

# A New Proposed Energy Baseline Model for a Data Center as a Tool for Energy Efficiency Evaluation

Spiros Livieratos<sup>1\*</sup>, Spiros Panetsos<sup>1</sup>, Avgoustinos Fotopoulos<sup>2</sup> and Michael Karagiorgas<sup>3</sup>

<sup>1</sup> Dept. of Electrical and Electronic Engineering, ASPETE, Athens, Greece

<sup>2</sup> Mechanical Engineer, Athens, Greece

<sup>3</sup> HVAC & RES laboratory of Dept. of Mechanical Engineering, ASPETE, Athens, Greece  
Email: [slivieratos@aspete.gr](mailto:slivieratos@aspete.gr)

**Abstract.** Evaluating the energy efficiency of a data center is very important to understand how energy-demanding it is and to investigate any potential energy saving measures. The new energy baseline model introduced in this article expresses the energy consumption as a function of the outdoor air temperature and the IT (information technology) equipment consumption, while at the same time it indicates how a data center performs. The proposed energy baseline model is validated against the actual data that have been collected from one of the data centers of the Greek Research and Technology Network (GRNET) located in Attica prefecture. Moreover, the carry out of a dynamic sliding analysis of the energy baseline in the form of sequential 12-month cycles brings to the fore potential changes in energy management operation and/or strategy as the time goes on.

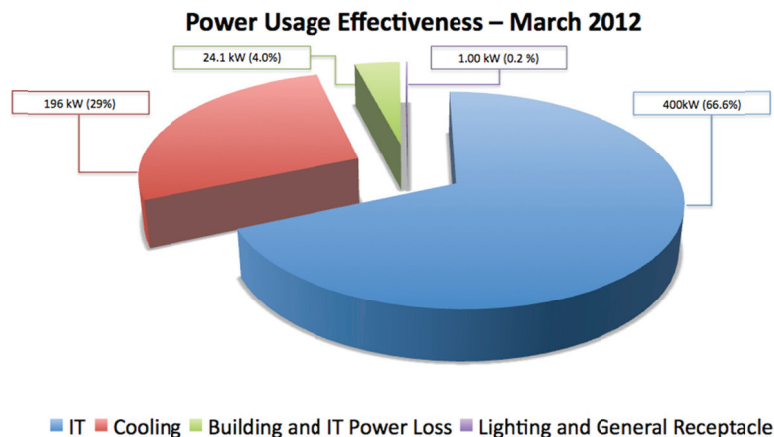
**Keywords:** Data center, energy efficiency, energy baseline, energy saving.

## 1 Introduction

The term “data center” means differently to different people. Some of the names used include data center, data hall, data farm, data warehouse, computer room, server room, R&D software lab, high-performance lab, hosting facility, colocation, and so on [1]. The U.S. Environment Protection Agency defines a data center as a “Primarily electronic equipment used for data processing (servers), data storage (storage equipment), and communications (network equipment). Collectively, this equipment processes, stores, and transmits digital information” or “Specialized power conversion and backup equipment to maintain reliable, high-quality power, as well as environmental control equipment to maintain the proper temperature and humidity for the ICT (information and communication technology) equipment.” [1]. In any case data centers are amongst the most complex and energy demanding building indoor environments, as a result of their high internal loads, low indoor temperature and humidity settings and continuous (uninterrupted) operation. Data centers are characterized by a very high concentration of IT equipment, peripherals and facility equipment, for example, power distribution units (PDUs), standby generators, uninterruptible power supplies (UPS), cooling systems (chillers, fans, pumps etc.), computer room air conditioning (CRAC) or computer room air handling (CRAH) units, cooling towers or chiller plants, artificial lighting and ancillary services [2]. The energy consumption of data centers has been increasing with the growth of demand for ICT services such as cloud computing and high performance computing and their associated concerns such as data security and privacy, which have also increased the need for designing and constructing sophisticated data centers. Further development of the data centers is expected when augmented reality applications come into the market. Apparently, data centers’ energy consumption is one of the highest concerns due to its significant environmental impact. In 2014, data centers consumed 2% of the total energy consumption of USA, a staggering 70 billion kWh [3], which was equal to the energy demand of 6.4 million average American households. The energy demand of data centers doubled during the last five years [4].

Based on the report entitled “Power Usage Effectiveness, March 2012” prepared by Lawrence Berkeley National Laboratory, 33.4% of the total energy consumed in a data center is used for powering and cooling and the 66.6% is the energy required by the IT equipment (Figure 1). For a typical server, the 30% of the power is consumed by its processor and the 70% by its peripheral equipment including the power supply, memory, fans, drive, and so on. A server’s utilization efficiency is estimated to be

around 20% [5], which is a quite disappointing level.



**Figure 1.** Average power consumption 400kW IT, 1kW Lighting, 24.1kW Building and IT Losses, 196kW Cooling

The rapid growth of data centers worldwide and the resultant carbon emissions, have spurred organizational bodies to recommend data centers be compliant with certain energy standards, such as ISO 50001 [6], as a measure of evaluating their energy efficiency and identifying energy saving opportunities. On the other hand, governmental authorities have set into force mandatory measures such as energy audits so that data center administrators are obliged to perform on a tactical basis (i.e. every 4 years) in order to validate the energy performance of their infrastructure. Both approaches require the building of the energy baseline of the data center to be the reference profile of its energy consumption, normalized with respect to the most critical factors that affect it, over an agreed reference period. Having built the energy baseline, the assessment of the energy performance of a data center is strongly facilitated as it makes the comparison of its energy consumption possible between different time-periods under normalized operational conditions.

