

A Taguchi-Based Method for Assessing Data Center Sustainability

Sidney L. Pendelberry

*Computing and Information Sciences
Rochester Institute of Technology, Rochester, New York
sidney.pendelberry@rit.edu*

Sophie Ying Chen Su

*Golisano Institute for Sustainability
Rochester Institute of Technology, Rochester, New York
sophie.chensu@rit.edu*

Michael Thurston

*Golisano Institute for Sustainability
Rochester Institute of Technology, Rochester, New York
michael.thurston@rit.edu*

Abstract: This paper presents a methodology for measuring sustainability performance of data center facilities. Treating data center operation as a continuous production process, a Taguchi framework is used to estimate the sustainability impact due to power utilization from the facility by calculating the loss to society as the facilities power utilization becomes less efficient than the Environmental Protection Agency's guidelines. By tracking data center sustainability performance with continuous production process metrics, it is easy to identify and study performance characteristics, which if out of specification, have broader environmental impacts. This method evaluates the data center's excess energy consumption as a performance quality control issue, and estimates the cost of poor performance. Average power usage effectiveness (PUE) is a standard data center performance metric. In this case study PUE is tracked continuously to provide real-time operations and controls feedback. Further, PUE is mapped using the Taguchi framework and the U.S. EPA data center benchmark to the societal cost of non-conformance with the performance benchmark performance deviation based on carbon emissions and associated carbon exchange markets. It is expected that the method used in this paper may provide useful information to policy makers, engineers, and decision makers regarding sustainable IT data centers.

Keywords: Taguchi; data center; sustainability; PUE; carbon emission

1. INTRODUCTION

In 2010, the power consumption of United States data centers was projected to be 95.5 billion kWh (ENERGY, 2007). With data center power consumption at 1.5% of the U.S. total and growing rapidly, "Greener" data centers offer a large opportunity to reduce U.S. power consumption and thus reduce U.S. carbon emissions. This paper asserts that a data center is considered to be "Green" in the United States when the measure of sustainability, Power Usage Effectiveness (PUE) meets or exceeds the specification of the United States Environmental Protection Agency's (EPA) best practice benchmark (ENERGY, 2007). PUE is defined as the ratio of the total power consumed by a data center to the power

consumed by the IT equipment that populate the facility (Belady, Rawson, Pfleuger, & Cader, 2008):

$$PUE = \frac{TotalFacilityPower}{ITEquipmentPower} \quad (1)$$

In effect, the PUE represents the energy overhead associated with data center support systems. Total facility power is the sum of the IT equipment power demands and the support system power (security, environmental controls, lighting, etc.)

Treating sustainability as a continuous process of managing change, statistical process control techniques commonly used in manufacturing apply to the sustainability process to ensure that data center process variables such as PUE and carbon emissions stay in control (Di Mascio & Barton, 2001; Montgomery, 2009). A premise of this paper is that a data center facility will maximize its sustainability process as it approaches a PUE characteristic of 1.0, and further that a sustainability loss occurs when the PUE moves away from 1.0. A PUE of 2.0 means that for every watt of energy used to power IT equipment, an additional watt of energy is used to provide cooling, lighting, etc. While a PUE of 1.0 is ideal (given no internal data center renewable energy generation), this cannot be achieved but only asymptotically approached. We suggest a more practical PUE benchmark of 1.5 for the “product” specification (ENERGY, 2007; Google, 2010).

The Taguchi loss function was established to measure the impact of statistical process deviation from a target where a loss or cost function is calculated based on variance around the target value. On target processes incur the least overall loss (Pyzdek, 2003). The intent of this method is to support decision making regarding additional control and maintenance of a process. This paper applies the Taguchi approach (Di Mascio & Barton, 2001; Taguchi, Elsayed, & Hsiang, 1989) to gauge the sustainability process for data centers and the loss to society (as measured by a carbon emission metric) when the actual PUE deviates from the benchmark. The carbon emission related metric is used to quantify and unitize the loss function in sustainability terms, based on societal impact associated with green house gas emissions.

Using the EPA’s benchmark PUE as the ideal value for the performance characteristic (Greenberg, Mills, Tschudi, Rumsey, & Myatt, 2006), the closer to the actual value is to the ideal value, the “Greener” the data center. This is in essence Taguchi’s quality approach. The original purpose of the Taguchi-based measures was to quantify quality control in economic terms. It has been used in many applications, from biotechnology to software reliability.

