Development of an energy management system for a naphtha reforming plant: A data mining approach

David Velázquez [a], Rocío González-Falcón [a,*], Luis Pérez-Lombard [a], L. Marina Gallego [a], Iñigo Monedero [b], Félix Biscarri [b]

[a] School of Industrial Engineering, Energy Engineering Department, University of Seville, Camino de los Descubrimientos s/n, 41092 Seville, Spain
[b] School of Computer Science and Engineering, Electronic Technology Department, Av. Reina Mercedes s/n, 41012 Seville, Spain

Keywords:

Energy management systems
Improve energy efficiency
Data mining
Naphtha reforming

A B S T R A C T

Despite the industrial sector accounts for about a quarter of total final consumption worldwide and great efforts have been carried out to reduce its energy use in the last decades, there are still substantial opportunities to improve industrial energy efficiency. Among those opportunities, energy management systems (EMSs) are one of the most successful and cost-effective ways to significantly reduce energy use, energy costs and environmental impact without affecting production and quality. This paper describes the development of an energy management system for a naphtha reforming plant by the use of a data mining approach. The paper shows how these techniques have been applied to identify key influence variables on energy consumption and to develop an energy performance model of the plant. Energy baseline and energy targets have been derived for the assessment of achieved and potential energy savings. Plant results show how savings may be achieved after the implementation of the EMS by tracking and adjusting performance against energy targets.

1. Introduction

Industry is a significant energy end-use sector comprising 27% of the world's total final consumption in 2009 [1]. Energy-intensive industries (bulk chemicals, refining, paper products, iron and steel, aluminum, food, glass, and cement) dominate industrial energy demand, accounting for nearly two-thirds of industrial delivered energy consumption [2]. Additionally, worldwide projections on industrial energy consumption predict a growing trend over the next 25 years with an annual average rate of 1.5% [3]. However, substantial opportunities to improve industrial energy efficiency have been shown by the International Energy Agency (IEA) [4], and much of this potential can be captured through policies for the promotion of the implementation of energy management systems (EMSs). In this sense, the IEA has advised governments to require high energy-intensive industries to comply with ISO 50001 [5] or an equivalent energy management protocol.

Refining is a high-consuming industry, accounting for about 7% of total energy consumption in US in 2002 [6]. A large variety of opportunities exist within petroleum refineries to reduce energy consumption while maintaining or enhancing productivity, as clearly shown by competitive benchmarking data indicating that most refineries can economically improve energy efficiency by 10–20% [7]. Particularly, implementing EMS is one of the most successful and cost-effective ways to bring about energy efficiency improvements.

Energy management aims to minimize energy costs and environmental impact without affecting production and quality [8] by the achievement of continuous improvement of energy performance, energy efficiency and energy conservation. Energy performance indicators (EPIs) should be identified to assess energy performance and to subsequently evaluate progress towards objectives and targets. Thus, measuring baseline performance, setting goals and tracking performance against those goals are the keystones of every EMS [9].

Industrial patterns of energy use are complex, especially when production rates are highly variable, when the product mix varies, or when several interacting processes coexist at a single site. Recently, the availability of massive performance data has increased the interest in the application of data mining to industrial energy management. Data mining is the process of extracting valid, new and comprehensive information from massive data in order to improve and optimize business decisions [10]. A multitude of methods are available for carrying out data mining procedures, however, within the industrial field, they focus on process monitoring and control, soft sensors, expert systems, fault detection and diagnosis [11]. Many industries have applied data mining by the use of statistical methods [12–14] and neural networks or decision trees [15–19]. In the last decade, the research on crude distil-

* Corresponding author. Tel.: +34 954 482283.
E-mail address: rgf@us.es (R. González-Falcón).