In this context, researchers have paid much attention to the problems of (i) the modelling of the total energy consumed in a data center, (ii) the development of control methods to optimize its energy performance, (iii) the recovery of the wasted energy (heat) and (iv) the employment of renewable energy sources. Pelley et al. [7] have developed a total data center power model for detailed simulation according to which two external factors primarily affect the data center power usage, the aggregate IT load and the ambient outside air temperature. In simulation, each component's utilization, (processing demand for servers, current draw for electrical conditioning systems, thermal load for cooling systems, etc.), is related to its power draw. Hence, simulation requires precise accounting of per-component utilization and the connections between each component, resulting in a cumbersome and time-consuming analysis that represents the typical and not the actual energy use of the data center. Fang et al. [8] have developed a predictive control method for the optimization of the energy consumption of a data center maintaining at the same time the quality of service. This dynamic model employs time-consuming computational interactions and thermal relationship amongst the components of the data center and formulates a constrained non-linear optimal control problem to minimize the energy consumption of both IT and cooling equipment. Ebrahimi et al. [9] present various low-grade waste-heat recovery technologies that can be adopted by the data center industry as energy saving practices. Oro et al. [10] present available energy efficiency strategies and renewable energy integration into data centers and their characterization using numerical models. However, it is necessary to develop dynamic models and metrics in order to properly understand and quantify the energy consumption and the benefits of applying the incoming energy efficiency strategies and renewable energy sources to the operation of data centers. Rahmani et al. [3] have developed a modular simulation model for a data center with a certain number of servers. This model consists of a detailed power consumption modelling for each component and of interactions between all components. Its advantage, unlikely to other models in the literature, is that it can take different design structures of data centers into account and provide hourly power consumption profiles for each component. However, it remains quite complex for a quick analysis and it

does not capture the actual use as it builds various theoretical scenarios for typical uses.

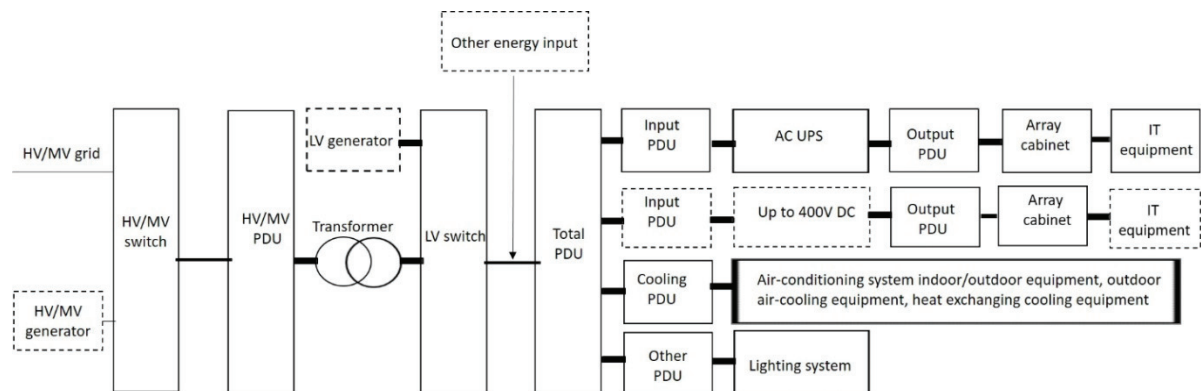
The present article focuses on the development of a simple, fast and effective model for the energy baseline of a data center representing its actual use and not its typical/theoretical one. On the one hand, the proposed model provides data center administrators with what an energy audit and/or ISO 50001 certification require, while, on the other hand, it brings into light energy management inefficiencies as well as energy saving opportunities. Moreover, it can be the benchmark against similar data center facilities in order to classify them in terms of energy performance. Subsequently, the proposed model prioritizes potential energy intervention projects under budgeting and cost/benefit constraints, starting from zero-cost interventions and going on higher levels of expenditure required by the facilities that need a retrofit. In addition, the significant facilitation of the energy forecasting process can be supportive to budgeting calculations.

In the proposed energy baseline model, the most important and statistically significant predictors (independent variables), that affect the energy consumption (dependent variable), are the outdoor air temperature and the IT equipment consumption as in [7]. By analyzing the relationship between energy use, IT processes and outdoor air temperature, a bivariate linear relationship has been adopted for predicting electricity consumption as a function of the IT electric consumption and the outdoor air temperature. The proposed model can be also used to predict future energy costs for budgeting, to establish baseline energy consumption, to measure energy savings and it can be applicable to many other relevant purposes [11]. It is also validated against actual data that show the real operation of the data center, which is affected by the actual energy strategy and management practices (time-plan, temperature and humidity set-points, etc.) and potential equipment malfunctions, and not its typical use. Moreover, it can identify data centers with the worst performance and quantifies how weather- and IT-normalized energy consumption changes over time by carrying out a sliding analysis over sequential 12-month cycles.

## 2 Theoretical Background of the Data Center Energy Baseline

### 2.1 Data Center Power Flow and Architecture

Figure 2 [12] below, depicts the power flow and the architecture of a data center:



**Figure 2.** Example of the data center energy consumption architecture

The input energy in Figure 2 comes from grid, generator or any other source of energy. The AC power distribution system consists of high voltage (HV) or medium voltage (MV) distribution, transformer, power distribution cabinet, distribution cable, cooling system power distribution, safety power distribution, lighting power distribution, array cabinet, IT equipment cabinet power distribution, etc. The UPS system includes UPS equipment and battery strings, etc. The up to 400V DC system is optional and usually adopted when telecommunication equipment is also collocated in a data center. The cooling system can consist of air conditioner (including indoor and outdoor equipment), cooling tower, pump, outdoor air equipment, heat exchanging equipment, etc. Other equipment can address

lighting, safety, anti-fire, sensors, monitoring, etc.

In the actual operation of a data center, electrical energy can be consumed by various subsystems other than the IT equipment, for example, transformers, UPS, distribution wiring, fans, air conditioners, pumps, humidifiers, and lighting. Some of these devices, such as UPS and transformers, are along the energy path that supplies the IT equipment. On the other hand, cooling and lighting provide auxiliary services, but do not supply energy to the IT equipment. The valuable computing and storage work executed in the data centers owns its successful and continuous performance to the electrical energy supplied to the IT equipment. All other types of energy consumption in a data center is wasted energy in the context of the present analysis as in reality it is just heat produced due to the inefficiency of the other subsystems of the energy path.

In addition, energy consumption varies depending on the operational mode of the IT equipment, which can be shutdown, sleep or active [13]. Hence, depending on the workload, the reduction of the energy consumption may be achieved by setting the appropriate power state of the ICT equipment.

