ETM

ENGINEERING & TECHNOLOGY MANAGEMENT

Technology Forecasting of Google Data Centers

Course: ETM 590 Engineering and Technology Management Synthesis

Instructor: Dr. Anderson

Term: Summer 2011

Team members: Chih-Hsun Lin, Tin Nguyen, Gaurav Sonny Katoch

Table of Contents

Abstract

The purpose of this research paper is to measure improvement rate of Power Usage Efficiency (PUE) with Information Technology (IT) equipment power and Total Facility Power (TP) in order to forecast the data center performance. We examined eleven Google Data Centers located in different geographical regions within the United States between 2005 and 2011. We used two mathematical approaches, Malmquist Index and the concept of Technology Forecast Data Envelope Analysis (TFDEA). The results were analyzed to determine the trend of PUE and used to forecast the PUE performance for the Google data centers for the next 10 years.

Introduction

Since cloud computing generated enormous amounts of discussions, more and more data center operators such as Google, IBM, Facebook, and Dell have been interested in building new data centers [1][2][3][4]. On the other hand, these data center operators are also excited about how to optimize their existent data centers' performance because it can drive the cost lower dramatically [5].

For most data center operators, one of the challenges for optimizing the data center performance is controlling power consumption. According to ICF International Report, data centers consume 1.5% of the total power in the U.S. Predicted growth over the next five to ten year is expected to require a similar increase in power generation [6]. Today, we know data center are big power consuming with high density. Garner Inc warns: "If they are not fully aware of the problem, data center managers run the risk of doubling their energy costs between 2005 and 2011. If we assume that data center energy costs continue to double every five years, they will have increased 1,600 percent between 2005 and 2025" [7].

In order to reduce power consumption, many data centers invested R&D resources in the direction of green technology. Organizations like Emerson, Google, and Hewlett Packard are just a few which have gone towards green technology.

Emerson invested $50 million in energy-efficient IT facility located in Missouri "…boasting more than 550 solar panels that can generate 100 kilowatts…" [8]. Tom Christian, Senior Research Scientist in HP's Sustainable IT Ecosystem Lab proposed the idea of cow-powered data centers. Why? Well cows produce methane and this can be captured to power electrical generators. Research presented by the American Society of Mechanical Engineers (ASME) International Conference on Energy Sustainability shows how a farm of 10,000 dairy cows could generate 1MW of electricity, enough to power a typical modern data center and still support other needs on the farm [9]. This is great opportunity for data center organizations to reduce costs and concurrently reduce their carbon footprint.

Our paper focus on two mathematical approaches such as Malmquist Index and the concept of TFDEA to evaluate performance of data center through the trend of PUE and use to forecast the PUE performance for the Google data centers for projection of next 10 years.

Background

Data Centers

According to Roger, one of contributor in datacenter website defined data center is a centralized repository, either physical or virtual, for the storage, management, and dissemination of data and information organized around a particular body of knowledge or pertaining to a particular business [10].

Data center are ravenous power consumers. Between 2000 and 2006, data center electricity consumption doubled in the United States. It is on pace to double again by 2011 to more than 100 billion kWh, equaling to $7.4 billion in annual electricity costs[11]. Data centers use nearly 10-30 times more energy per square foot than office space [12]. This vast consumption of energy will divert important resources (i.e. money) from being invested into other areas of the organizations.

A typical data centers spends about $1,000 a square foot. However, Google is spending nearly $3,000 a square foot on its new data center project in North Carolina, roughly three times the going rate for developing premium data center [13]. Why would Google spend money; when the cost of energy is increasing? This increase in cost per square foot would indicate that Google would be including green technology in order to increase its efficiency of its data center.

Literature Review

i. Green Data Centers

Defining Green:

Currently, the Data Center industry has no definition on green data centers; however, Douglas Alger author of the book *Grow a Greener Data Center* incorporates the definition of "green" buildings as a foundation. Douglas Alger has more than 20 years of professional experience with more than 12 years in Data Center physical design, Data Center operations, IT project management, construction project management, and IT infrastructure management.

A "green" building uses resources; such as, energy, water, and materials, more efficiently and has less impact upon people and the environment. When we replace building with data center we get "Computing environment that uses resources in a more efficient manner and has less impact upon people and the environment."

"Truly sustainable commercial buildings are extremely rare. Even achieving a Data Center that uses mostly sustainable is a major step forward from past server room designs and can provide significant benefits to your company." [14 Alger]

Reasons to go Green:

As organizations face growing demand on power and energy; leaning towards green data centers will be beneficial to the overall organization. "In an August 2007 report by the United States Environment Protection Agency or EPA estimates that U.S. Data Center power usage doubled in 6 years, consuming 61 billion kilowatt hours (kWh) of energy by 2006. Additionally, the report projects that unless Data Centers make efficiency improvement to both facilities and IT components, that power consumption will reach 100 billion kWh by 2011."[15 Alger]

When debating to go green the following four reasons are evaluated:

1. *Trade-offs in functionality and availability:* Does a green Data Center have more or less capacity, i.e., power, cooling, and connectivity, than other server environments? Are its physical infrastructure components more or less susceptible to downtime?

2. *Cost implications:* Is a green Data Center more or less expensive to build than a facility that doesn't bother with environmental considerations? Is it more or less expensive to operate? Is there enough return on investment that retrofitting an existing facility to be greener is worthwhile?

3. *Use of technologies uncommon to the Data Center industry:* What operational changes or new expertise does a green Data Center require?

4. *Ancillary issues:* Data Centers are not islands. They are a key piece of how a company functions. What issues outside of the hosting space are influenced by having a green Data Center?