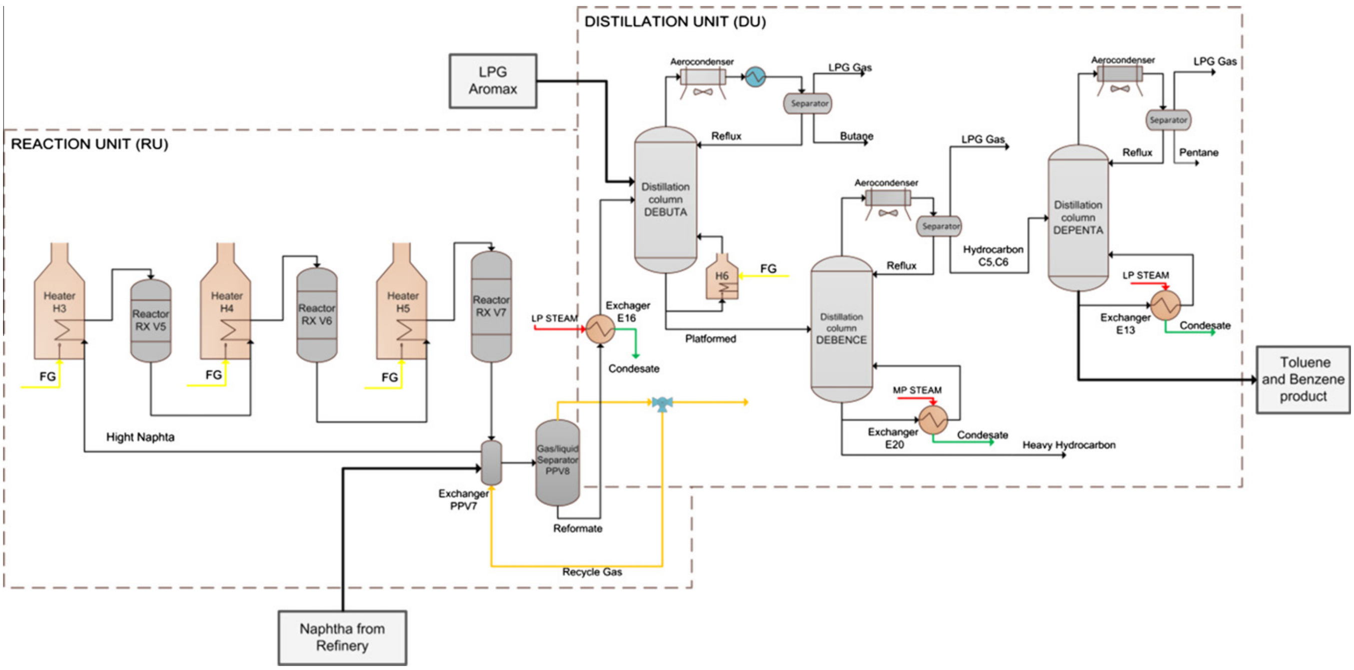

Fig. 1. Process flow diagram of the platforming unit.

Table 1
Annual energy consumption by energy source and unit.

Unit	Energy source	Equipment	Energy use (GWh/year)	Energy use (%)
RU	Fuel gas	Heater H3	62.8	27
	Fuel gas	Heater H4	68.5	29
	Fuel gas	Heater H5	25.9	11
DU	Fuel gas	Desbutanizer Heater H6	37.6	16
	Med pres. steam	Desbencenizer Boiler E20	25.5	11
	Low pres. steam	Despentanizer Boiler E13	12.9	6
Total			233.2	

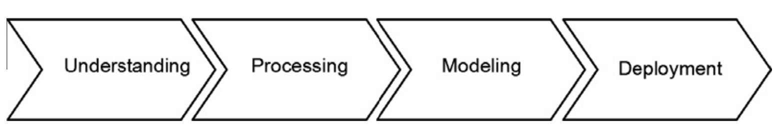

Fig. 2. Phases of a data mining project.

lation processes has put the focus on control and optimization. Data mining has also been applied to predict quality of distillation products [20,21] and to estimate the optimal operating conditions of distillation process [22]. However, data mining has not been used so far for the development of an EMS, despite energy management has been identified as a priority for energy optimization in refineries [23].

This paper presents the application of a data mining approach to the definition, development and implementation of an EMS in a naphtha reforming plant with the aim of assessing energy performance and evaluating progress towards efficiency targets. The paper shows the way to select the key influence variables for energy performance, to identify the best performing periods in the past in order to set energy performance targets for the future, and to develop a baseline model. The main contribution of the paper is the development of an EMS for a naphtha reforming plant that allows measuring baseline performance, setting targets and tracking energy performance against those targets.

The paper starts with the description of the naphtha reforming plant (Section 2) and the definition of EMS requirements (Section 3). Then, in Section 4, the research methodology is presented. Section 5 provides the main results of the application of the data mining approach to the naphtha reforming plant. Finally, concluding remarks, benefits and limitations of the research are discussed.

2. Naphtha reforming plant

This section briefly describes the platforming unit and analyses its energy structure in order to define the requirements for the implementation of the EMS.

Catalytic reforming of heavy naphtha is a key process in the production of gasoline. The major components of petroleum naphthas are paraffins, naphthenes, and aromatic hydrocarbons. The aim of catalytic reforming is to transform naphthas hydrocarbons with low octane to hydrocarbons with high octane. The chemical

Table 2
Data mining project objectives and requirements.

EMS phases	DM phases	DM objectives	DM inputs
Plan	Understanding	– Understand process – Analyse energy use – Define energy performance indicator – Identify PIV	– Process flow diagrams – Historical energy data – Other historical data
	Processing	– Obtain a suitable data set with EPI and PIV	– Energy consumption data – PIV data – Suitable data set with EPI and PIV
	Modeling	– Obtain discriminant functions – Identify KIV – Classify external and controllable variables – Obtain EPI regression models	
	Deployment	– Obtain EPI targets – Obtain EPI baseline	– Discriminant functions – EPI regression models – Actual measurement of KIV – Past measurement of EPI and KIV
Check	Deployment	– Monitor actual EPI against targets – Calculate achieved energy savings – Calculate potential energy savings – Calculate efficiency ratio	– Actual measurement of EPI – EPI targets – EPI baseline

Table 3
List of potential influence variables (PIVs).