Cooling equipment is also energy-hungry. It is composed of indoor and outdoor units [13]. This equipment controls the temperature and the amount of the refrigerants supplied from the indoor unit to the ICT equipment, in response to the heat generation due to the operation of the ICT equipment.

## 2.2 Data Center's Subsystems Energy Efficiency

The industry metric proposed for the evaluation of the efficiency of a data center has been PUE (Power Usage Effectiveness) until now. This specific metric highlights the amount of the aggregated power supplied to a data center, including the IT equipment and the corresponding aggregated power amount required by the facility infrastructure in order to only meet the IT needs. The algebraic PUE definition is:

$$PUE = \frac{\text{Total Facility Power}}{\text{IT Equipment Power}} \quad (1)$$

Manufacturers provide efficiency data for power and cooling equipment. For power equipment, the efficiency is typically expressed as percentage. Cooling equipment's efficiency is expressed in a number of ways, typically energy efficiency ratio (EER) for heat pumps and rooftop systems, which is the ratio of heat (kW) removed to electrical power (kW) consumed, and kW/ton for chillers. However, the published values of these metrics are calculated using a single point at standard temperature and humidity and therefore they provide indicative values under certain typical (not actual) conditions. The actual efficiency of the cooling equipment will vary depending on the weather conditions and the IT load. The cycling operational behavior of the various cooling subsystems result in time varying consumption even if the IT load remains the same and subsequently in time varying metrics. In addition, the efficiency will vary over time due to IT load changes, outdoor conditions and changes of the cooling mode. Annual PUE seems to be the intended metric, because it is easily related to the average energy consumption over the lifetime of a data center.

### 2.2.1 Data center's energy efficiency effect of the IT load

The IT load can vary from moment to moment due to instantly different computing needs, while the removal/addition/replacement of IT equipment can also cause permanent changes in the IT load. In any case, the energy efficiency is highly affected by the IT load. As the installed IT load increases, although the energy consumed increases as well, PUE tends to be lower. This means that if a well-designed data center experiences a higher than expected PUE, one possible explanation could be to check if the installed IT load is significantly lower than the nominal one. Normally, if the actually installed IT load capacity reaches 70% of its nominally planned capacity the performance of the data center remains approximately the same, all other factors being equal. On the contrary, no matter how much efficiently a data center is designed, it will experience a rather poor PUE if its installed IT load is low, e.g. 10-25% of its nominally planned capacity.

### 2.2.2 Data center's energy efficiency effect of outdoor conditions

Outdoor conditions vary with time and affect the energy efficiency of data centers. While various factors such as sunlight, humidity, and wind speed can affect data centers' energy efficiency, the most

important variable seems to be the outdoor air temperature. The energy efficiency of a data center declines as the outdoor air temperature increases because heat rejection systems consume more energy when processing the waste heat of a data center. Clearly, the efficiency varies with outdoor temperature, especially when a data center is equipped with economizer cooling modules.

### 2.2.3 Data center's energy efficiency effect of user configuration and settings

Data center operators can take various actions to affect PUE such as changing temperature or humidity set points, moving or adding vented floor tiles, failing to clean air filters, etc. These effects are highly variable and depend on the exact design of the power and cooling systems as well as the operation and maintenance plan of the facility. In the context of the present work, these operator-dependent settings are treated as attributes of the data center design and operation, and not as variable conditions such as IT load and outdoor temperature. That is the reason why these remain out of the scope of the theoretical analysis that follows.

## 2.3 Data Center's Energy Baseline New Proposed Mathematical Model

A mathematical model for the energy baseline that accurately represents the energy consumption of a specific data center, and has as inputs the IT load, outdoor temperature, etc., could be used effectively in a data center energy management program. Here is a list of some of the benefits provided by an effective energy baseline model that cannot be obtained by simply metering and recording data:

- Forecast the energy consumption of an operating data center.
- Estimate the efficiency of an operating data center for conditions that are not practical to measure, such as different IT loads and unusual outdoor temperatures, or industry benchmark conditions.
- Benchmark the performance of similar but different data centers under the same conditions via normalization.

The consumption of the air conditioning subsystem is very important for a data center as it removes the wasted heat and ensures appropriate environmental conditions for the IT equipment. The cooling load of a data center mainly comes from the IT equipment, envelope, lighting, people and infiltration, etc. [14]. Among these, the cooling load caused by people and infiltration compared to the amount generated by the IT equipment is insignificant and can be neglected. Therefore, the total cooling load of a data center is calculated as follows:

$$Q_{cooling} = UA_{envelope}(T_{out} - T_{in}) + Q_{IT} \quad (2)$$

where,

- $Q_{cooling}$  is the cooling load (W)
- $U$  is the thermal conductivity of the building ( $W/m^2/K$ )
- $A_{envelope}$  is the envelope of the building ( $m^2$ )
- $T_{out}$  is the outdoor air temperature ( $^{\circ}C$ )
- $T_{in}$  is the indoor air temperature ( $^{\circ}C$ )
- $Q_{IT}$  is the IT wasted thermal load (W)

Given above, it is qualitatively indicated that the outdoor temperature and the IT load are expected to mostly affect the energy consumed in a data center. The source data for the development of the proposed energy baseline model for a data center is: (i) total monthly electricity consumption, (ii) monthly IT electricity consumption, (iii) monthly AC consumption, and (iv) average monthly outdoor air temperature. Monthly bills provide information about the total electricity consumed. However, submeters have to be installed and commissioned for collecting IT and AC consumption data usually every e.g. 15min. The average monthly temperature can be measured by the data center administrators using a private meteorological station that could be installed near the facility, or provided by the closest public meteorological station.