Data center facility energy demands are a function of both the data center design and the operations and control strategies. While PUE is often used for evaluating data center design, the proposed Taguchi loss metric, calculated continuously, can be used a means of optimizing control in order to minimize overall sustainability impact while meeting service requirements. Once a data center has achieved the benchmark 1.5 average PUE, maintaining that level is a continuous process. The benchmark PUE value may also move (reduce) with technology advances in either data center facilities and IT equipment (Salim, 2009). Therefore, maintaining the same sustainability rating will require continual process or facility improvements. This state of change can be shown as a Deming Cycle (Deming, 2000) for the ongoing process of measuring and implementing data center efficiency. If at some point, the gap between best practice and current practice becomes too great, the process is considered to be out of control.

The PUE rating alone does not tell the whole story of data center energy efficiency. The PUE provides a measure of efficiency relative to the IT equipment consumption, but does not provide a measure of the total energy consumed by a data center or the associated carbon emissions. For example, two data centers of similar capacity and PUE, but with different IT equipment efficiency can have very different carbon emissions.

The next section, Methodology, describes in more detail the general framework for Taguchi’s quality assessment, and the suggested quality measure for data center sustainability. Following the Methodology section, Rochester Institute of Technology’s data center is used as a case study where the Taguchi framework is applied and the result shown. This paper concludes with a discussion of the results and some examples of common energy saving strategies that can aid in achieving the Taguchi product specification for a data center.

2. METHODOLOGY: TAGUCHI FRAMEWORK FOR SUSTAINABILITY ASSESSMENT

Taguchi’s loss function focuses on industrial production, and the need to produce a product within the tolerances as specified in the design. Taguchi argued that quality should start with the understanding of quality costs in various situations. He carried this one step further by considering the cost to society of manufacturing defects. Outside of the costs of rework, any part that is manufactured outside of its design specification and delivered to the customer would result in some loss to the customer through early wear-out; and problems throughout the system a chain effect due to the non-conformance of one part. These external losses, Taguchi argued, would eventually find their way back to the manufacturer. The general form of Taguchi’s loss function is shown in Equation 2 (Di Mascio & Barton, 2001; Taguchi et al., 1989).

$$L = k(y - m)^2 \quad (2)$$

In equation (2), L is the loss in terms of dollars. For sustainability loss, an example would be: if a given data center is not meeting the best practice benchmark in Table 1, the loss in dollars is a function of the extra CO₂ produced for the extra power consumed multiplied by the cost of carbon emission credits to offset the additional CO₂. Parameter m is the target value of the measured output. For the data center sustainability application, it is defined as the expected CO₂ emissions for the current data center if it achieved the “best practice” PUE. Finally, variable y is the currently measured (calculated) output value. For our data center sustainability formulation, it is defined as the current CO₂ emission based on the actual PUE.

TABLE 1. EPA PUE PROJECTION STANDARD BY 2011 (ENERGY, 2007)

PUE Benchmarks Scenario for Mid-Tier Datacenters	Year	
	2006	2011
Current practice	2	1.9
Improve operation	2	1.7
Best practice	2	1.5
State of art	2	1.5

The variable k is a scaling constant defined as the cost of the counter-measure or action taken to account for an individual part being outside of the specification, divided by the absolute value of the acceptable variation. This paper defines k for data center sustainability, as shown in Equation 3. It is the cost of the carbon credit divided by the absolute value of the difference between the CO₂ released if the data center were operated at the “State of the Art (SA)” PUE and the CO₂ released if the data center were operated at “Best Practice (BP)” PUE. The BP and SA reference values are taken from EPA recommendations and change over time. As technology improves, SA is adjusted to lower PUE values as design and implementation practices improve over time to utilize the new technology, the BP benchmark approaches the new SA. As noted earlier, the PUE benchmark asymptotically approaches an ideal value of 1.0, unless credit is given for local renewable power generation within the data center facility.