ID	Variables	Unit	Description
1	Tin_RX_V5	°C	Input temperature reactor RX_V5
2	Tin_RX_V6	°C	Input temperature reactor RX_V6
3	Tin_RX_V7	°C	Input temperature reactor RX_V7
4	Tout_RX_V5	°C	Output temperature reactor RX_V5
5	Tout_RX_V6	°C	Output temperature reactor RX_V6
6	Tout_RX_V7	°C	Output temperature reactor RX_V7
7	P_PP_V8	bar	Pressure in separator PP_V8
8	RGD	kg/Nm^3	Recycle gas density (P_PP_V8)
9	PU_load	m^3/h	Input stream to the platforming unit.
10	LPG Aromax	m^3/h	Input stream from Aromax unit to debutanizer
11	WAIT	°C	This variable of the RU explain the catalyst degradation
12	F_PP-V8	m^3/h	Output stream from separator PP_V8 to debutanizer
13	F_PP_V11	m^3/h	Input stream to debutanizer
14	FL_PG GAS	m^3/h	Gas fraction of the top flow of debutanizer
15	FL_PG LIQ	m^3/h	Liquid fraction of the top flow of debutanizer
16	FG_H3	m^3/h	Fuel gas consumption in heater H3
17	FG_H4	m^3/h	Fuel gas consumption in heater H4
18	FG_H5	m^3/h	Fuel gas consumption in heater H5
19	FG_H6	m^3/h	Fuel gas consumption in heater H6
20	VM_E16	kg/h	Low pressure steam consumption in reboiler E16
21	DEBE_load	m^3/h	Bottom stream of debutanizer
22	DEBE_Top flow	m^3/h	Gas fraction of the top flow of debenzenizer
23	DEBE_Bottom flow1	m^3/h	Liquid fraction of the bottom flow of debenzenizer to YT_135
24	DEBE_Bottom flow2	m^3/h	Liquid fraction of the bottom flow of debenzenizer to gasolines
25	VM_E20	kg/h	Medium pressure steam consumption in reboiler E20
26	DEPE_Top flow	m^3/h	Gas fraction of the top flow of depentenizer
27	DEPE_Bottom flow	m^3/h	Liquid fraction of the bottom flow of depentenizer
28	VM_E13	kg/h	Medium pressure steam consumption in the reboiler E13
29	RR debuta		Reflux ratio: debutanizer's top stream/liquid return to the column
30	RR debenze		Reflux ratio of debenzenizer
31	RR depenta		Reflux ratio of depentenizer
32	P_V13	Bar	Pressure operation of debenzenizer
33	T_top_V13	°C	Top gases output temperature of the debenzenizer
34	T_bottom_V13	°C	Bottom liquid output temperature of the debenzenizer
35	P_V11	Bar	Pressure operation of debutanizar
36	T_top_V11	°C	Top gases output temperature of the debutanizer
37	T_bottom_V11	°C	Bottom liquid output temperature of the debutanizer
38	P_V12	Bar	Pressure operation of depentenizer
39	T_top_V12	°C	Top gases output temperature of the depentenizer
40	T_bottom_V12	°C	Bottom liquid output temperature of the depentenizer
41	DT_total	°C	Global Delta T in reactor
42	T ambient	°C	Ambient temperature
43	DT_RX_V5	°C	Delta T in PP_V5 reactor
44	DT_RX_V6	°C	Delta T in PP_V6 reactor
45	DT_RX_V7	°C	Delta T in PP_V7 reactor

reactions that lead to these changes are guided by a catalyst under well-defined operating conditions [24].

Naphtha reforming plants are composed of different operating units to separate fractions and improve their quality. In particu-

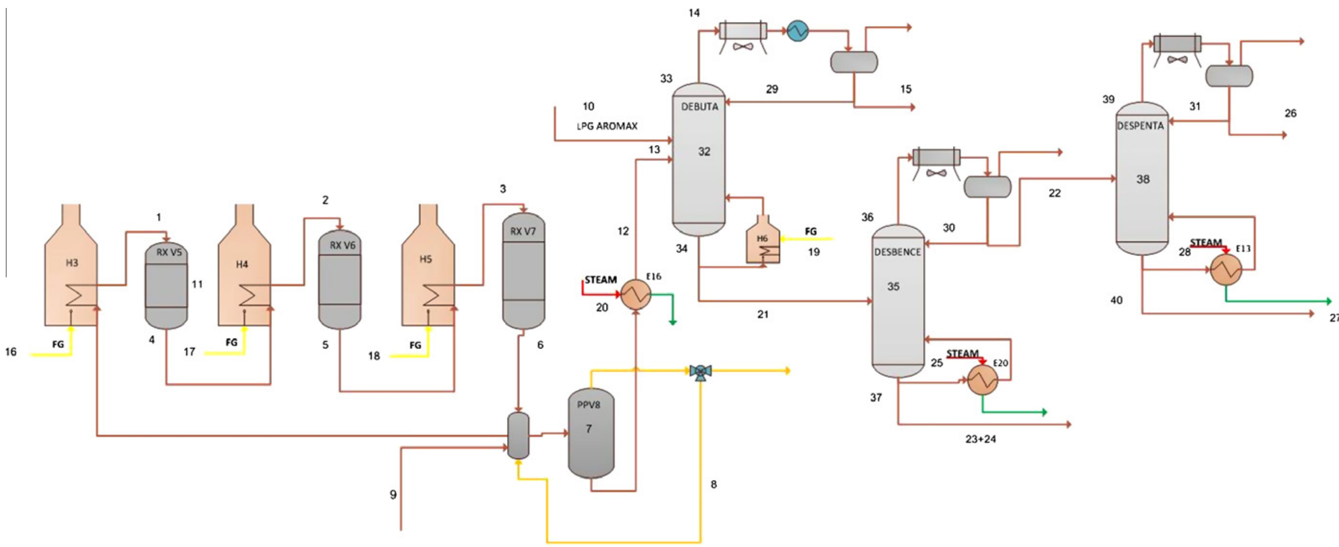

Fig. 3. Location of the potential influence variables in the platforming unit.

Table 4
Categorization of the energy performance indicators (EPIs).