The proposed energy baseline model is [15]:

$$E(MWh) = a_E * T(^{\circ}C) + b_E * IT(MWh) \quad (3)$$

where

- $E$  is the total monthly electricity consumption
- $T$  is the average monthly outdoor air temperature
- $IT$  is the monthly consumption of the IT equipment
- $\alpha_E$  is the temperature slope of  $E$  and expresses the additional total energy consumed if the temperature is increased by 1K
- $b_E$  is the IT slope of  $E$  and expresses the additional total energy consumed if the IT consumption is increased by 1MWh

A similar model to that expressed in (3) above can be adopted for the total electricity consumption of the air conditioning subsystem. In particular, the model becomes:

$$AC(MWh) = a_{AC} * T(^{\circ}C) + b_{AC} * IT(MWh) \quad (4)$$

where

- $AC$  is the total monthly electricity consumption of the entire air conditioning subsystem
- $T$  is the average monthly outdoor air temperature
- $IT$  is the monthly consumption of the IT equipment
- $\alpha_{AC}$  is the temperature slope of AC and expresses the additional AC energy consumed if the temperature is increased by 1K
- $b_{AC}$  is the IT slope of AC and expresses the additional AC energy consumed if the IT consumption is increased by 1MWh

Given the expressions (3) and (4) above, it is deduced that the parameters  $\alpha_E$  and  $\alpha_{AC}$  have to be equal because the increase in energy consumption - if the temperature is increased by 1K (all other factors being equal) accounts the same for both E and AC. Alternatively, the consumption due to the outdoor air temperature, addresses the AC and not the IT subsystem or any other electrical loads such as lighting, miscellaneous, etc. Hence,

$$\alpha_E = \alpha_{AC} \quad (5)$$

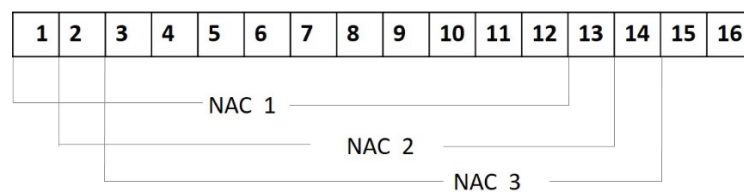
The parameters  $\alpha_E$ ,  $\alpha_{AC}$ ,  $b_E$  and  $b_{AC}$  are involved in a very complicated manner, which has not been algebraically identifiable so far, in characteristics of the energy performance of a data center such as the overall building envelope conductance, the ventilation or infiltration flow rate, the density of air, the external heating/cooling coefficient, the efficiency of the space heating/cooling equipment, the thermostat set-point temperature, the internal loads from electricity use, solar gain and people. Numerical changes of the above parameters of the energy baseline model indicate changes in the data center equipment, operation or energy management.

## 2.4 Normalization of Consumption of a Data Center and Dynamic 12-month Sliding Analysis

Utility bills show the actual monthly energy consumption of a data center. However, the energy consumption may be affected by unusual weather or computing processes. This makes it difficult to assess the energy performance of a data center over time when outdoor air temperature or IT equipment needs change. Both of these problems are easily eliminated by driving the proposed energy baseline model with “typical” temperature and IT input from historical records. The resulting annual energy consumption is called Normalized Annual Consumption (NAC). Thus, NAC represents the energy consumption of a data center after changes due to abnormal weather and IT variances have been removed. As such, NAC reveals the true energy characteristics of a data center, and allows the comparison of the energy consumption over time.

The change in energy characteristics of a data center can be determined by comparing its NAC during sequential 12-month cycles. This is called sliding NAC analysis [16]. To calculate the sliding NAC, an energy baseline is statistically created, via bivariate regression as previously mentioned, for each set of 12 sequential months, and then it is driven with typical temperature and IT data. In essence, using typical data means that for each month, e.g. for each April, the temperature and the IT needs remain the same. Hence, if the energy baseline is obtained by employing the actual IT consumption and temperature data of this particular 12-month cycle, and, if subsequently this baseline is fed with typical temperature and IT input data, it can lead to the typical monthly energy consumption estimation by

summing up over a 12-month cycle to NAC, which shows how much energy the data center would consume if the independent variables of T and IT took the typical values. The energy baseline model parameters, which are obtained with this sequential process, change over time too. From these changes, malfunctions and possible changes in the energy strategy may be identified. For example, the comparison of the actual 12-month consumption (AAC, Actual Annual Consumption) with the respective NAC can show when high or low energy use is caused by unusually hot or cold weather, rather than by fundamental changes in the data center energy system which is expressed by the model parameters. In this manner, a data center can be assessed over time as a standalone facility regarding how it performs compared to previous 12-month cycles under the same ambient temperature and IT needs. Moreover, two or more similar but different, data centers can be benchmarked as long as the percentages of their IT installed capacities with respect to their planned IT capacities respectively are almost the same. In such cases an additional normalization should take place taking into account the active IT density ( $\text{kWh}/\text{m}^2$ ). The dynamic 12-month sliding NAC analysis illustrates how the data center fundamental energy consumption characteristics change over time. Figure 3 shows a graphical representation of how a sliding NAC is calculated using the monthly sequential dataset.



**Figure 3.** Graphical representation of sequential 12-month cycles

As previously mentioned, the industry metric proposed so far for evaluating the efficiency of a data center, is PUE. In the present paper the idea of PUE, which involves electric power, is extended to electric energy. Hence, a new indicator is defined called EUE (Energy Usage Effectiveness), as follows:

$$EUE = \frac{\text{Total Facility Energy (per month or per annum)}}{\text{IT Equipment Energy (per month or per annum)}} \quad (6)$$

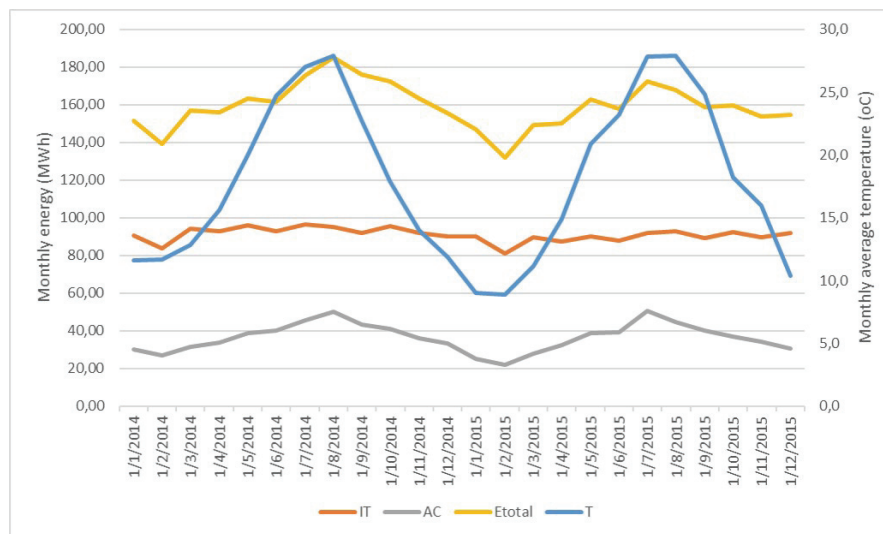
The variations of EUE can provide useful information regarding the energy performance of a data center as it is shown in the following section, which is referring to the numerical results. Moreover, a partially differentiation of expression (3) leads to:

$$b_E = \frac{\partial E}{\partial IT} \quad (7)$$

which, in other words, means that the marginal rate of the energy consumed by the facility with respect to IT energy needs is  $b_E$ . This is the total electricity that the facility consumes for every additional unit of energy consumed by the IT equipment. In a sense, one can say that  $b_E$  is the marginal EUE, while  $b_E < EUE$ . In a similar manner, one can say that  $b_{AC}$  is the AC energy consumed for every additional unit of energy consumed by the IT equipment. By the above, it is rather straightforward that the difference  $b_E - b_{AC}$  represents the additional energy consumed for every additional unit of IT energy, which is not related to pure AC and IT needs such as lighting, UPS losses, power distribution losses, etc. For simplicity purposes, this difference is called “other losses coefficient”.

### 3 Numerical Results

The proposed model is evaluated taking into account actual data from one of the data centers of GRNET which provides Internet connectivity, high-quality e-Infrastructures and advanced services to the Greek Educational, Academic and Research community. The data center is located in the northern part of the Attica prefecture and the period examined is from 1-1-2014 to 31-12-2015. Temperature data is available from the closest station of the meteo.gr network [17]. The IT consumption data is measured by using sub-meters installed to play this particular role. First, available data visualization is provided in Figure 4, where E, AC, IT and T are illustrated for the entire reference period.



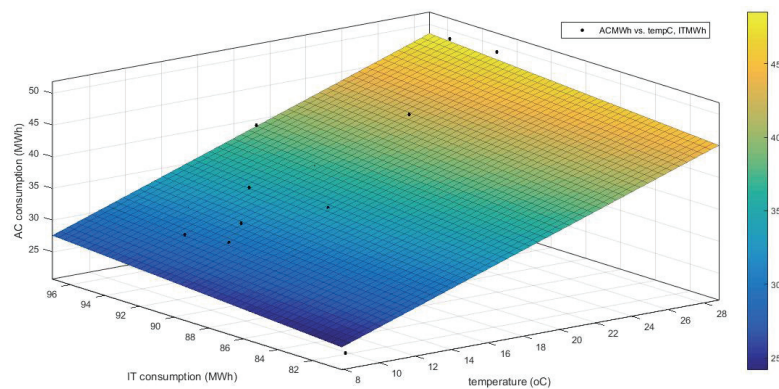
**Figure 4.** Monthly total energy E, AC energy, IT energy and average monthly outside air temperature T are illustrated.

It becomes obvious that the increase of E and AC occurring during the summers of 2014 and 2015 is due to the increase of T. On the other hand, the variations of IT consumption are very slight throughout the entire period (varying at around 90MWh per month). The parameters  $\alpha_{AC}$  and  $b_{AC}$  can be calculated via a bivariate statistical regression analysis, which is conducted using state-of-the-art mathematical software tools. Initially, AC is statistically correlated with T and IT as in expression (4) and the results are:

$$\alpha_{AC}=1.001 \text{ (MWh/K)}, \text{ obtained within 95\% confidence bound} \quad (8.1)$$

$$b_{AC}=0.2023, \text{ obtained within 95\% confidence bound} \quad (8.2)$$

while the goodness of fit of the regression analysis, expressed as  $R^2$ , is 90.82% and the RMSE (root mean square error) is 2.3MWh. The latter means that to this percentage the relationship of AC, T and IT is as proposed, while the complementary 9.18% depends on other unknown factors. The former, on the other hand, expresses a measure of the average error between the actual/observed and predicted values of monthly AC. A low  $R^2$  could indicate e.g. a bad control of the data center cooling system, while a high RMSE a large prediction error. A simple interpretation in the present case is that for every additional MWh/month of IT, 0.2023 MWh/month of cooling energy is required. The relevant graph follows (Figure 5):



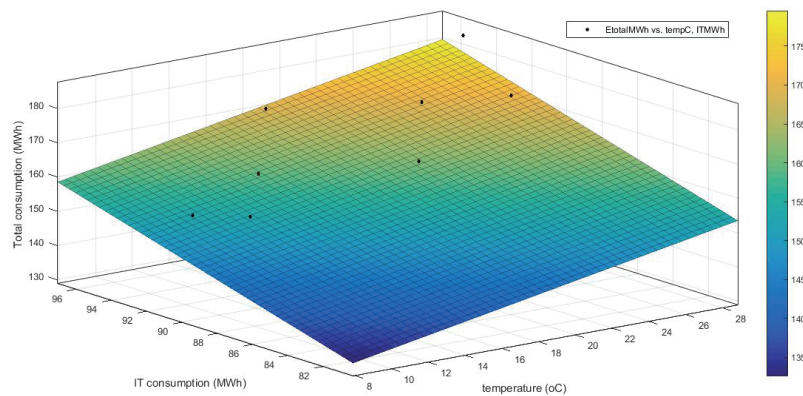
**Figure 5.** Monthly AC consumption as a function of average monthly temperature T and monthly IT consumption