The EPA's report to Congress on Server and Data Center Energy Efficiency displays the trend of U.S. Data Center power consumption between 2000 and 2011. The figure below illustrates the findings in the report with various energy-efficiency measurements:

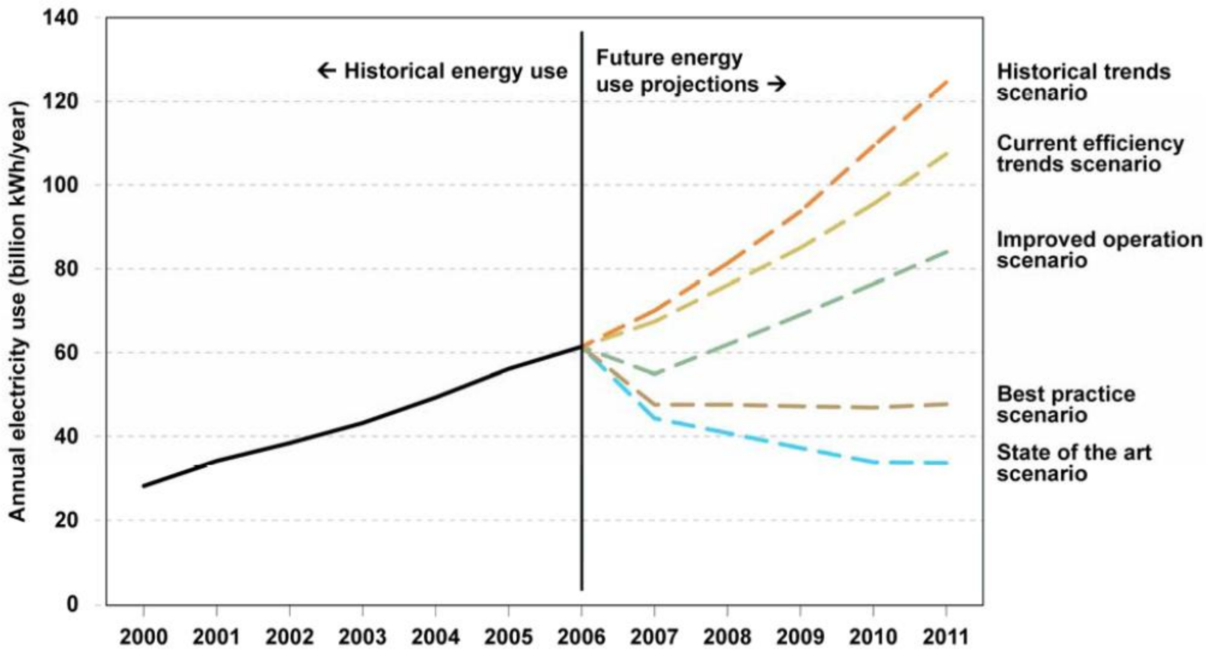


Fig. 1 Estimated U.S. Data Center Power Consumption Trends, With Green Measures

Green Incentives:

As more and more organization lean towards going green; several agencies now offer financial incentives to encourage the push for green. Government program and utility companies are just two examples of agencies which encourage the push for green.

Utility companies offer rebates and strategies for improving power efficiency. Why does an organization that is in the business of selling power be giving rebates to use less power? Ultimately, conserving energy benefits the utility companies by relieving demand on major power grids during peak times (typically during hot summer days and cold winter mornings) and lessening wholesale electric prices.

For example, “California-based Pacific Gas and Electric Co. (PG&E), offers rebates for organizations that consolidate older servers, paying 9 cents per kWh plus $100 per kW of demand reduction-up to 50 percent of the total cost of the project. The utility company offers similar rebates for upgrading disk storage equipment. PG&E founded the IT Energy Efficiency Coalition in 2007 to facilitate the creation of uniform incentive program, and more than 24 utilities from the United States and Canada now participate. If the programs are implemented as proposed a company with Data Centers in multiple locations can consolidate servers in each of them and count on receiving similar rebates."[16 Alger]

In addition, national and local governments offer tax break incentives to encourage a range of green adoption by the companies. These tax break incentive also play a role in which location projects; such as, Oregon's Google Dallas Data Center. Google had multiple locations to build their $600 million state-of-the-art data center, but eventually The Dallas, Oregon won the bid with better incentives.

"In the United States, federal tax credits for energy efficiency include a tax deduction of $1.80 per square foot for cutting a new or pre-existing commercial building's heating, cooling, ventilation, water heating, and interior lighting energy cost by 50 percent."[17 Alger]

ii. Power Usage Effectiveness or PUE:

Currently the most widely implemented metric in benchmarking data center efficiency. Developed in 2006 by The Green Grid (TGG); "a global consortium of organizations, government agencies, and educational institutions dedicated to advancing energy efficiency in data centers and business computing ecosystems" [18]. It is vital to measure energy consumption in data centers for four basic reasons:

1. Power is the most precious resource.

2. Power consumption is the most expensive operational cost.

3. Power is the common element among disparate data center subsystems.

4. Power consumption largely defines a data centers environmental impact.

An organization can calculated their PUE with a simple mathematical equation. The two variables are required to measure an organization's PUE are: Total Facility Power (TP) and IT Equipment Power (IT). Once these are obtained, and then simply divide the data centers Total Power by the IT equipment power. See equation below:

$$PUE = \frac{TP}{IT}$$

The typical real-world data centers' energy allocation is show as below. IT power normally only has less than 50% of the total facility power. Some powers are wasted in cooling, air movement, electricity transfer/UPS and Lighting [19].

