Management for Power and Industrial Energy, In: ESCAPE-17, Bucharest.
- Mařík K., Schindler Z. and Stluka P., 2008, Decision support tools for advanced energy management, Energy, Volume 33, Issue 6, 858-873, Elsevier
- Schindler Z. Efficient load forecasting and optimal resource allocation, 2004, In: AICARR international conference, Milan.
- Weron R., Modeling and Forecasting Electricity Loads and Prices: A Statistical Approach, 2006, John Wiley & Sons, Ltd.