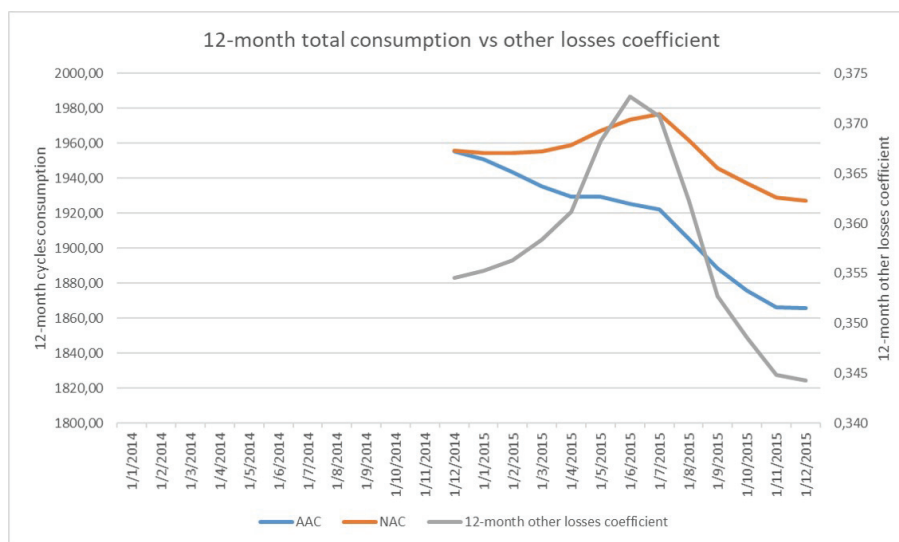
The total monthly energy  $E$  is statistically correlated with  $T$  and  $IT$  as in (3), taking into account the constraint expressed in (5). The result is:

$$b_E = 1.552, \text{ obtained within 95\% confidence bound} \tag{9}$$

while the goodness of fit of the regression analysis, expressed as  $R^2$ , is 85.61% and the RMSE is 4.6MWh. As above, the latter means that to this percentage the relationship of  $E$ ,  $T$  and  $IT$  is as proposed and the complementary 14.39% depends on other unknown factors. The former, on the other hand, expresses a measure of the average error between the actual/observed and predicted values of monthly  $E$ . A low  $R^2$  could indicate e.g. a bad control of the data center energy operation, while a high RMSE a large prediction error. A simple interpretation in the present case is that for every additional MWh/month of  $IT$ , 0.552 MWh/month of additional energy is required ( $IT$  consumption is subtracted). The relevant graph follows (Figure 6):



**Figure 6.** Total monthly energy consumption  $E$  as a function of average monthly temperature  $T$  and monthly  $IT$  consumption.



**Figure 7.** NAC and AAC vs other losses coefficient.

In Figure 7, which follows, the sliding NAC analysis of  $E$  is illustrated along with the sliding analysis of the difference  $b_E - b_{AC}$ , alternatively called “other losses coefficient” as previously mentioned. As shown, NAC increases until mid-summer 2015 while AAC decreases for the same period. This indicates that, although a lower actual consumption has been experienced, it is not due to an improvement of the energy strategy followed as NAC increases over the same period but it is due to an unexpected but

favorable influence of the other parameters IT and largely T. Figure 4 shows that the summer of 2014 was a bit hotter than the one of 2015 resulting, thus, in a slightly higher cooling and subsequently total consumption even though IT needs almost remain the same. The energy management deterioration can be captured by the fact that the “other losses coefficient” increases from January 2015 (0.355) to July 2015 (0.373) leading to this increase of the total energy consumed, while after July 2015 it gets progressively lower (0.344) than its initial value (0.355).

It is of high importance to examine the 12-month sliding profile of the proposed indicator EUE. However, it will first be explained how the typical 12-month EUE is calculated. From the definition of EUE it is suggested that

$$12\text{-month } EUE_{typical} = \frac{\text{Normalized Annual Consumption (NAC)}}{12\text{-month } IT_{typical} \text{ Consumption}} \tag{10}$$

Similar conclusions to the above mentioned can be obtained if one considers Figure 8, which represents the sliding analysis of 12-month EUE vs “other losses coefficient”. It is shown that until mid-summer 2015, both typical and actual 12-month EUE increase as a result of the increase of the “other losses coefficient”, while subsequently changes in the energy strategy lower the “other losses coefficient” and result in a continuous decrease of both typical and actual 12-month EUE till the end of 2015.

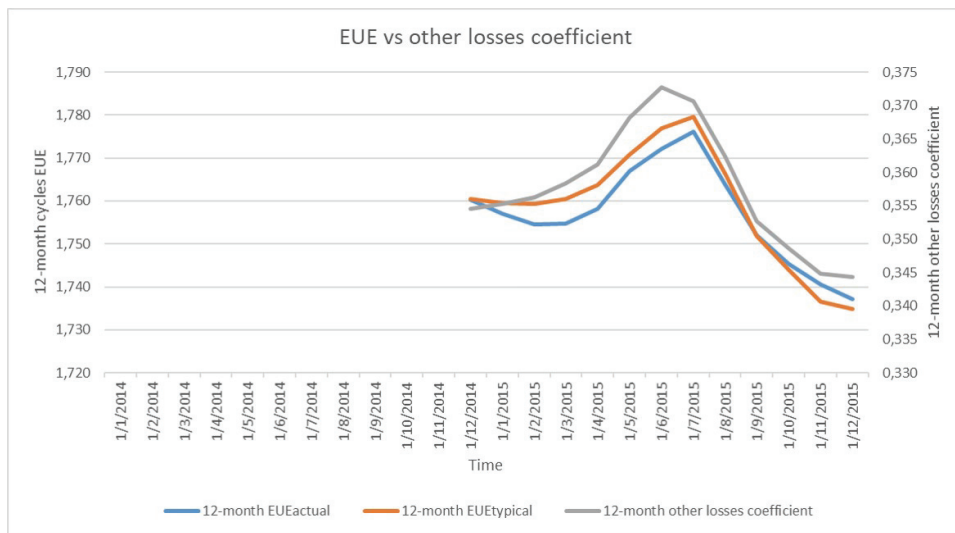


Figure 8. Sequential 12-month cycles of typical and actual EUE vs other losses coefficient.