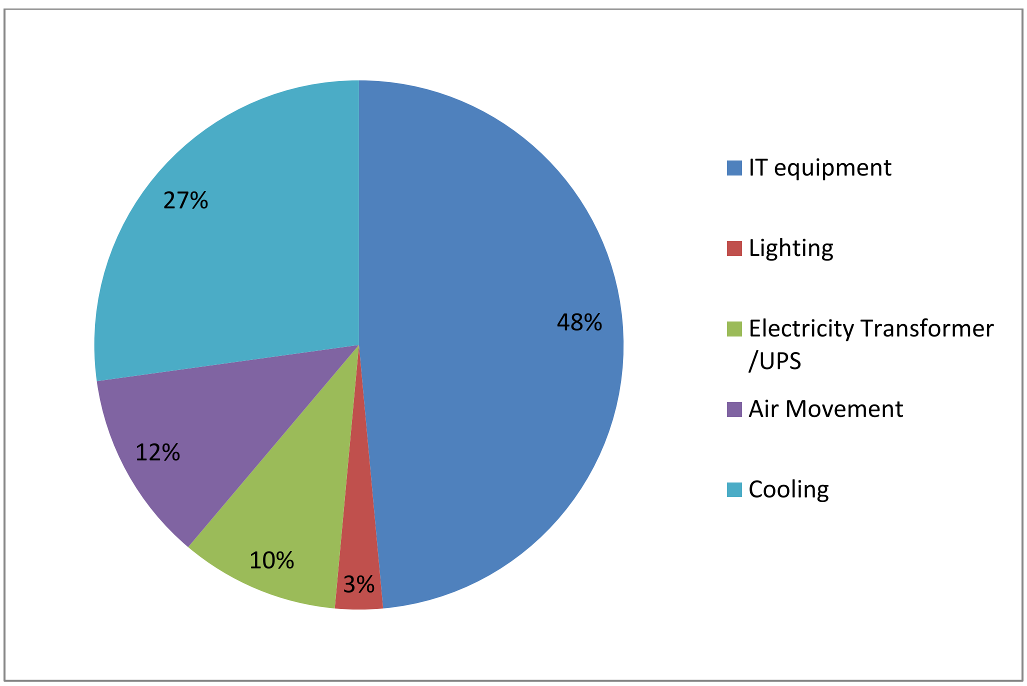


Fig. 2 Typical Real World Energy Allocation in Today's Data Center

Source: EYP Mission Critical Facilities Inc., New York

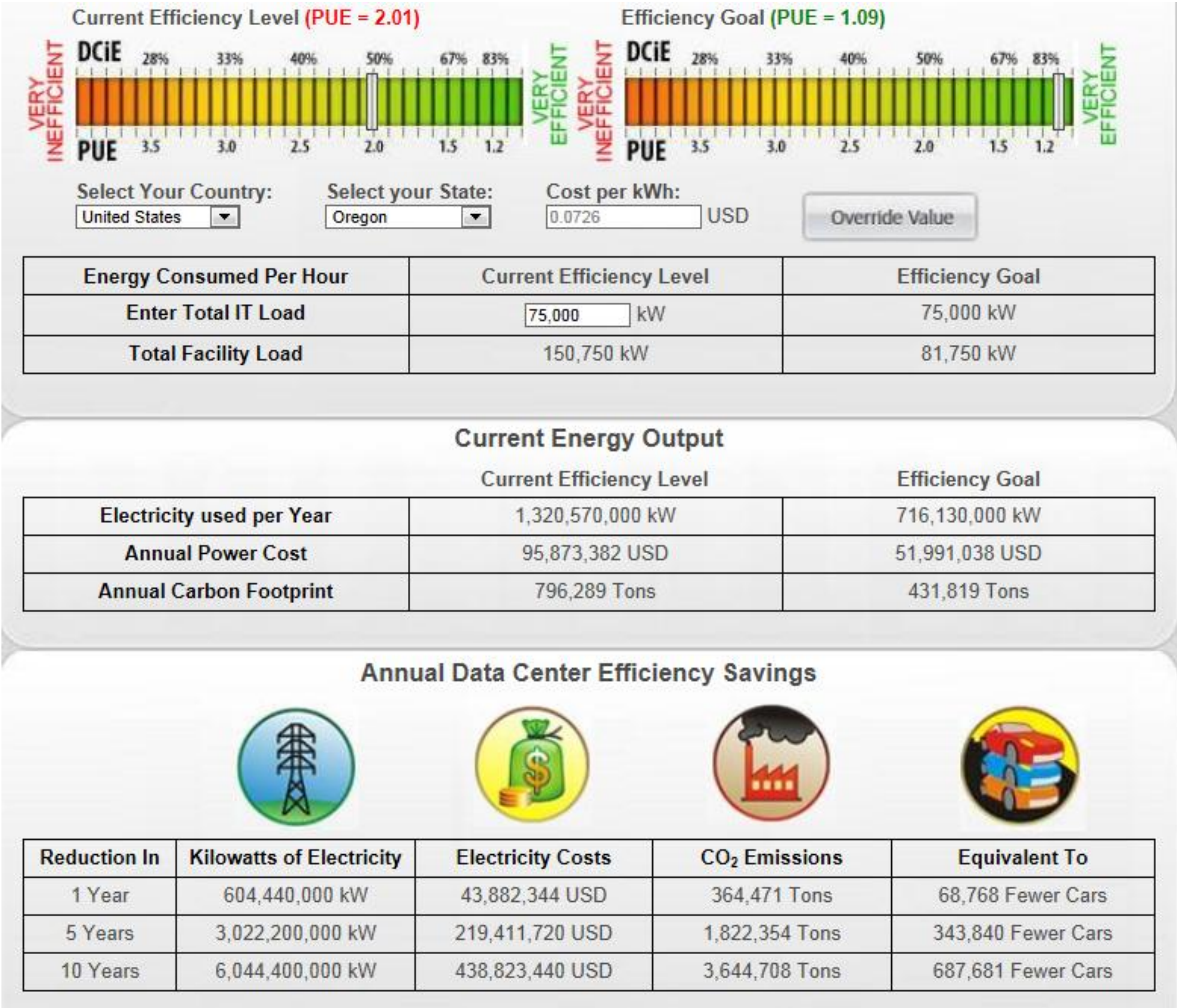
How does this aid an organization's data center become more energy efficient? The Green Grid (TGG) has also developed a table rating the calculated PUE and correlating the results to five levels of efficiency. Table 1 shows the detail of these five levels.

PUE	Level of Efficiency
3.0	Very Inefficient
2.5	Inefficient
2.0	Average
1.5	Efficient
1.2	Very Efficient

Table 1 the Relationship between PUE and Efficiency

Knowing where an organization stands in PUE rating is vital to help them save in operational cost. "As much as 50% of a data center's energy bill is from infrastructure (power & cooling equipment).

Furthermore, in 42U company (members of Green Grid) website, provides a transferring tool that helps us to see how decreasing of PUE value can reduce the electricity use, electricity cost, the quantity of carbon emission. Table 2 exhibits a result of reducing PUE value from 2.01 to 1.09 when IT power is 75,000 KW. According to the calculation, if this situation happens in one data center located in Oregon, it will save 604,440,000KW, 43,882,344 US dollars, and 364,471 tons of carbon emission (equivalent to 68,768 fewer cars) in one year.



Energy Consumed Per Hour	Current Efficiency Level	Efficiency Goal
Enter Total IT Load	75,000 kW	75,000 kW
Total Facility Load	150,750 kW	81,750 kW

Current Energy Output

	Current Efficiency Level	Efficiency Goal
Electricity used per Year	1,320,570,000 kW	716,130,000 kW
Annual Power Cost	95,873,382 USD	51,991,038 USD
Annual Carbon Footprint	796,289 Tons	431,819 Tons

Annual Data Center Efficiency Savings

Reduction In	Kilowatts of Electricity	Electricity Costs	CO_2 Emissions	Equivalent To
1 Year	604,440,000 kW	43,882,344 USD	364,471 Tons	68,768 Fewer Cars
5 Years	3,022,200,000 kW	219,411,720 USD	1,822,354 Tons	343,840 Fewer Cars
10 Years	6,044,400,000 kW	438,823,440 USD	3,644,708 Tons	687,681 Fewer Cars

Table 2 PUE Calculator Model

iii. Google Data Centers

Google is committed to using technology to build a clean energy future. They have worked hard to be carbon neutral as a company, launched multiple initiatives and have invested in several projects and companies going green.