$$k_{dc} = \frac{CarbonCreditCost}{|SA(CO_2) - BP(CO_2)|} \quad (3)$$

Using the scaling coefficient defined above, our formulation of the Taguchi loss function for data centers is given in equation 4 below

$$L_{dc} = k_{dc} \cdot (u[Ac(CO_2(P_{dc}(t))) - BP(CO_2(P_{dc}(t)))]^2 \quad (4)$$

where P_{dc} represents the time varying power consumption of the data center, u is a unit (Heaviside) step function, $Act(CO_2)$ represents the current actual CO_2 emissions of the data center facility at the operating loads and efficiencies, and $BP(CO_2)$ is as defined above. The unit step function makes the Loss Function a one-sided loss function, so-as not to penalize performance exceeding the target value. The $Act(CO_2)$ parameter represents the Taguchi control variable (y) and the $BP(CO_2)$ represents the Taguchi target value (m).

2.1 PUE and Carbon Emissions

The data center loss function formulation assumes that Green House Gas (GHG) impact (as measured by Carbon Emission) is the key environmental control metric for data center operations. More comprehensive and sophisticated metrics could be derived, however there is a move internationally to regulate carbon emissions and active trading markets are developing, putting a real value (or cost) on these emissions. Equation 4 requires conversion of electrical power consumption to carbon emissions. While this conversion also varies by region and over time due to the mix of generation sources, this paper uses a conversion based upon the United States national average (ENERGY, 2007).

$$DC(CO_2)(\text{tonne}) = \left(P_{dc}(kWhr) \times 1329.35 \left(\frac{lb}{MWhr} \right) \times 1 \left(\frac{MW}{1000kW} \right) \times 1 \left(\frac{\text{tonne}}{2204lb} \right) \right) \quad (5)$$

2.2 PUE - US Standard

Using PUE as a benchmark standard is the first step for setting process control limits (Salim, 2009). The U.S. Environmental Protection Agency (EPA) published a PUE projection in 2006 for data centers in the U.S. (ENERGY, 2007) Depending on the different scenarios (e.g. server room / server closet, mid-tier, enterprise) the PUE is projected for mid-tier as shown in Table 1. The desired PUE number by 2011 is 1.5 for the best practice; this paper uses that benchmark as the control. The European Union is currently gathering PUE data of data centers from various agencies for monitoring individual data center performance (CENTRE, 2008).

2.3 Carbon Credits

Table 2. Carbon Credit Prices variation in the US (EcoBusinessLinks, 2009)

Carbon Offset Provider	Price (US\$/Metric ton CO2)
Verus Carbon Neutral	\$2.75*
Carbonfund.org	\$10.00
Terrapass	\$13.94
ClearSky Climate Solutions	\$15.00
Bonneville Environmental Foundation	\$29.00

*varies with the Chicago Carbon Exchange Price

Taguchi’s loss function is traditionally formulated such that the loss is represented in economic units. Equation 3 assumes that there is an economic cost (or value) that can be assigned to carbon emissions based on a market in Carbon Credits. Converting CO_2 emissions to currency values using carbon credit equivalents are subject to location, market pricing and energy usage. Table 2 illustrates the current variation in carbon credit prices. The large degree of variation observed in the prices is reflective of lack of standardization in the basic definition of the credit, as well as lack of regulation in the markets. As these markets become more regulated, there will still be variation in value over time, as well as between different economies as they try to meet their international commitments to reduce emissions. However, within a particular economy, at a point in time, the variation in pricing should be significantly less. For the purpose of this study, we selected the lowest cost option for achieving data center carbon neutrality (\$2.75/ton).

3. ROCHESTER INSTITUTE OF TECHNOLOGY DATA CENTER CASE STUDY

temperature map and also temperature transients in order to stay within equipment operating specifications. Figure 3 below provides an overlay of the data center temperature with the data center current draw, illustrating the variation in these parameters over a one-day period in February of 2010. The data shows good temperature control (~1 deg C total variation) and a variation in current draw (30 minute moving average) of about 25A, or 10%.