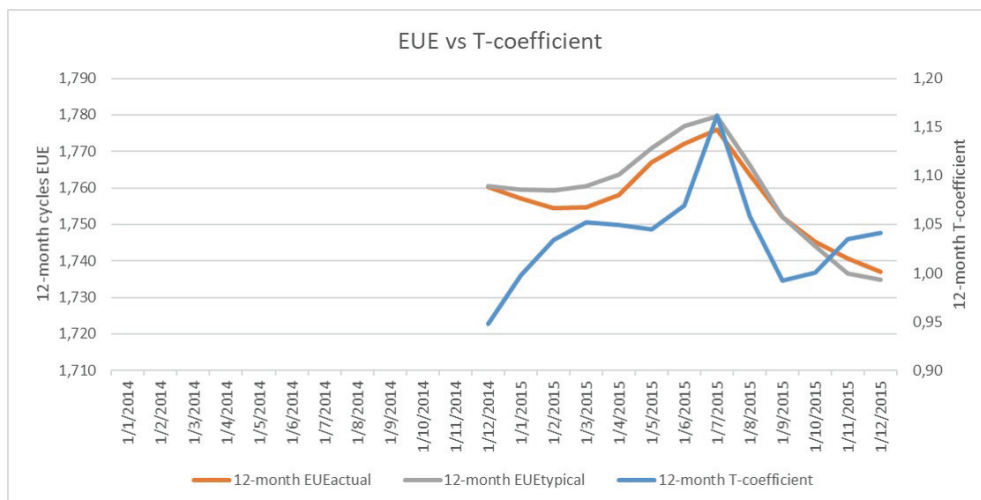


Figure 9. Sequential 12-month cycles of EUE vs T-coefficient.

In Figures 9 and 10 the sliding analysis of 12-month EUE vs  $a_E$  and  $b_E$  is illustrated respectively. It is worth noticing that an increase in EUE during both summers is largely due to an increase in  $a_E$  (T-coefficient), which reflects the fact that during both summers the performance of the cooling subsystem gets worse, while during the winters EUE is more affected by the variations of  $b_E$  (IT-coefficient).

The final Figure 10, for the case study of the GRNET data center, comprises an important conclusion, according to which the IT-coefficient, that reflects the energy use efficiency in the plant, is equal to 1.575 for the first year of measurements (dated 31-12-2014) while it is equal to 1.53 for the second year (dated 31-12-2015). This represents a drop of 2.86% due apparently to partial load conditions or to other operation parameter improvements that must be further analyzed.

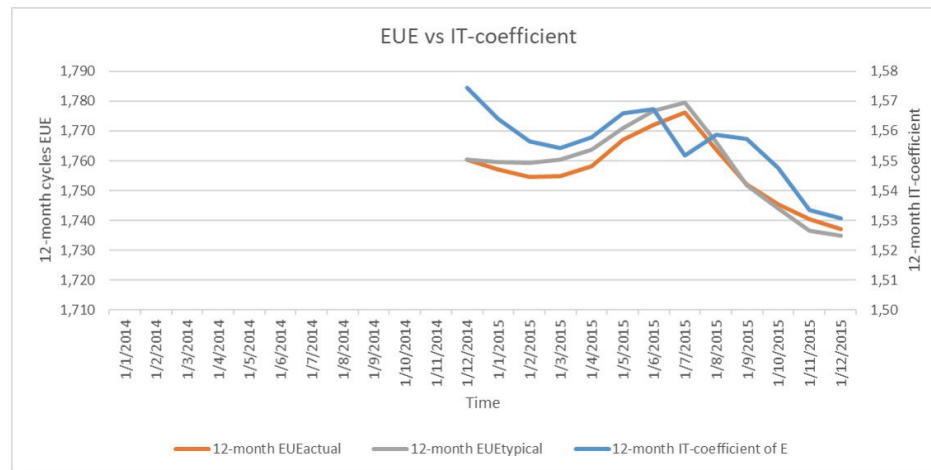


Figure 10. Sequential 12-month cycles EUE vs IT-coefficient.

## 4 Discussion and Analysis

The proposed energy baseline model has been successfully implemented in the case of the data center of GRNET and has been validated against actual temperature and energy consumption data. The regression accuracy achieved is acceptable and the energy performance of the facility has been captured in a relatively simple manner. It is of high importance that the proposed model is built taking into account the actual consumption and not the typical energy profiles. This is the differentiating point of the proposed method against other similar ones because it focuses on how the facility works and not on how it is planned to work. Moreover, it is a quick and cost-efficient method to apply and it can provide data center administrators with useful indications and recommendations for reviewing the energy management planning.

Initially, the two bivariate regressions of AC and E with respect to T and IT were conducted. From the regressions the numerical values of the parameters  $\alpha_E$ ,  $\alpha_{AC}$ ,  $b_E$  and  $b_{AC}$  were obtained. Having got the parameter values of the model, the two baselines of AC and E have been formed. In this point, one can say that the typical mandate of creating an energy baseline, which is foreseen within ISO 50001 energy standard or in an energy audit protocol, is fully met. In addition, the obtained models are used to forecast the cooling energy and/or the total energy required for the months to come and develop various budgeting scenarios regarding the energy spend. Such scenarios support alternative ways of negotiating energy contracts from energy service providers. This can be of high importance particularly for large data center organizations which have high energy needs and many facilities, and can lead to lower energy supply costs possibly by employing loyalty and/or volume discounts or even energy financial derivatives.

Further on, the proposed model supports a dynamic view of the energy performance which is based on the sequential 12-month sliding analysis. The dynamic analysis of the above parameters along sequential 12-month cycles provide a normalized and comparable overview of how a data center performs over time from the energy efficiency perspective. In other words,  $a_E$  and  $b_E$  describe the data center facility as time goes on and their variations can bring in light changes in the energy strategy adopted by the data

center administrator or malfunctions/aging of the facility equipment, both of which can indicate energy saving opportunities. As previously mentioned, the period examined is from 1-1-2014 to 31-12-2015. The actual data is collected from utility energy measurements and sub-measurements as well as temperature logs. Taking as reference the first 12-month cycle, from 1-1-2014 to 31-12-2014, the monthly values of  $T$  and  $IT$  of this cycle are used for the same months of other 12-month cycles, for which the baseline models have been calculated after regression as in (3) and (4). For unchanged energy management conditions, it is expected that NAC remains almost the same for sequential 12-month cycles. Likewise, slight differences are expected in the values that the parameters  $\alpha_E$ ,  $\alpha_{AC}$ ,  $b_E$  and  $b_{AC}$  take for the respective sequential 12-month cycles as a result of the statistical regression carried out. If the sliding analysis leads to variations of NAC, changes in energy management or in equipment performance are expected to happen. Increase of NAC indicates a deterioration of the energy management or equipment malfunction, while its decrease is associated with operational improvements.