According to Bill Weihl, Green Energy Czar, and Charles Baron, Google.org, Clean Energy Team below are factors that make Google a green leader:

1. ***Energy innovation pays off big:*** *Apply BAU or "business as usual" to scenarios with breakthroughs in clean energy technologies. On top of those, we layered a series of possible clean energy policies. We found that by 2030, when compared to BAU, breakthroughs could help the U.S.:*

- *Grow GDP by over $155 billion/year ($244 billion in our Clean Policy scenario)*
- *Create over 1.1 million new full-time jobs/year (1.9 million with Clean Policy)*
- *Reduce household energy costs by over $942/year ($995 with Clean Policy)*
- *Reduce U.S. oil consumption by over 1.1 billion barrels/year*
- *Reduce U.S. total carbon emissions by 13% in 2030 (21% with Clean Policy)*

2. ***Speed matters and delay is costly:*** *Our model found a mere five year delay (2010-2015) in accelerating technology innovation led to $2.3-3.2 trillion in unrealized GDP, an aggregate 1.2-1.4 million net unrealized jobs and 8-28 more giga-tons of potential Green House Gas or GHG emissions by 2050*[20].

Policy and innovation can enhance each other: Combining clean energy policies with technological breakthroughs increased the economic, security and pollution benefits for either innovation or policy alone. Take GHG emissions: the model showed that combining policy and innovation led to 59% GHG reductions by 2050 (vs. 2005 levels), while maintaining economic growth[21].

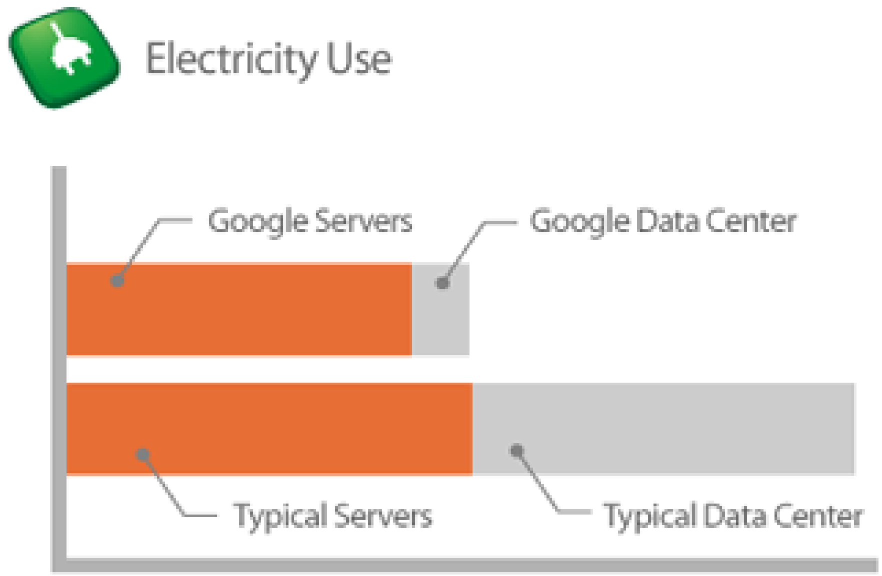


Fig. 3 The Power Consumption comparison between Google Data Centers and typical Data Centers

From figure 3 shows the differences of power efficiency between Google Data Centers and typical Data Centers. The gray bars in this figure indicate the total power consumptions and the orange bars indicate the power usage for servers. We can see that Google Data Centers have lower power consumption compared to the typical Data Centers.

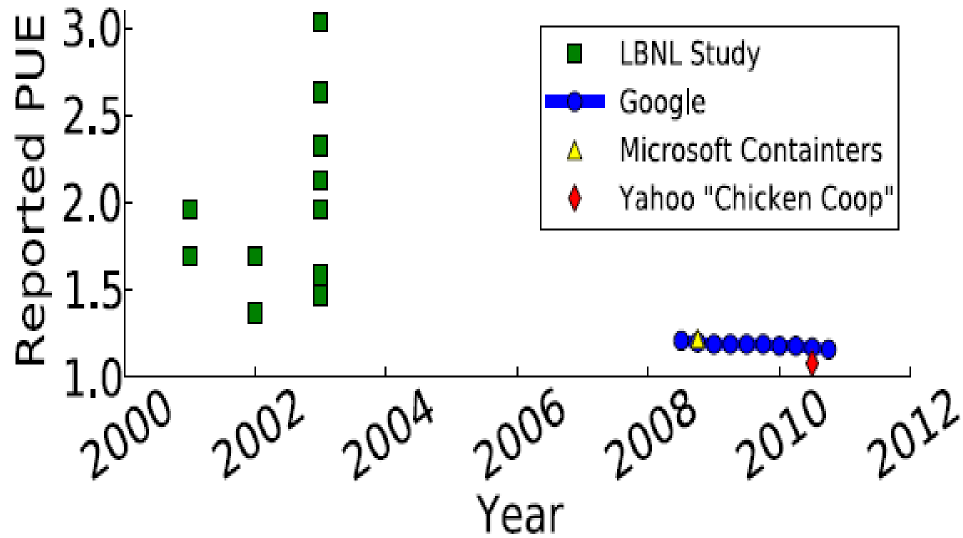


Fig. 4 The trend of PUE in variety of data centers [22]

From figure 4, we can see from 2008 to 2012, PUE for Google is 1.2 seem be efficient compare other data center test benchmarking with ineffectiveness of the efficiency.

iv. Technology Forecasting

The goal of technology forecasting is to prediction future technology capabilities, attributes, and parameters [23]. The major techniques for technological forecasting may be categorized under two general headings: methods based on numeric data and judgmental methods.

Technique	Advantages	Disadvantages
Monitoring	Unsophisticated First step to any good TF technique Low cost	Must review a great quantity of material Time-consuming Collection/summarization technique Not a "predictor"
Scenarios	Aids in understanding present Develops a plan of action for future	Dependent on select few (the writers) High cost Too general
Morphology	Goal-setting method Exhaustive Precise methods Breaks down whole into component parts	Extremely time-consuming Must know all alternatives Extremely high cost Impossible combinations must be recognized
Relevance trees	Goal-setting method Structures goal achievement	Unsophisticated Too general Bad project may not be easily seen
Delphi	First step to other TF techniques Many people can participate Eliminates personality conflict	Emphasis on consensus Time-consuming Does not relate final event to means to achieve final event
Cross-impact	Moderates some problems with Delphi Can be computerized Highlights lack of specific knowledge	Very laborious Requires Delphi analysis prior to use Time-consuming Must use same people as in Delphi study

Table 3 Advantages and Disadvantages of Judgment Technological Forecasting Techniques

From Table 3, we can see the advantages and disadvantage of using Judgmental technological forecasting technique based on monitor, scenarios, morphology, relevance trees, Delphi and cross-impact. For example the advantage of the Delphi Technique is involve with many people participate but the disadvantage of this method is time consuming.

