Expert Systems with Applications 39 (2012) 12061–12070

Contents lists available at SciVerse ScienceDirect

Expert Systems with Applications

journal homepage: www.elsevier.com/locate/eswa

Evaluating decision-making performance in a grid-computing environment using DEA

A. Fernández-Montes [a,*], F. Velasco [b], J.A. Ortega [a]

[a] Departamento de Lenguajes y Sistemas Informáticos, Universidad de Sevilla, Spain
[b] Departamento de Economía Aplicada I, Universidad de Sevilla, Spain

ARTICLE INFO

Keywords:
DEA
Energy policies
Efficiency
Performance
Return-to-scale

ABSTRACT

Energy saving involves two direct benefits: sustainability and cost reduction, both of which Information Technologies must be aware. In this context, clusters, grids and data centres represent the hungriest consumers of energy. Energy-saving policies for these infrastructures must be applied in order to maximize their resources. The aim of this paper is to compare how efficient these policies are in each location of a grid infrastructure. By identifying efficient policies in each location and the slack in inputs and outputs of the inefficient locations, Data Envelopment Analysis presents a very useful technique for comparing and improving efficiency level. This work enables managers to uncover any misuse of resources so that corrective action can be taken.

1. Introduction

Data Envelopment Analysis (DEA) is a nonparametric method to provide a relative efficiency assessment (called DEA efficient) for a group of decision-making units (DMU) or for productive efficiency (aka technical efficiency) with a multiple number of inputs and outputs. DEA was first proposed in Charnes, Cooper, and Rhodes (1978) and is commonly used in operations research and economics to empirically measure productive efficiency of DMUs. In order to determine whether a DMU is efficient is as easy as checking if the DMU is on the "frontier" of the production possibility set. In this way, DEA identifies a "frontier" on which the relative performance of all utilities in the sample can be compared.

In recent years, a great variety of applications of DEA have appeared for the evaluation of the performances of many kinds of entities engaged in various contexts. DEA is especially useful when examining the nature of complex (often unknown) relations between multiple inputs and multiple outputs. DEA has been used both in private (Amirteimoori & Emrouznejad, 2012; Chiang & Hwang, 2010; Eilat, Golany, & Shtub, 2008; Emrouznejad, Parker, & Tavares, 2008) and in public contexts (Afonso, Schuknecht, & Tanzi, 2010; Gonzalez-Rodriguez, Velasco-Morente, & González-Abril, 2010).

Regarding energy efficiency studies, DEA is commonly applied for the study and comparison of the performance and efficiency of energy industries, above all in the electricity industry, see (Pérez-Reyes & Tovar, 2009; Pombo & Taborda, 2006; Tovar, Javier Ramos-Real, & de Almeida, 2011; Vaninsky, 2006; Weyman-Jones, 1991). More recently, it has also been applied to IT companies in Serrano-cinca and Fuertes-calle (2005). Recently, it has also been popularized in environmental performance measurement due to its empirical applicability.

In this work, DEA is used as a method to compare energy-consumption efficiency between each Grid'5000 location, where productive efficiency is measured as the energy consumed to run Grid'5000 jobs at each location.

The rest of this paper is structured as follows: Section 2 includes a brief introduction to DEA methodology used in this paper. Various on–off policies, designed to save energy are presented, and a comparison between current energy consumption and the results of each on–off policy are given in Section 3. The way in which jobs can be scheduled between resources is shown in Section 4. Software developed for testing and simulation is explained in Section 5 and the dataset used for DEA is described and presented. Finally, in Sections 6 and 7, results are given and conclusions are drawn.