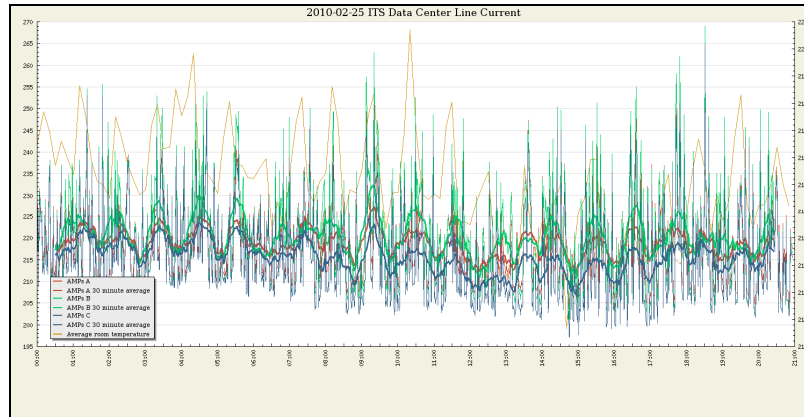


Figure 3. Power Statistics from Point 4 in Real Time

3.2 Results

An engineering study of the RIT data center was commissioned in 2009 (Consultants, 2009) focusing on HVAC and UPS capacity. The data collection systems described above were used for the engineering study. The associated reports provide UPS and generator capacity, total annual power usage, and also break out the HVAC associated power and other auxiliaries. The data summaries from the 2009 report were used for the data center analyses that follow.

Table 3 below shows an analysis of RIT data center power consumption for 2008 and 2009, broken out by sub-system, and also total energy use. The following three years (2010-2012) are projections based on historical growth trends; a 15% annual growth rate for server power and cooling demand is assumed. For purposes of this analysis, it is assumed that the existing UPS system can meet the growing system demand although it is at near to 100% capacity.

Table 3. RIT Data Center Annual Average Power Usage Projections

Item	Year				
	2008	2009	2010	2011	2012
Server power (kW) (15% annual growth)	70	81	92	106	122
Cooling (kW) (15% annual growth)	73	84	97	111	128
Lights (kW) 19 out of 33 lights are on 24 hours	2.8	2.8	2.8	2.8	2.8
UPS power loss (kW)	15	15	15	15	15
Total Power (MW-hr)	1410	1602	1813	2058	2348

Using Equation 2, Equation 3 and Table 3, Carbon footprint and PUE of RIT data center is calculated. Table 4 show the carbon emission projection for the RIT data center.

Table 4. RIT Data Center PUE, Carbon Emission

Item	Year
------	------

	2008	2009	2010	2011	2012
Average PUE	2.3	2.27	2.24	2.21	2.19
Annual CO ₂ (tonne)	851	965	1095	1245	1418

Table 5 shows the calculated Taguchi loss function for the RIT data center, using the EPA values and projections for state of the art and best practice, and using a carbon credit cost of \$2.75. Using the actual data center carbon emission for 2008 (851 ton) and the selected carbon cost would suggest that RIT should purchase \$2340 of carbon credits to be carbon neutral. The suggested Taguchi loss function formulation calculates a loss of \$6360 for 2008. This value is higher than the directly calculated value due to the small spread between SA and BP PUE. The variation of RIT's data center performance from BP is much higher than the variation between SA and BP. This provides an additional economic penalty for excessive deviation from best practice. Table 5 also gives the calculated Taguchi loss for the RIT data center through 2011, assuming no changes in the existing infrastructure.

Table 5. Taguchi Loss Calculation For RIT Data Center

Year	Loss (\$)	k \$/tonCO ₂ ²	y tonCO ₂	m tonCO ₂	SA PUE	BP PUE	Actual PUE
2008	6360	0.185	851	666	1.76	1.8	2.3
2009	6327	0.107	964	722	1.64	1.7	2.27
2010	6884	0.070	1095	782	1.52	1.6	2.24
2011	7811	0.0488	1245	845	1.4	1.5	2.21

3.3 Change Analysis

There are several ways to reduce data center power consumption and associated carbon emissions. Upgrading to higher efficiency IT equipment reduces total energy consumption but may actually result in an increase in PUE. Increasing the overall data center energy effectiveness (reducing non IT related power) results in PUE reductions and associated energy savings. Below are a few examples of changes that can make a positive impact on PUE (examples are relative to the 2008 baseline).