In our report, we used numeric analyzed approach to forecast the improving rate of PUE performance in Google Data Centers. Two mathematic tools we chose are TFDEA and Malmquist index.

Technology Forecasting Data Envelopment Analysis (TFDEA) was created in 2001 as an alternative quantitative approach for technology forecasting. It is a powerful new technique for predicting complex technological trends and time to market for new product development.

TFDEA has been applied broadly in different industries. Inman, O. L, Anderson, T. R and Harmon, R. R used TFDEA to predict U.S. jet fighters between 1944 and 1982. The result shows that TFDEA is more accurate to predict the jet fighters' future designs than the traditional regression model [24]. TFDEA also was used in reviewing the most important forecasting in semiconductor industry, Moore's Law [25]. In this paper, Anderson, T, Fa¨re, R., Grosskopf, S., and Inman, L found that TFDEA not only could work well in the long term period but also could be a good instant performance measurement tool in fast-changing technology. In addition, Anderson, T. R, Daim, T. U, and Jisun Kim expanded the application of TFDEA in wireless communication technologies [26]. This paper provides us a clear framework how to identify input and output, analyze the result and forecast the state-of-art technology.

Malmquist index was introduced in 1982 was name after Sten Malmquist. It's call Malmquist productivity index. Malmquist productivity index is a bilateral index number enabling a productivity of comparison in data center j in different year. Malmquist index are quite popular method that use in field of production analysis.

Malmquist Index has been use in variety of field. Normally Malmquist productivity index (MPI) means measure the productivity changes over time. According to Journal of Construction Engineering and Management, the author mention the MPI is used to measure the productivity changes of Chinese construction industry from 1997 to 2003[27]. This tool provides good support of set up the policies and strategic decision by going through the improvement of Chinese construction industry performance through the difference region.

Methodology

i. Data collection

Table 4 summarizes the data about Google data centers. The cells colored in yellow are the data we collected from the websites and/or reports. The reason(s) behind inputting the same number(s) of TP in 2009 and 2011; first, is due to the fact that we only found the average numbers in these years, and second, Google announced its TP is similar in each of its data centers. We made two assumptions; the TP in 2008 was 80,000 MKW and the TP for 2011 was 103,000 MKW for all Google data centers.

		Input	Output	
Year	Site	TP (MKW)	IT (KW)	PUE (TP/IT)
2005	Oregon	45,457	22,729	2.00
2007	A	60,004	47,247	1.27
2007	B	60,004	51,728	1.16
2007	C	60,004	48,196	1.25
2007	D	60,004	49,590	1.21
2008	A	68,940	55,374	1.25
2008	B	68,940	57,811	1.19
2008	C	68,940	56,624	1.22
2008	D	68,940	57,331	1.20

2008	E	68,940	58,424	1.18
2008	F	68,940	56,508	1.22
2009	A	80,000	65,574	1.22
2009	B	80,000	70,796	1.13
2009	C	80,000	66,116	1.21
2009	D	80,000	67,227	1.19
2009	E	80,000	69,565	1.15
2009	F	80,000	69,565	1.15
2009	G	80,000	65,574	1.22
2009	H	80,000	62,992	1.27
2009	I	80,000	64,516	1.24
2010	A	91,003	75,521	1.21
2010	B	91,003	80,533	1.13
2010	C	91,003	75,209	1.21
2010	D	91,003	76,634	1.19
2010	E	91,003	79,305	1.15
2010	F	91,003	79,305	1.15
2010	G	91,003	78,790	1.16
2010	H	91,003	77,780	1.17
2010	I	91,003	76,634	1.19
2010	J	91,003	79,478	1.15
2011	A	103,000	88,034	1.17
2011	B	103,000	93,636	1.10
2011	C	103,000	84,426	1.22
2011	D	103,000	86,555	1.19

2011	E	103,000	94,495	1.09
2011	F	103,000	94,495	1.09
2011	G	103,000	91,964	1.12
2011	H	103,000	92,793	1.11
2011	I	103,000	90,351	1.14
2011	J	103,000	93,636	1.10

Table 4 Data used for two mathematical approaches

As we mentioned before, most organizations keep their power consumption information a secret, it is hard to get accurate and sufficient data. Due to the difficulty in obtaining the total facility power in 2007, 2008, and 2010, we made another assumptions on the unknown parts based on the calculations of the improving rate in TP. The formula below details the steps and assumptions which were made to get our results.

Total Facility Power (TP) in year j = (TP in year k) × (1+i_x*)^tp----------------------------------(1)*

i_x: the improving rate in time period, x=1,2,…..n

tp: the time period between year j to year k

From formula (1), we can get to formula 2 and 3 based on our collected TP data in 2005, 2009 and 2011.

45,457(TP in 2005) = 80,000(TP in 2009) × (1+i_1) ^ (2009-2005) ----------------------------- (2)

45,457(TP in 2005) = 103,000(TP in 2009) × (1+i_2) ^ (2011-2005) --------------------------- (3)

From the results of the calculations, we got i_1 around 0.1518 and i_2 around 0.1461. It is interesting that the improving rate of TP is similar for the time period. The average of i_1

and i_2 is 0.1489; this was used to calculate the improving rate index of TP and the unknown numbers of TP in 2007, 2008 and 2010. These numbers are colored in blue on our table.

Finally, based on the formula, IT equals TP divided by PUE; thereby, achieving all of the numbers of IT for each year. We colored these cells in prink, as shown in Table 4.

ii. The concept of TFDEA

Based on the data we collected in table 1, we applied the concept of TFDEA to measure the improving rate of PUE in these years. Table 2 displays the calculated results of TFDEA. In table 5, Φ_R represents the ratio of the PUE in data center i in year j compared to the lowest PUE in the released year k. Therefore, the cells under Φ_R column with the score 1.000 are considered as the data centers that have obtained the best PUE performance before the released year. On the other hand, Φc represents the ratio of the PUE in data center i in year j compared to the lowest PUE in the current year, 2011. For example, the number of Φc of Data Center A in 2009, 1.1193, was calculated by its PUE (1.22) divided by the lowest PUE number (1.09). Finally, we used the numbers of Φ_R and Φc to calculate the Gamma or the improving rate. To calculate Gamma we used the formula below:

$$\text{Gamma} = \Phi c \ \hat{}\ (1/\ (2011\text{- year j in data center i}),\ \forall\ \Phi_{R} \leq 1.0000$$

Accordingly, the calculation results of improving rates (cells in green color) are listed below:

- 1.1065 in 2005.
- 1.0157 in 2007.
- 1.0200 in 2008.