Concept hierarchies	EPI discrete value	SC_{PU} (kWh/m^3)	SC_{RU} (kWh/m^3)	SC_{DU} (kWh/m^3)
Very low	−2	<405	<281	<114
Low	−1	$405 \leqslant x < 462$	$281 \leqslant x < 328$	$114 \leqslant x < 129$
Normal	0	$462 \leqslant x \leqslant 576$	$328 \leqslant x \leqslant 423$	$129 \leqslant x \leqslant 158$
High	1	$576 < x \leqslant 633$	$423 < x \leqslant 470$	$158 < x \leqslant 172$
Very high	2	$x > 633$	$x > 470$	$x > 172$

lar, this research focuses on the platforming unit (PU). A process flow diagram of the PU is presented in Fig. 1. The unit might be divided into the reaction unit (RU), where the catalytic reforming is carried out, and the distillation unit (DU). The RU consists of three adiabatic reactors (RX V5, RX V6 and RX V7) in series with intermediate heaters (H3, H4 and H5) through which naphtha is converted to high-octane aromatics and others hydrocarbons. Then, the stream enters the product separator (PPV8) where flash separation of hydrogen and some light hydrocarbons is carried out. The desired reformate is then pumped to the DU where high-octane aromatics are separated of light hydrocarbons traces. Separation of products is made in three distillation columns: debutanizer, debencenizer and depentanizer. The first column is fed with the liquid flow from the RU and a stream of aromatic LPG from other refinery processes. Basic products are butane to the top and platformed stream (rich in aromatics with high octane) to the bottom. Platformed stream is fed to the debencenizer to obtain hydrocarbon C5–C6 to the top, which is fed to depentanizer, and heavy hydrocarbon to the bottom. Finally, the desired products of depentanizer are toluene and benzene to the bottom. The bottom product of the debencenizer and the top product of debutanizer and depentanizer are stored up or sent to other units of the refinery.

Table 1 shows annual energy consumption by energy source (Fuel gas, medium pressure steam and low pressure steam) and by unit (Reaction and Platforming). Fuel gas is consumed in the reactors and desbutanizer heater while steam is used in desbencenizer and despentanizer boilers. The major energy consumers are the three reactors and the debutanizer column. Annual energy costs of the platforming unit are around 10.9 M€.

Table 5
Description and classification of key influence variables of SC_{PU}.

ID	KIV	Units	Type
8	Recycle gas density (P_PP_V8)	kg/Nm^3	External
9	Input Stream to the Platforming unit	m^3/h	External
10	Input stream from Aromax unit to debutanizer	m^3/h	External
16	Fuel gas consumption in heater PP_H3	m^3/h	Controllable
17	Fuel gas consumption in heater PP_H4	m^3/h	Controllable
18	Fuel gas consumption in heater PP_H5	m^3/h	Controllable
19	Fuel gas consumption in heater PP_H6	m^3/h	Controllable
25	Medium pressure steam consumption in the debenzenizer	kg/h	Controllable
29	Reflux ratio: debutanizer's top stream/liquid return to the column	–	Controllable
30	Reflux ratio of debenzenizer	–	Controllable
32	Pressure operation of debenzenizer	Bar	Controllable
42	Ambient temperature	°C	External
43	Delta T in PP_V5 reactor	°C	External
45	Delta T in PP_V7 reactor	°C	External

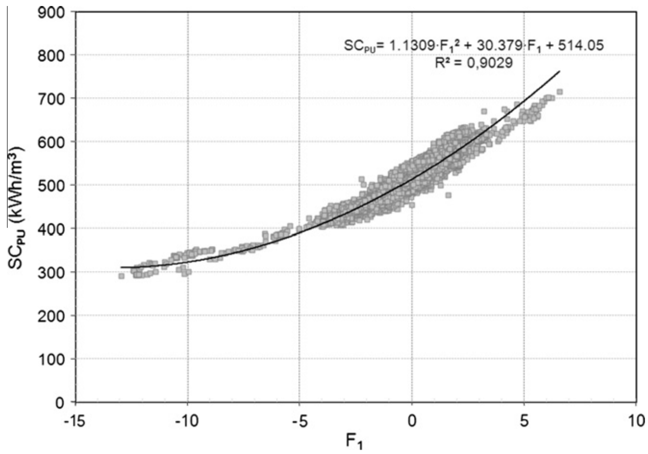

Fig. 4. Specific consumption of the PU (SC_{PU}) versus discriminant function (F_1).

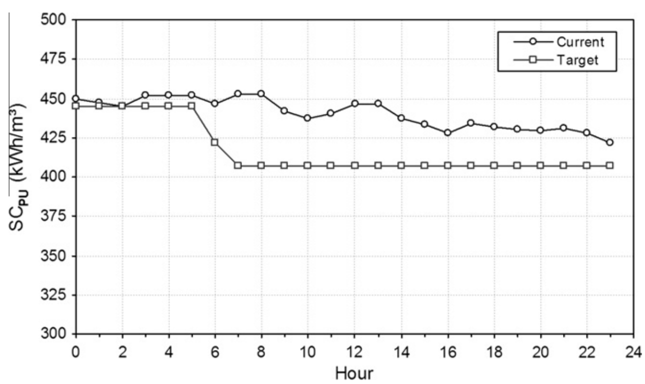

Fig. 5. Example of management report, showing hourly progress of current and target SC_{PU}.

The energy consumption of the plant in 2009 was 233.1 GWh, 67% in the RU and the remaining 33% in the DU.

3. EMS requirements

The aim of an EMS is to achieve continual improvement of energy performance, energy efficiency and energy conservation. Most energy management systems are based on the Plan-Do-Check-Act (PDCA) continual improvement framework [25,26]. 'Plan' phase aims to set the objectives and targets to deliver results in accordance with the opportunities to improve energy performance and the energy policy of the organization. 'Do' means implementing the plan. In the 'check' phase, energy performance is monitored and measured to assess plant energy performance and to evaluate progress towards targets. Finally, during the 'act' phase, corrective and preventive actions are identified and implemented in order to guarantee the accomplishment of the objectives.