1. Reduction in lamp fixtures from 19 to 6 results in an improvement of approx 0.03 in PUE and a reduction of 10 tonne of CO₂ emissions per year.
2. If in addition to turning the reduction in operating fixtures to 6, the remaining fixtures are converted to LED lighting, PUE is reduced by approx 0.04 and CO₂ emissions by 15 tonne/year.
3. HVAC cooling reduction – The RIT data center uses Liebert Chillers that are set to utilize free air passive cooling when the outside ambient air temperature drops below 45 deg F. Rochester New York is in a temperate climate and can utilize free air-cooling for approximately 5 months of the year. The 45 deg passive cooling limit is set based upon the targeted data center interior control temperature of 65 deg F. If the data center control temperature is raised to 75 deg F, which would allow passive cooling for approximately 7.5 months of the year, resulting in an overall 25% reduction in cooling load. This improvement results in a PUE improvement of 0.26 and a CO₂ reduction of 96 tonnes of CO₂.

Increasing equipment operating temperatures raises a concern of reduced equipment reliability or durability. As an estimate of this effect, using the Arrhenius time to failure model indicates that for each 10 °C (18 °F) temperature rise component life is reduced by 50%(Feng & Hsu, 2004; Pinheiro, Weber, & Barroso, 2007). Experience at RIT indicates that the current physical life of IT equipment is longer than the useful life due to technology obsolescence. Additional work research will be needed to insure that increased operating temperatures don't result in actual reductions in data center reliability and availability.

4. CONCLUSIONS

This paper has described how sustainability metrics can be used for process control using a Taguchi loss function based method. In this case, sustainability performance

(impact on environment and society) is measured using the overall carbon emission. The standard form of the Taguchi loss function penalizes adverse process behavior and variation relative to a performance benchmark. We have suggested that the CO₂ emission associated with benchmark data center PUE is an appropriate benchmark.

The employed form of the Taguchi loss function penalizes data center performance deviation that significantly exceeds the variation between best practice and state of the art benchmarks, resulting in increasing cost associated with non-compliance. These results indicate that the RIT data center infrastructure is out of date and requires significant improvements, if not complete overhaul, in order to achieve best practice and zero the loss function. The loss function formulation also results in no economic penalty when the data center is at benchmark performance. The justification for this is that advancement beyond best practice is often not cost effective and there may be better (more cost effective) opportunities to reduce carbon emissions elsewhere.

5. REFERENCES

- Belady, C., Rawson, A., Pfleuger, D., & Cader, S. 2008. Green Grid data centre power efficiency metrics: PUE and DCIE. *Consortium Green Grid*.
- CENTRE, E. C. D.-G. J. J. R. 2008. Code of Conduct on Data Centres Energy Efficiency Version 1.0: EUROPEAN COMMISSION.
- Consultants, R. I. o. T. 2009. RIT Campus Data Center Power distribution Study.
- Deming, W. 2000. *The new economics: For industry, government, education*: The MIT Press.
- Di Mascio, R., & Barton, G. 2001. The economic assessment of process control quality using a Taguchi-based method. *Journal of Process Control*, 11(1): 81-88.
- EcoBusinessLinks. 2009. Ecobusinesslinks.com Carbon Offset Survey.
- ENERGY, S. 2007. Report to Congress on Server and Data Center Energy Efficiency Public Law 109-431. *Public Law*, 109: 431.
- Feng, W., & Hsu, C. 2004. *The origin and evolution of green destiny*.
- Google. 2010. Data Center Efficiency Measurements, *Efficient Computing*.
- Greenberg, S., Mills, E., Tschudi, B., Rumsey, P., & Myatt, B. 2006. Best practices for data centers: Lessons learned from benchmarking 22 data centers. *Proceedings of the ACEEE Summer Study on Energy Efficiency in Buildings in Asilomar, CA. ACEEE, August*, 3: 76-87.
- Montgomery, D. 2009. Statistical quality control: a modern introduction.
- Pinheiro, E., Weber, W., & Barroso, L. 2007. *Failure trends in a large disk drive population*.
- Pyzdek, T. 2003. *The Six Sigma handbook: a complete guide for greenbelts, blackbelts, and managers at all levels*: McGraw-Hill New York.
- Salim, M. 2009. Energy In Data centers: Benchmarking and Lessons Learned, *esmagazine*.
- Taguchi, G., Elsayed, E., & Hsiang, T. 1989. *Quality engineering in production systems*: McGraw-Hill College.