- 1.0182 in 2009.
- 1.0367 in 2010.

Year	Site	Input: TP (MKW)	Output: IT (KW)	PUE (TP/IT)	TFDEA Concept: Φ_R	Φ_c	Gamma
2005	Oregon	45,457	22,729	2.00	1.0000	1.8349	1.1065
2007	A	60,004	47,247	1.27	1.0948	1.1651	
2007	B	60,004	51,728	1.16	1.0000	1.0642	1.0157
2007	C	60,004	48,196	1.25	1.0733	1.1422	
2007	D	60,004	49,590	1.21	1.0431	1.1101	
2008	A	68,940	55,374	1.25	1.0733	1.1422	
2008	B	68,940	57,811	1.19	1.0280	1.0940	
2008	C	68,940	56,624	1.22	1.0496	1.1170	
2008	D	68,940	57,331	1.20	1.0366	1.1032	
2008	E	68,940	58,424	1.18	1.0000	1.0826	1.0200
2008	F	68,940	56,508	1.22	1.0517	1.1193	
2009	A	80,000	65,574	1.22	1.0796	1.1193	
2009	B	80,000	70,796	1.13	1.0000	1.0367	1.0182
2009	C	80,000	66,116	1.21	1.0708	1.1101	
2009	D	80,000	67,227	1.19	1.0531	1.0917	
2009	E	80,000	69,565	1.15	1.0177	1.0550	
2009	F	80,000	69,565	1.15	1.0177	1.0550	
2009	G	80,000	65,574	1.22	1.0796	1.1193	
2009	H	80,000	62,992	1.27	1.1239	1.1651	

2009	I	80,000	64,516	1.24	1.0973	1.1376	
2010	A	91,003	75,521	1.21	1.0664	1.1055	
2010	B	91,003	80,533	1.13	1.0000	1.0367	1.0367
2010	C	91,003	75,209	1.21	1.0708	1.1101	
2010	D	91,003	76,634	1.19	1.0509	1.0894	
2010	E	91,003	79,305	1.15	1.0155	1.0528	
2010	F	91,003	79,305	1.15	1.0155	1.0528	
2010	G	91,003	78,790	1.16	1.0221	1.0596	
2010	H	91,003	77,780	1.17	1.0354	1.0734	
2010	I	91,003	76,634	1.19	1.0509	1.0894	
2010	J	91,003	79,478	1.15	1.0133	1.0505	
2011	A	103,000	88,034	1.17	1.0734	1.0734	
2011	B	103,000	93,636	1.10	1.0092	1.0092	
2011	C	103,000	84,426	1.22	1.1193	1.1193	
2011	D	103,000	86,555	1.19	1.0917	1.0917	
2011	E	103,000	94,495	1.09	1.0000	1.0000	
2011	F	103,000	94,495	1.09	1.0000	1.0000	
2011	G	103,000	91,964	1.12	1.0275	1.0275	
2011	H	103,000	92,793	1.11	1.0183	1.0183	
2011	I	103,000	90,351	1.14	1.0459	1.0459	
2011	J	103,000	93,636	1.10	1.0092	1.0092	

Table 5 the calculation results from TFDEA concept

iii. Malmquist Index

We used the Malmquist Index to see productivity of rate of change over time. From table 4, we use j to represent data center site located in different region within US. For example, in 2007 j1 would equal to a PUE of 1.27 due to the fact that it was the lowest PUE rating for that year. Next, we would identify the PUE value corresponding to j2 (similar to finding j1, we locate the lowest PUE value for 2011). Once the corresponding values are identified we divide the numerator j1 by the denominator j2; the result of which is raised to the power of (1/tp). Using this approach we can compare the rate of change between a set time periods and the same data center location. We repeated the steps above to develop a clear understanding of the rate of change.

$$\textbf{Malmquist Index}= \left(\frac{Q_{j1}}{Q_{j2}}\right)^{\frac{1}{tp}} \quad \textbf{(Equation 2)}$$

The symbol below are meaning of equation

- tp : Number of Years
- Q_{j1} : PUE of 1.27 in 2007
- Q_{j2} : PUE of 1.17 in 2011

Year	Site	Input	Output		Malmquist Index	
		TP (MKW)	IT (KW)	PUE (TP/IT)	Change	Malmquist
2005	Oregon	45,457	22,729	2.00		

2007	A	60,004	47,247	1.27	1.0855	1.0207
2007	B	60,004	51,728	1.16	1.0545	1.0134
2007	C	60,004	48,196	1.25	1.0205	1.0051
2007	D	60,004	49,590	1.21	1.0168	1.0042
2008	A	68,940	55,374	1.25	1.0641	1.0209
2008	B	68,940	57,811	1.19	1.0841	1.0273
2008	C	68,940	56,624	1.22	0.9980	0.9993
2008	D	68,940	57,331	1.20	1.0105	1.0035
2008	E	68,940	58,424	1.18	1.0826	1.0268
2008	F	68,940	56,508	1.22	1.1193	1.0383
2009	A	80,000	65,574	1.22	1.0427	1.0211
2009	B	80,000	70,796	1.13	1.0273	1.0135
2009	C	80,000	66,116	1.21	0.9918	0.9959
2009	D	80,000	67,227	1.19	1.0000	1.0000
2009	E	80,000	69,565	1.15	1.0550	1.0272
2009	F	80,000	69,565	1.15	1.0550	1.0272
2009	G	80,000	65,574	1.22	1.0893	1.0437
2009	H	80,000	62,992	1.27	1.1441	1.0696
2009	I	80,000	64,516	1.24	1.0877	1.0429
2010	A	91,003	75,521	1.21	1.0299	
2010	B	91,003	80,533	1.13	1.0273	
2010	C	91,003	75,209	1.21	0.9918	
2010	D	91,003	76,634	1.19	0.9979	
2010	E	91,003	79,305	1.15	1.0528	
2010	F	91,003	79,305	1.15	1.0528	