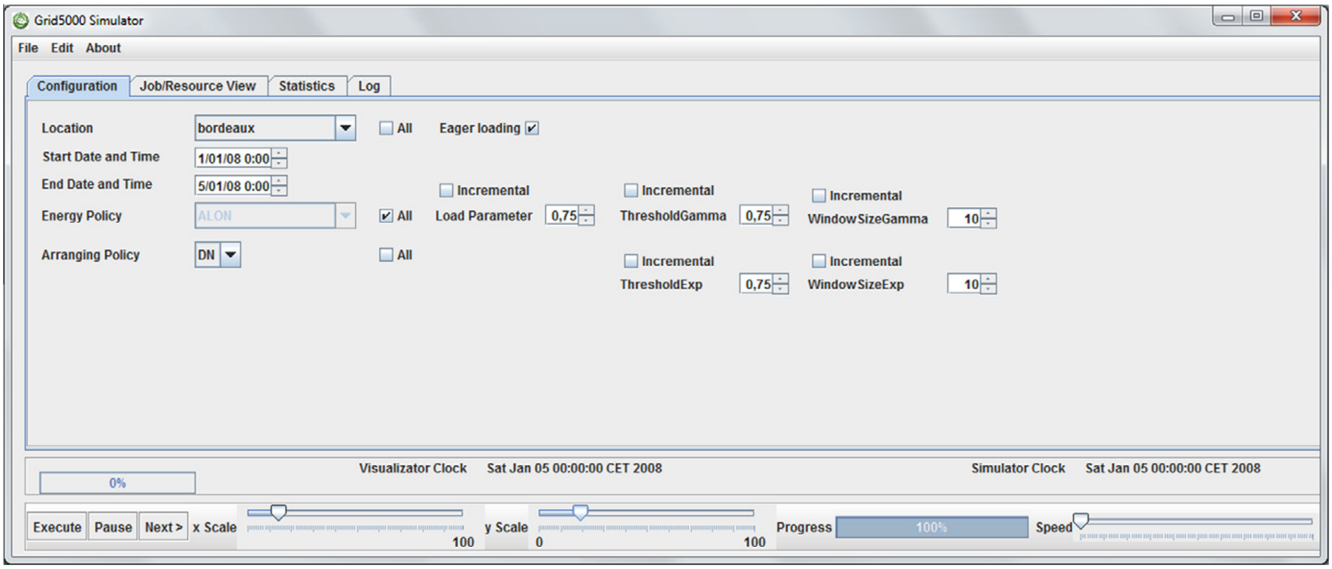
2. Data Envelopment Analysis

DEA has been successfully applied to several sectors. The method establishes a best-practice production frontier (or envelop) based on the empirical input and output data on DMUs. It determines the level of production inefficiency of a DMU by projecting the unit onto the frontier. The original DEA model, introduced in Charnes et al. (1978), was set up with input orientation and assumes constant returns to scale (CRS). In an input-oriented model, the desired output level is achieved by minimizing the production inputs. The CRS

---

* Corresponding author. Tel.: +34 626215333; fax: +34 954557139.
E-mail address: afdez@us.es (A. Fernández-Montes).

0957-4174/$ - see front matter © 2012 Elsevier Ltd. All rights reserved.
http://dx.doi.org/10.1016/j.eswa.2012.04.028



Location: bordeaux ☐ All
Eager loading ☑
Start Date and Time: 1/01/08 0:00
End Date and Time: 5/01/08 0:00
Energy Policy: ALON ☑ All
☐ Incremental Load Parameter: 0,75
☐ Incremental ThresholdGamma: 0,75
☐ Incremental WindowSizeGamma: 10
Arranging Policy: DN ☐ All
☐ Incremental ThresholdExp: 0,75
☐ Incremental WindowSizeExp: 10
0%
Visualizator Clock: Sat Jan 05 00:00:00 CET 2008
Simulator Clock: Sat Jan 05 00:00:00 CET 2008
Progress: 100%

Fig. 1. Configuration tab presenting setup parameters for a batch of simulations.

assumption suggests that an increase in the amount of inputs utilized would lead to a proportional increase in the amount of outputs generated. The original model has been subsequently extended and numerous variations of DEA. For example, a DEA model can be set up to be output-oriented (Charnes, Cooper, & Rhodes, 1981), which attempts to maximize outputs with a set of available inputs. Another significant development of the DEA model by Banker, Charnes, and Cooper (BCC) (Banker, Charnes, & Cooper, 1984) allows for variable returns to scale (VRS). The VRS assumption suggests that an increase in the amount of inputs utilized can lead to a proportional or non-proportional change in the amount of outputs generated (Barkhi & Kao, 2010).

3. Energy policies at a glance

Energy policies establish the managing of grid resources. While other research works try to reduce the make-span (Tseng, Chin, & Wang, 2009), the policies shown in this work try to describe and compute what to do with a resource once a job finishes its execution. Thus, each energy policy decides whether to leave a resource switched on or to switch it off depending on the purpose of the policy. The following subsections show energy policies implemented in Grid'5000 Toolbox.

3.1. Always On

This is the simplest energy policy. It never switches resources off, under any condition, and hence resources stay idle, waiting for a new job to be run. Grid'5000 is currently running this way, and therefor these consumption results can be used for comparison with other energy policies in order to know how much energy would have been saved. The number of times resources are switched off or on are always zero, and therefor the stress upon the resource is minimal.

3.2. Always Off

This policy always switches resources off, under any condition, and hence a resource starts shutting down immediately after any job finishes, and remains switched off. If a new job arrives, resources assigned have to be booted to run that job. This booting is carried out within reservation limits, and hence the user cannot make effective use of the resources until they are booted. This policy is usually the best regarding energy consumption results, but the number of times a resource is booted up and shut down is always maximum, and the stress produced on the hardware components is the highest, which is seldom desirable.