Important conclusions on the efficiency of the energy management can be obtained from the comparison of the NAC with the AAC. If NAC and AAC go in parallel, this means that, even though the energy performance still depends on weather conditions and/or IT needs and might need improvement, the energy management does not change. If there are changes in the monotony of the relevant NAC and AAC time series, there are operational changes. It is most desirable to have NAC declining, which means that the energy consumed is getting less in result of a better energy strategy. In this case-study, the performance was getting worse by mid-summer 2015 but afterwards the trend was reversed and the performance was much better by the end of 2015.

The 12-month sliding analysis can also apply to other metrics, such as the EUE proposed in the present work, under similar assessment guidelines, e.g. it is desirable to have EUE typical and actual time series going in parallel and declining so that energy management gets more effective. If these target patterns are not met, then 12-month sliding analysis may indicate possible malfunctions and assist to identify energy saving opportunities. For example, from Figures 7 and 8 it is rather obvious that the increase in energy consumption and EUE is due to the increase of other losses such as lighting, power distribution, UPS, transformer, etc. which in turn might be investigated for potential failures or malfunctions.

It must be noted here that the results of sliding analysis can contribute in prioritizing facilities that need energy interventions. This is based on two fundamental facts: (i) which facilities have an increasing NAC and (ii) which facilities have a high NAC. These two categories have to be prioritized compared to the others and this classification is made based on normalized figures, e.g. same temperature and IT load so that data centers located in different areas with different local climatic conditions and different IT needs to be comparable.

## 5 Conclusions

This paper introduces and describes how a new method can be used to analyze monthly electric utility bills, temperature and IT requirements data to calculate the energy baseline of a data center and carry out a sliding analysis based on sequential 12-month cycles. The proposed method is fast and simple compared to time consuming simulation approaches, it is quite powerful, it accurately describes the energy performance of a data center and it captures changes in energy efficiency independently of weather and computation variations. It does not depend on the main source of electricity of the data center supplied by the main grid or by renewable energies through micro-grids. It facilitates energy budgeting calculations as well as the process of evaluating the energy savings that can be achieved from particular energy interventions as the proposed baseline normalizes the consumption under the same climate and IT load conditions. In addition, it easily quantifies the benchmarking among similar facilities, contributes to the estimation of the impact of the equipment aging (if executed periodically). What is more, this new method is of great assistance to both the internal and external energy auditors as they have a first view of the facilities' energy performance based on a relatively small dataset of measurements prior to their site survey and detailed audit, which is focused on particular fields in order to point out potential malfunctions or inefficiencies. The flexibility, scalability, comprehensiveness and modularity of this model provides the researchers and administrators with a powerful tool for energy analysis, management and planning of data centers with different designs and locations. The proposed

method is validated against actual data collected from a data center located in Attica, Greece owned by GRNET.

## References

1. G. Hwaiyu, *Data Center Handbook*, John Wiley & Sons, 2015.
2. C.A. Balaras, J. Lelekis, E.G. Dascalaki, D. Atsidaftis, "High Performance Data Centers and Energy Efficiency Potential in Greece", *International Conference on Sustainable Synergies from Buildings to the Urban Scale*, SBE16, 2017, pp.107-114.
3. R. Rahmani, I. Moser, M. Seyedmahmoudian, "A Complete Model for Modular Simulation of Data Centre Power Load", *Journal of IEEE Transaction on Automation Science and Engineering*, Vol. 14, No. 8, Aug. 2017, pp. 1-9.
4. D. Arnone, A. Barberi, D. La Cascia, E. R. Sanseverino, and G. Zizzo, "Smart grid integrated green data centres as ancillary service providers" in *Clean Electrical Power (ICCEP)*, 2015 International Conference on. IEEE, 2015, pp. 170-177.
5. Growth in data center electricity use 2005 to 2010, Available: <http://www.analyticspress.com/datacenters.html>
6. ISO 50001 standard, <https://www.iso.org/standard/69426.html>
7. S. Pelley, D. Meisner, T. F. Wenisch, J. W. VanGilder, "Understanding and Abstracting Total Data Center Power", in *Workshop on Energy Efficient Design*, 2009.
8. Q. Fang, J. Wang, Q. Gong, "QoS-Driven Power Management of Data Centers via Model Predictive Control", *IEEE Transactions on Automation Science and Engineering* 13(4), July 2016, pp. 1557-1566
9. K. Ebrahimi, G. F. Jones, and A. S. Fleischer, "A review of data center cooling technology, operating conditions and the corresponding low-grade waste heat recovery opportunities", *Renewable and Sustainable Energy Reviews*, Volume 31, March 2014, pp. 622-638
10. E. Oró, V. Depoorter, A. Garcia, J. Salom, "Energy efficiency and renewable energy integration in data centres. Strategies and modelling review", *Renewable and Sustainable Energy Reviews*, Volume 42, February 2015, pp. 429-445.
11. J.K. Kissock, C. Eger, "Measuring Industrial Energy Savings", *Applied Energy*, 85, pp. 347-361, 2008.
12. ITU-T, L.1302, *Assessment of energy efficiency on infrastructure in data centers and telecom centers*, 2015.
13. ITU-T, L.1300, *Supplement on rationale for minimum data set for evaluating energy efficiency and for controlling data center equipment in view of power saving*, 2014.
14. L. Ling, Q. Zhanga, L. Zeng, "Performance and energy efficiency analysis of data center cooling plant by using lake water source", *10<sup>th</sup> International Symposium on Heating, Ventilation and Air Conditioning*, ISHVAC2017, 19- 22 October 2017, Jinan, China.
15. S.N. Livieratos, S.L. Panetsos, A.K. Fotopoulos, M. Karagiorgas, "Energy Signature Model for Data Center Infrastructure", *ASHRAE 7<sup>th</sup> International Conference, "ENERGY in BUILDINGS"*, Athens, Greece, November 2018.
16. N. Lammers, K. Kissock, B. Abels, F. Seven, "Measuring Progress with Normalized Energy Intensity", *SAE Int. J. Mater. Manuf.* 4(1), pp. 460-467, 2011.
17. Meteorological data <http://www.meteo.gr/index-en.cfm>