This paper presents the application of a data mining approach to the Plan-Do-Check phases of an EMS. The specific requirements of these phases are described below.

3.1. Planning the EMS

Our planning of the EMS aims to define the energy performance indicators and to set energy targets and energy baseline. EPI are quantitative indices to assess, monitor and measure energy performance. Energy targets are detailed energy performance objectives to be reached and energy baseline is a quantitative reference for comparison of energy performance.

First of all, energy performance indicators must be defined. EPI generally refer to specific consumption and may combine energy sources associated with a particular process. However, in some processes the use of energy sources may not be closely related, and individual EPI for each source may provide a clearer picture of performance. Anyway, there is no EPI that can be applied in every situation and appropriate indicators have to be defined on a case by case basis [27].

Once EPI have been defined, a target for each EPI must be set. To ensure accurate monitoring of EPI with continuous operational improvement, we propose to set targets using best practices. A best practice target identifies what a process or plant could achieve if it would be best operated according actual conditions. These targets can be set through the best value of the EPI that the process has achieved in the past.

In addition to setting targets for continuous operational improvement, the EMS must clearly show that actions taken to reduce energy use have been successful to justify ongoing investment, to validate energy-saving decisions, and to demonstrate that improvements have been achieved. However, energy savings cannot be directly measured, but are to be calculated from a comparison of the energy baseline with the post implementation energy consumption [28]. The energy baseline is the energy consumption that would have occurred if no direct measures had been taken to influence energy consumption [29] and it should be defined before implementing any energy efficiency projects. We propose the use of a predictive model to set a baseline based on actual conditions to obtain reliable energy savings.

3.2. Checking the EMS

In the checking phase, EPI must be regularly measured and monitored against targets and baseline to assess record and report energy-savings. Achieved energy savings are calculated by comparing baseline with current performance (Eq. (1)), while potential savings are the difference between baseline and target energy consumptions (Eq. (2)). Then, the efficiency ratio (Eq. (3)) could be a measure of the success in achieving potential energy savings.

$$\begin{aligned} Achieved\ savings = {} & Baseline\ consumption \\ & - Actual\ consumption \end{aligned} \tag{1}$$

$$\begin{aligned} Potential\ savings = {} & Baseline\ consumption \\ & - Target\ consumption \end{aligned} \tag{2}$$

$$Efficiency\ ratio = \frac{Achieved\ savings}{Potential\ savings} \tag{3}$$

4. Methodology

Data mining (DM) is a hybrid discipline that integrates technologies of databases, statistics, machine learning, signal processing and high performance computing. It is helpful to analyze, understand or even visualize massive data gathered from business and scientific applications [30]. A systematic approach is required to be successful in a DM project. In particular, Cross Industry Standard Process for Data Mining (CRISP-DM) [31] generally involves four basic phases: understanding, processing, modeling and deployment. The main phases of CRISP Data Mining project could be represented in Fig. 2.

Understanding the business is the initial phase and focuses on defining project objectives and requirements from a business per-

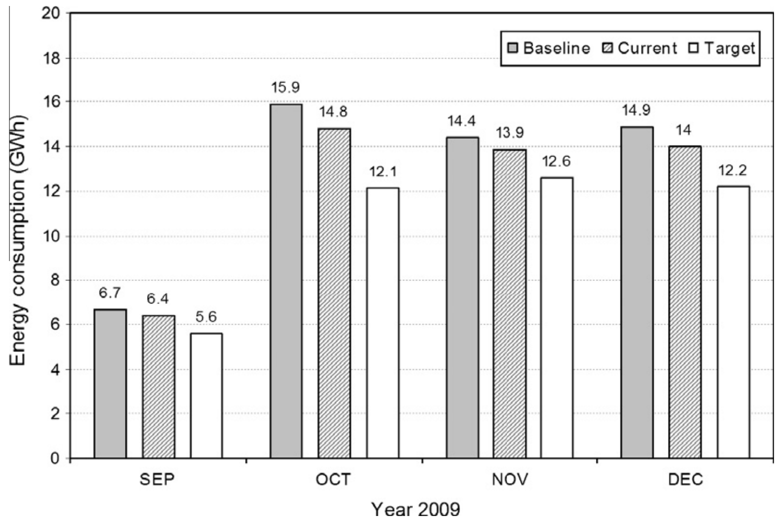

Fig. 6. Management report showing monthly current, target and baseline consumptions.

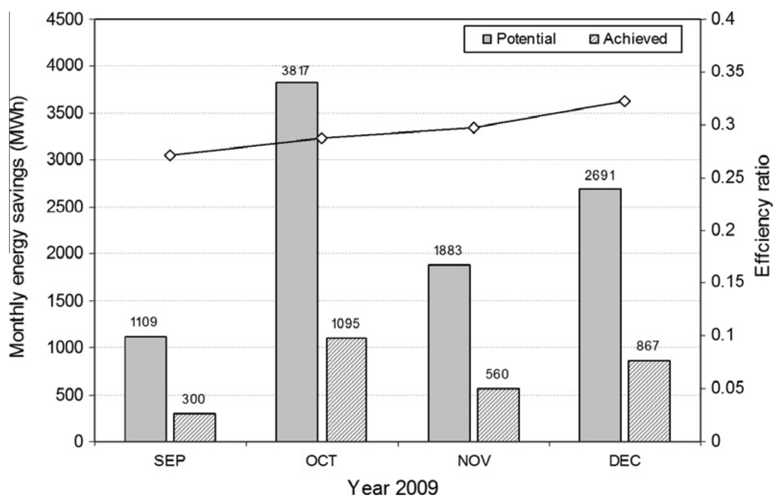

Fig. 7. Monthly potential and achieved energy savings.