2010	G	91,003	78,790	1.16	1.0313
2010	H	91,003	77,780	1.17	1.0541
2010	I	91,003	76,634	1.19	1.0417
2010	J	91,003	79,478	1.15	1.0409
2011	A	103,000	88,034	1.17	
2011	B	103,000	93,636	1.10	
2011	C	103,000	84,426	1.22	
2011	D	103,000	86,555	1.19	
2011	E	103,000	94,495	1.09	
2011	F	103,000	94,495	1.09	
2011	G	103,000	91,964	1.12	
2011	H	103,000	92,793	1.11	
2011	I	103,000	90,351	1.14	
2011	J	103,000	93,636	1.10	

Table 6 the calculation results from Malmquist Index

Data Results

Figure 3 shows the trend of PUE in Google Data Centers from 2005 to 2011 based on TP versus IT. The fine lines on the figure indicate the state of art performance of PUE in Google Data Centers in different years, and the bold green line indicates the lowest PUE value which is equivalent to 1. After tracing the slope of these lines, we go two implications. First, both TP and IT increased from 2005 to 2011 in all data centers. Second, all of these Google Data Centers are closer and closer to their terminal goal, the lowest PUE (1.0).

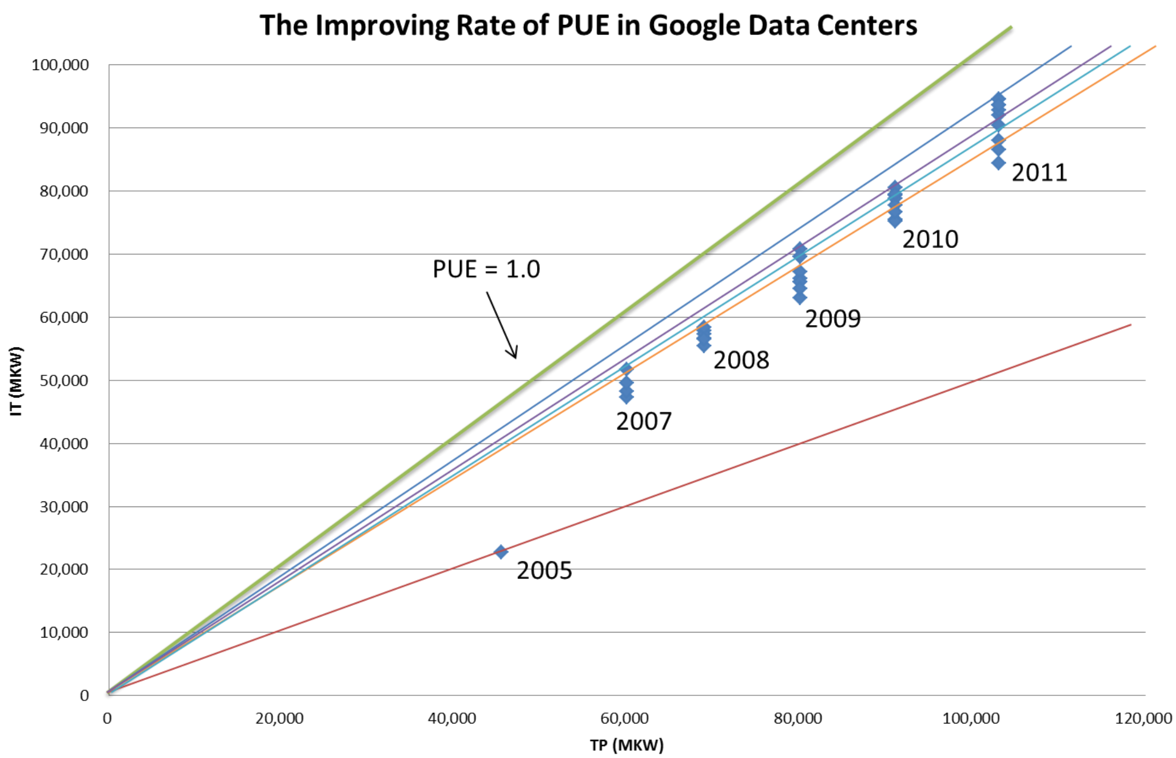


Fig 5 the trend of PUE from 2005 to 2011

Further, when we plot the curve based on PUE versus year from 2005 to 2011, we can see clearly that the improving rate in Google Data Centers increased quickly from 2005 to 2007 but it slowed down after 2007. In order to predict when PUE value will be around 1.0, we used the average numbers of improving rate. In TFDEA approach, the average improving rate (Gamma) is around 1.0226. Following the TFEA concept, we predicted PUE will be closed to 1.0 around 2014 to 2015. On the other hand, in Malmquist Index approach, the result shows that PUE will be closed to 1.0 around 2015 to 2016.