spective. Processing collected data is mandatory to identify data quality problems and to discover first insights into the data. Data must be prepared to ensure that useful knowledge is derived from data and to increase the data quality before the data mining models can be successful applied. Typically, up to 90% of the time and effort in a data mining project is spent on data processing. In the modeling phase, various modeling techniques are selected and applied and their parameters are calibrated to optimal values. Finally, through the deployment phase, insight and actionable information can be derived. Deployment involves the integration of data mining models within applications, data warehouse infrastructure, or query and reporting tools. Table 2 shows the relations between the basic steps for the development of the EMS and the four phases of our DM approach. This table provides the objectives and inputs for each DM phases.

The initial DM phase focuses on understanding the process and the energy use through the analysis of process flow diagrams and the available historical data. It is important to focus on those areas which a high consumption to define energy performance indicators and to identify potential influence variables (PIVs) with may have a large impact on energy performance.

Once EPI have been defined and PIV have been identified, input data must be processed to obtain a suitable data set for the modeling phase. Before a useful model can be developed, data must be provided in the amount, structure, and format suited to modeling procedures. Processing phase starts with acquisition, description and quality assessment of data and ends with data preparation.

In the modeling phase, some techniques may have specific requirement on data and consequently, stepping back to the processing phase is often necessary. In our case a Discriminant Function Analysis (DFA) technique is selected to extract key influence variables (KIVs) from PIV and model EPI against these variables. DFA is a multivariate statistical technique that is commonly used to build models based on observed predictor variables. EPI must be categorized into several groups and the potential influence variables are normalized. DFA measures the relevance of potential influence variables and let us obtain discriminate functions, which identify KIV and show how these variables can be linearly combined to best classify EPI. Normalized variables (x_i^*) and discriminant functions (F) are defined as follows:

$$x_i^* = \frac{x_i - x_{i,mean}}{S_i} \tag{4}$$

$$F = \sum_{1}^{i} v_i \cdot x_i^* \tag{5}$$

where x_i is the potential influence variable i; $x_{i,mean}$ is the mean value of variable i; S_i is the standard deviation of variable i; v_i is the weighting coefficient for the variable x_i^*; and x_i^* is the normalized potential influence variable i.

KIV may be classified as external or controllable. An *external* influence variable is a variable that influences energy performance but over which operators have no control, while a *controllable* influence variable is, or needs to be, controlled.

In the last step of the modeling phase, a regression model of EPI versus discriminant function is obtained. This model condenses the dependence of energy performance on KIV and will be used to obtain EPI baseline.

Finally, in the deployment phase, the knowledge gained from DM must be organized and presented in a useful way for energy management. In this phase EPI targets and EPI baseline are calculated to be compared with actual EPI on a regular basis. EPI targets represent best-practice performance and are calculated with the best energy performance achieved in the past at current values of the external influence variables. The progress towards energy targets is tracked through efficiency ratio and potential energy savings, both periodically calculated and reported. EPI baseline is obtained through EPI regression model depending on. For equipment improvements there is no action on the operation conditions, so baseline would be obtained from actual key influence conditions. For operation improvements, which imply changes of controllable variables to improve energy performance, baseline would be adjusted to actual external influence conditions. This allows a correct evaluation of achieved energy savings from any implemented energy efficiency improvement.

5. Results and discussion

In this section we give details of the application of the DM approach to the planning and checking of the EMS of the naphtha reforming plant. Results of the understanding, processing, modeling and deployment stages are presented and discussed below. The software SPSS Modeler Clementine has been used as the basic tool for our data mining analysis.

5.1. Understanding

We aim to develop an EMS to track the energy performance of the PU and consequently, we must identify, define and measure EPI of the plant. In our case, we have chosen the specific consumptions of platforming, reaction and distillation units (Eqs. (6)–(8)) for energy performance assessment.

$$SC_{PU} = \frac{\text{PU energy consumption}}{\text{Platforming load}}\left(\frac{\text{kWh}}{\text{m}^3}\right) \tag{6}$$

$$SC_{RU} = \frac{\text{RU energy consumption}}{\text{RU load}}\left(\frac{\text{kWh}}{\text{m}^3}\right) \tag{7}$$

$$SC_{DU} = \frac{\text{DU energy consumption}}{\text{DU load}}\left(\frac{\text{kWh}}{\text{m}^3}\right) \tag{8}$$

Many variables are measured and collected in the PU. Some are direct measures of energy use, while others could affect it. Initially, 45 potential influence variables are selected and collected in an hourly frequency from March 2008 to August 2009. The 45 potential influence variables are described in Table 3 and Fig. 3 indicates the location of each one.

The whole sample is divided for analysis: the training set contains 12149 records, while the validation set contains the remaining 9402 records. The first is used to train the model, while the second is used to test its final performance.

5.2. Processing

In a first step of processing phase, 120 outliers corresponding to the days where the plant had unusual operation were filtered and deleted from the training sample. Besides, the resulting sample was filtered on the basis of two quality requirements: percentage of benzene should be less than 1% and percentage of toluene should be less than 10% at the bottom of the debenzenizer. After this filtering, 132 records are deleted from the training sample.

Once the training sample is filtered and cleaned, data set is prepared for the modeling phase. Data preparation includes EPI categorization and potential influence variables normalization. Categorization techniques would be used to reduce the number of values for a given continuous attribute, particularly for the EPI. Concept hierarchies are used to replace low-level concepts, such as numeric values, by higher level concepts, such as very high, high, normal, low and very low energy performance. EPI are categorized based on partitioning rules that split data distribution of the EPI into five groups and define the variation range for each one. The standard deviation (σ) of the EPI histogram defines the partitioning rules: very high ($>2\sigma$); high (σ,2σ]; normal [$-\sigma$,σ]; low (-2σ,σ] and very low ($<-2\sigma$). Table 4 shows variation ranges for the five EPI categorization groups.

5.3. Modeling

The main objectives of the modeling phase are obtaining discriminant functions, identifying KIV and developing EPI regression models. Two canonical discriminant functions are obtained for each EPI, but only one (F_1) has statistical and practical significance because it covers more than 90% of the variation between groups. For instance, in the DFA of SC_{PU} a total of 14 KIV are identified from the 45 potential influence variables and the discriminant function may be written as follows:

$$\begin{aligned} F_1 = &-0.388 \cdot x_8^* - 1.897 \cdot x_9^* - 0.057 \cdot x_{42}^* + 0.112 \cdot x_{10}^* \\ &- 0.701 \cdot x_{43}^* - 0.422 \cdot x_{45}^* + 0.19 \cdot x_{29}^* + 0.181 \cdot x_{30}^* \\ &+ 1.131 \cdot x_{16}^* + 1.160 \cdot x_{17}^* + 0.967 \cdot x_{18}^* - 0.378 \cdot x_{19}^* \\ &+ 0.486 \cdot x_{25}^* + 0.093 \cdot x_{32}^* \end{aligned} \tag{9}$$

The KIV are described and classified as external or controllable in Table 5. A total of six external influence variables are identified, being input stream to the PU and delta temperature in PP_V5 reactor the variables that most influence the EPI. The rest of the KIV are controllable and can be used to bring EPI closer to target through system operation improvements.

The regression model of SC_{PU} is presented in Fig. 4, showing a strong correlation with F_1 (R^2 = 0.9029). This model represents past typical performance versus KIV and will be used to obtain EPI baseline in the deployment phase.

5.4. Deployment

The main goal of this phase is monitoring actual EPI against target and baseline on a regular basis. The comparison of EPI with target allows personnel to identify and implement continuous operational improvements. The comparison of EPI with the baseline allows managers to assess the progress of the EMS through achieved energy savings and the efficiency ratio.

The energy performance of the plant is tracked through energy management reports. It is important to ensure that the information contained in these reports is displayed in such a way that it can be easily understood, interpreted and applied by their potential users.

An example of the tracking process of SC_{PU} is shown in Fig. 5, where actual and target performances are compared on an hourly basis. Poor performance can be detected due to deviation from the *target* causing a performance alert and the operator can make process changes in response. Investigation into the operational conditions could then be shared across the business, resulting in controllable variable changes within safety limits and consistent energy savings across the shifts. Often, alerting personnel of poor performance is enough; since personnel may be experienced to understand the reasons for high energy use and take appropriate remedial action. Poor performance detected during a report period may also highlight a meter failure and help in maintaining the integrity of the measurement system. Failure to do so will make the energy performance checking process difficult. Regular energy reports provide a tool for the early detection of failure of critical meters.

It is also useful to summarize the results of the actual performance versus the target and baseline in a clear and concise way. The EMS was implemented in mid-September 2009 and it was found that the EPI tracking process produces savings from operational improvements. Fig. 6 represents monthly energy consumption in the post-EMS period, clearly showing achieved energy savings ranging from 4% to 7%, so energy consumption and energy cost can be approximately monthly reduced in 830 MWh and 37.500 € respectively, if average prices of different energy sources are assumed. In this figure, it has only been recorded half of September, so the consumption of this month is significantly lower.

Fig. 7 shows monthly potential and achieved energy savings during post-EMS period. We found that total potential energy savings during the post-EMS period were 9.5 GWh but only 2.82 GWh energy savings have been achieved. If a constant baseline value, equal to the average for the pre-implementation period is assumed, energy savings would be 2.03 GWh, 28% lower than our proposal, clearly showing that a constant baseline does not account for the actual values of KIV, and highlighting the importance of adjusting the baseline to current operating conditions to accurately calculate energy savings. Monthly efficiency ratio grows from 27% to 32%, showing that the EMS is helping to improve plant energy efficiency through changes in controllable influence variables within safety limits. Plant technical manager sets safety limits for every controllable influence variable to avoid potential risk. The EMS also allows a correct tracking of performance improvement to other energy efficiency measures.

6. Conclusions

This paper provides a methodology for the definition and implementation of EMS in industrial sites where massive historical data are available based on a data mining approach. The results show that data mining helps to characterize the energy performance of the plant, identifying the key influence variables and modeling energy performance indicators against these variables. Key influence variables allow indentifying the best past operation in order to set targets for tracking future energy performance, promoting a continuous operational improvement. The predictive models are used to adjust the baseline to current influence conditions and to calculate energy savings due to both operational and equipment improvements.

The data mining approach has been successfully applied to the definition, development and implementation of an EMS for a naphtha reforming plant. A total of 14 key influence variables have been selected from 45 potential influence variables and an energy performance model has been developed. The EMS allows setting baseline and targets in real time, taking into account influence of actual operation conditions and evaluates potential and achieved energy savings. The success achieving potential energy savings is assessed by an efficiency ratio which grows from 27% to 32%, showing that significant savings have been achieved after EMS implementation.

Further investigation is necessary to develop decision-making supporting methods that could help in taking strategic decisions and corrective actions within the EMS.