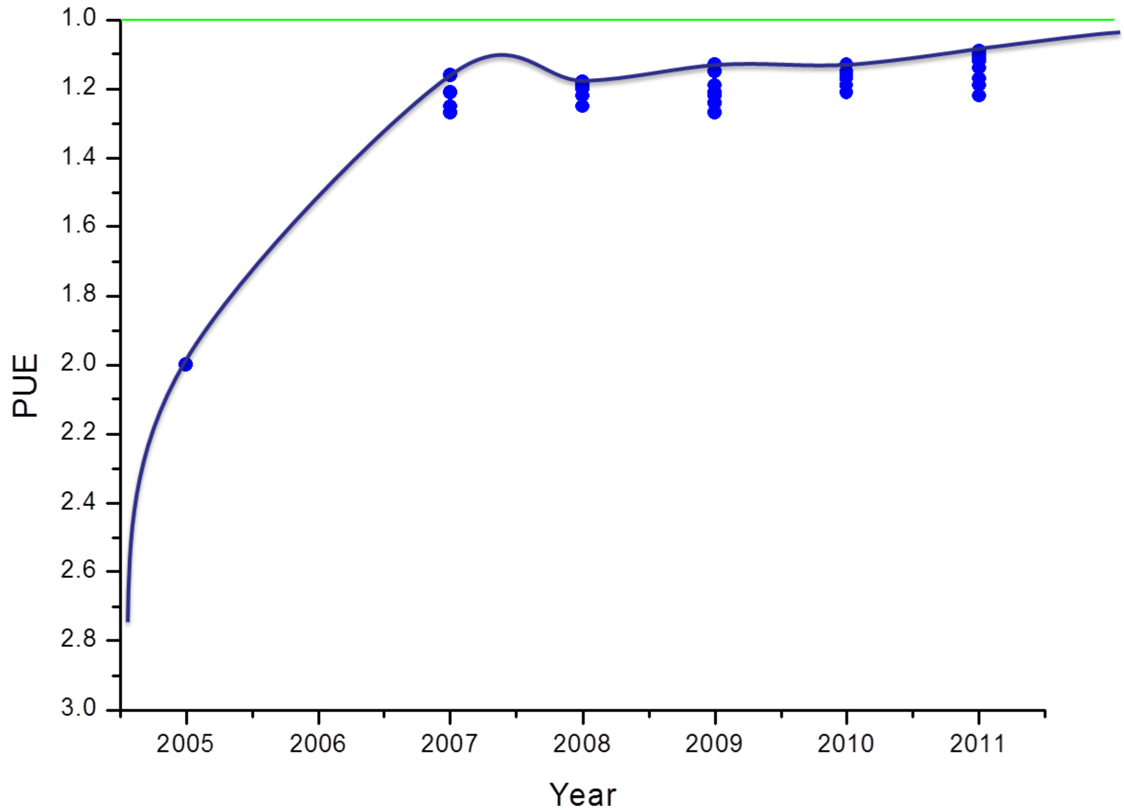


Fig 6 Plot curve of PUE from the year 2005 to 2011

Limitation & Recommendation

Since we couldn't get sufficient and accurate data from Google Data Centers, we made several assumptions in order to make a technology forecasting. We are not sure if the improving rate of PUE performance happens in the real world situation. Also, we can't prove if this technology forecasting can work in other data centers because it is probably other data center operators' improving rate of PUE is unstable every year due to the technology difficulty.

In order to make better forecasting in Google Data Centers, the further research should find the connection in Google and get more information. In this case, we could compare the differences and create a better model for technology forecasting. Furthermore, different technology forecasting tools can be applied in order to verify our results.

In addition, other research also can focus on measuring different metric's performance such as Carbon Usage Effectiveness (CUE). It probably will give us another insight of how "green" Data Centers are.

Conclusion

Our research provides a simple method how to make a technology forecasting with limited information. Based on the concept of TFDEA and Malmquist Index, we examined the power consumption efficiency in Google Data Centers. The results show that Google Data Centers kept walking on the pace of adopting green technology and successfully achieved this goal from 2005 to 2011. Moreover, we made a technology forecasting which shows when Google Data Centers could get the optimized value of PUE (1.0).

References

[1] Fantecchio, Mark. “ IBM to build ‘cloud computing’ data.” News Writer.

4 Aug. 2008 <http://searchcloudcomputing.techtarget.com/news/1355040/IBM-to-build-cloud-computing-data-center>

[2] Taylor, Colleen. “Facebook will build more data centers.” GigaWeb.

6 July. 2011<http://gigaom.com/cloud/facebook-data-growth-data-center/>

[3] Shad, Adam. “ Dell cloud computing data center opens in China.” IIDG New Service

6July.2011<http://www.infoworld.com/d/data-center/dell-cloud-computing-data-center-opens-in-china-273>

[4] Thibodeau, Patrick. “Apple, Google, Facebook turn N.C. into data center hub”. ComputerworldWeb.

3June.2011<http://www.computerworld.com/s/article/9217259/Apple_Google_Facebook_turn_N.C._into_data_center_hub>

[5] Normandeau, Kevin. “Optimizing Your Data for cloud computing.” Data CenterKnowlege Web.

24Feb.2011<http://www.datacenterknowledge.com/archives/2011/02/24/optimizing-your-data-center-for-cloud-computing/>

[6] ICF. (March 2009). Opportunities for Combined Heat and Power in Data Centers. Virgnia: Oak Ridge National Laboratory

[7] Kumar, R. (2008, October 20). *Media Relations*. Retrieved January 2009, from Gartner: http://www.gartner.com/it/page.jsp?id=781012

[8] Emerson Inc (July 2009. Emerson Unveils State-of-the-Art Global Data Center in St. Louis <http://www.emerson.com/en-US/newsroom/news-releases/emerson-corporate-news/Pages/Emerson-Global-Data-Center-St-Louis.aspx>

[9] Firth, Simon . " The Cow Powered Data Centers." HP.

2010 <http://www.hpl.hp.com/news/2010/apr-jun/wastefordatacenter.html>

[10] Godinho, Roger. "Data Center." DataCenterWeb.

(September 2000)< http://searchdatacenter.techtarget.com/definition/data-center>

[11] EPA.(2007). Report to Congress on Server and Data Center Energy Efficiency – Public Law 109-43. Washington: Environmental Protection Agency

[12] Caldow, J. (2008) "The Greening of Government::A Study of How Governments Define the Green Agenda".

<http://www-01.ibm.com/industries/government/ieg/pdf/green_gov_agenda.pdf p.8 Clarkea, L., Weyant, J., & Birky, A. (2006)>. On the sources of technological change: Assessing the evidence. *Energy Economics* , pp. 579-595

[13] Miller, Rich. "Google Data Center: $3,000 A Square foots?"

12 Nov.2007 http://www.datacenterknowledge.com/archives/2007/11/12/google-data-centers-3000-a-square-foot/

[14][15][16][17] D.Alger, *Grow a greener data center*, Pearson Education, 2009

[18] No Reference

[19] No Reference

[20] No Reference

[21] No Reference

[22] Meisner, David., Wenish, F, Thomas. "Does Low-Power Design Imply Energy Efficiency for Data Centers?" Advancy Computer Architecture Lab: University of Michigan

[23] Wiley College. <http://www.wiley.com/college/dec/meredith298298/resources/addtopics/addtopic_s_02p.html>

[24] O.L. Inman, T.R. Anderson, and R.R. Harmon, "Predicting US jet fighter aircraft introductions from 1944 to 1982: A dogfight between regression and TFDEA," *Technological Forecasting and Social Change*, vol. 73, 2006, pp. 1178–1187.

[25] T. Anderson, R., S. Grosskopf, L. Inman, and X. Song, "Further examination of Moore's law with data envelopment analysis," *Technological Forecasting and Social Change*, vol. 69, 2002, pp. 465–477.

[26] T.R. Anderson, T.U. Daim, and J. Kim, "Technology forecasting for wireless communication," *Technovation*, vol. 28, 2008, pp. 602–614.

[27] Xue. Xiaolong, Shen. Qiping, Wang. Yaowu, Lu. Jinfeng, "Measuring the Productivity of the Construction Industry in China by Using DEA-Based Malmquist Productivity Indices", Harbin Institute of Technology, vol. 134, 2008