References

[1] International Energy Agency. Online energy statistics. Energy balances; 2009. <http://www.iea.org/stats>.

[2] US Energy Information Administration. Annual energy outlook 2012. Early release overview; January, 2012. <http://www.eia.gov/forecasts/aeo>.

[3] US Energy Information Administration. International energy outlook 2011; September, 2011. <http://www.eia.gov/forecasts/aeo>.

[4] International Energy Agency. 25 energy efficiency policy recommendations by IEA; 2011. <http://www.iea.org/publications>.

[5] ISO 50001:2011. Energy management systems – requirements with guidance for use; 2011.

[6] US Department of Energy Office of Industrial Technologies. Energy and environmental profile of the US. Petroleum Refining Industry; November, 2007.

[7] Ernst Worrell, Christina Galitsky. Energy efficiency improvement and cost saving opportunities for petroleum refineries. An ENERGY STAR® guide for energy and plant managers. Ernest Orlando Lawrence Berkeley National Laboratory; 2005.

[8] Abdelaziz EA, Saidur R, Mekhilef S. A review on energy saving strategies in industrial sector. Renew Sust Energy Rev 2011:150–68.

[9] Van Gorp John C, CEM Services Marketing. Using key performance indicators to manage energy costs. Strategic planning for energy and the environment; 2005. p. 9–25.

[10] Hooke James H, Landry Byron J, Eng P, David Hart MAC. Energy management information systems: achieving improved energy efficiency: a handbook for managers. Engineers and operational staff. Office of Energy Efficiency of Natural Resources Canada; 2003.

[11] Platon Radu, Amazouz Mouloud. Application of data mining techniques for industrial process optimization. CANMET Energy Technology Centre – Varennes; 2007.

[12] Yoon S, Landry Jason, Kettaneh Nouna, Pepe William, Wold Svante. Multivariate process monitoring an early fault detection (MSPC) using PCA and PLS. In: Plant automation and decision support conference; 2003.

[13] Zhans Y, Dudzic MS. Online monitoring of steel casting processes using multivariate statistical technologies: from continuous to transitional operations. J Process Control 2006;16:819–29.

[14] Ahvenlampi T, Kortela U. Clustering algorithms on process monitoring and control application to continuous digesters. Informatics 2005;29:101–9.

[15] Zhou Y. Data driven process monitoring based on neuronal networks and classification trees. PhD dissertation. Texas A&M University; 2004.

[16] Edwardsy PJ, Murrayy AF, Papadopoulosy G, Wallacey AR, Barnard J. The application of neural networks to the paper-making industry. In: European symposium on artificial neuronal networks; 1999.

[17] Dayal Bhupinder S, MacGregor John F, Taylor Paul A, Kildaw R, Marcikic S. Application of feed forward neural networks and partial least squares regression for modelling kappa number in a continuous Kamyr digester. Pulp Pap Canada 1994;95(1):26–32.

[18] Devogelaere D, Rijckaert M, Leon OG, Lemus GC. Application of feed forward neural networks for soft sensors in the sugar industry. IEEE September Issue; 2002.

[19] Monedero I, Biscarri Félix, León Carlos, Guerrero Juan, González Rocío, Pérez-Lombard Luis. Decision system based on neural networks to optimize the energy efficiency of a petrochemical plant. Expert Syst Appl 2012;39:9860–7.

[20] Barbosa CH, Melo B, Vellasco M, Pacheco M, Vasconcellos LP. Bayesian neural networks on the inference of distillation product quality. In: VII Brazilian symposium on, neural networks; 2002.

[21] Fortuna L, Graziani S, Xibilia MG. Soft sensors product quality monitoring on debutanizer distillation columns. Control Eng Pract 2005;13:499–508.

[22] Motlaghi S. An expert system design for a crude oil distillation column with the neural networks model and the process optimization using genetic algorithm framework. Expert Syst Appl 2008:1540–5.

[23] California Energy Commission. Energy efficiency roadmap for petroleum refineries in California; April 2004. <http://www.eere.energy.gov/manufacturing/industries_technologies/petroleum_refining/>.

[24] Parkash Surinder, editor. Refining processes handbook. Elsevier; 2003.

[25] Hooke James H, Landry Byron J, David Hart MA. Energy management information systems – planning manual and tool. Office of Energy Efficiency of Natural Resources Canada; 2003.

[26] Gordić Dušan, Babić Milun, Jovičić Nebojša, Šušteršič Vanja, Končalović Davor, Jelic Dubravka. Development of energy management system – case study of Serbian car manufacturer. Energy Convers Manage 2010;51:2783–90.

[27] Bunse Katharina, Vodicka Matthias, Schönsleben Paul, Brülhart Marc, Ernst Frank O. Integrating energy efficiency performance in production management e gap analysis between industrial needs and scientific literature. J Cleaner Prod 2011:667–79.

[28] Common wealth of Australia. Energy savings measurement guide: how to estimate, measure, evaluate and track energy efficiency opportunities; 2008.
[29] Reichl Johannes, Kollmann Andrea. The baseline in bottom-up energy efficiency and saving calculations – a concept for its formalisation and a discussion of relevant options. Appl Energy 2011;88:422–31.
[30] Şencan Arzu. Modeling of thermodynamic properties of refrigerant/absorbent couples using data mining process. Energy Convers Manage 2007;48:470–80.
[31] Chapman Pete, Clinton Julian, Kerber Randy, Khabaza Thomas, Reinartz Thomas, Shearer Colin, et al. CRISP-DM 1.0. Step-by-step data mining guide. SPSS; 2000.