3.3. Switch off randomly

This policy randomly switches resources off or leaves them idle by following a Bernoulli distribution whose parameter is equal to 0.5 when a job finishes. Hence, the number of times resources are switched off or left idle tends towards 50%, and results tend to be half-way between those of the *Always Off* and *Always On* policies (regarding the times resources are switched off and those of energy consumption).

3.4. Load

Load can be defined as the percentage of resources that are *On* among the clusters of a location. This policy queries this information and leaves resources idle or switches resources off if the load when finishing a job is greater than a certain threshold or less than a threshold respectively. This threshold is a parameter selected from the GUI from 0 to 1.

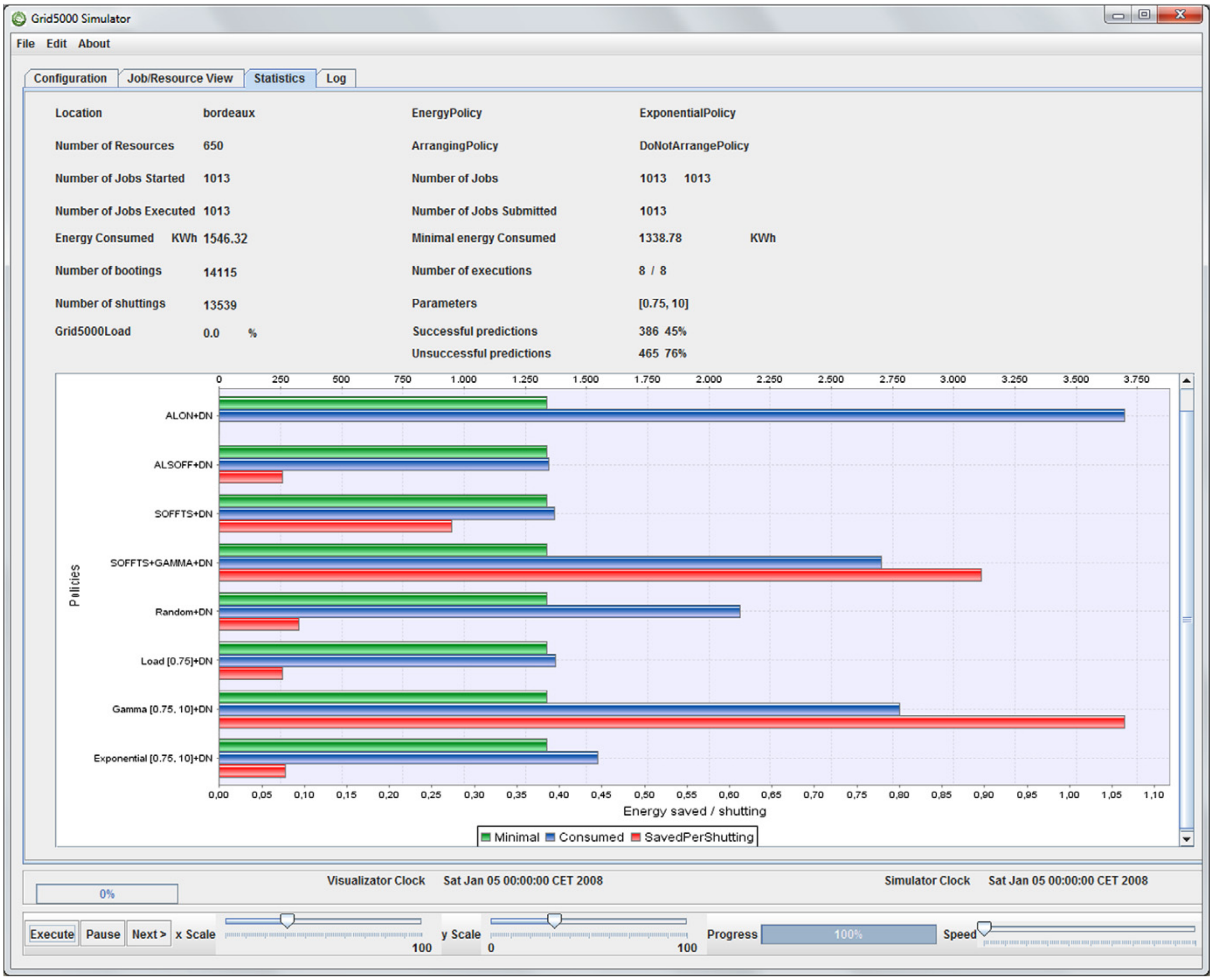
3.5. Switch off T_S

T_S is defined *as the minimum time which ensures an energy saving if a resource is switched off between two jobs* (Orgerie, Lefèvre, & Gelas, 2008). T_S can be computed as follows:

$$T_S = \frac{E_s - P_{Off} * \delta_{tot} + E_{On\rightarrow Off} + E_{Off\rightarrow On}}{P_{Idle} - P_{Off}}$$

where P_{Off} and P_{Idle} refer to the power consumption in watts of a given resource when it is *Off* and *Idle*, respectively. $E_{On\rightarrow Off}$ and $E_{Off\rightarrow On}$ refers to the required energy in joules for a given resource to boot or switch it off respectively. E_S is the energy saved during T_S seconds. Finally, $\delta_{tot} = \delta_{On\rightarrow Off} + \delta_{Off\rightarrow On}$, which is the total time a given resource needs for it to be switched off and switched on.

This energy policy queries the agenda to check if the next submitted jobs are going to be run in the grid in less than T_S. This policy computes the number of resources that are going to be



Location bordeaux
Number of Resources 650
Number of Jobs Started 1013
Number of Jobs Executed 1013
Energy Consumed KWh 1546.32
Number of bootings 14115
Number of shuttings 13539
Grid5000Load 0.0 %
EnergyPolicy ExponentialPolicy
ArrangingPolicy DoNotArrangePolicy
Number of Jobs 1013 1013
Number of Jobs Submitted 1013
Minimal energy Consumed 1338.78 KWh
Number of executions 8 / 8
Parameters [0.75, 10]
Successful predictions 386 45%
Unsuccessful predictions 465 76%

0%
Visualizator Clock Sat Jan 05 00:00:00 CET 2008
Simulator Clock Sat Jan 05 00:00:00 CET 2008

Fig. 2. Statistics tab presenting results for a batch of simulations.

needed within a time period less than T_S, and leaves idle or shuts resources down of the job which has just finished, accordingly. In this way, the simulator attempts to minimize the cycles of booting up and shutting down when these cycles are not going to save energy.

3.6. Exponential

The Exponential distribution, denoted by $Exp(\lambda)$, describes the time between events in a Poisson process, i.e. a process in which events occur continuously and independently at a constant average rate $(1/\lambda)$. Under the hypothesis that the arrival of new jobs follows an Exponential distribution, this energy policy attempts to predict the arrival of new jobs. Thus, to compute the λ parameter, every time a job finishes, then the mean time between the last jobs is computed, denoted by μ. Hence, $\lambda = 1/\mu$ according to the of method of maximum likelihood. The probability of the arrival of a new job is then computed by means of the Exponential cumulative density function (cdf) as $cdf(T_s) = 1 - e^{-T_s/\mu}$. Therefore, given a *threshold* value:

$$\begin{cases} \text{if } cdf(T_s) \geqslant \textit{threshold} \text{ then leave resources } \textit{Idle} \\ \text{if } cdf(T_s) < \textit{threshold} \text{ then switch resources } \textit{Off} \end{cases}$$

3.7. Gamma

The Gamma distribution, denoted by $\Gamma(\theta, \kappa)$, is frequently used as a probability model for waiting times, and is a more general model than that given by the *Exponential*. Under the hypothesis that the arrival of new jobs follows a Gamma distribution, this energy policy attempts to predict the arrival of new jobs. The parameters computed every time a job finishes are:

- Number of resources available, as *resourcesAvailable*. These are the resources that are *Idle* and ready to accept new jobs.
- Mean resources used by the last jobs, as *meanResources*. The total number of resources used by the last jobs is computed and divided by the number of jobs. The number of last jobs number is a selected window size.
- Mean duration of these last jobs, as *meanDuration*. The sum of the duration of the last jobs is computed and divided by the number of the last jobs.
- The floor of *resourcesAvailable*/*meanResources*, as *z*.

The parameters of the Gamma distribution are then estimated as: $\theta = 1/meanDuration$ and $\kappa = z + 1$. The probability of the arrival of a new job is then computed by means of the cumulative density function (cdf) with

Table 1
Summary of inputs and outputs.

Location	Outputs		Inputs	
	Saved energy (kW h)	# Jobs deployed	# Resources	# Bootings
		Always Off		
Bordeaux	128,697	345,218	650	4,036,514
Lille	238,159	62,451	618	327,408
Lyon	57,715	134,719	322	927,472
Nancy	94,932	73,934	574	1,668,946
Orsay	132,518	89,048	684	2,111,974
Rennes	152,832	57,987	714	2,328,890
Sophia	48,848	57,533	568	2,337,336
Toulouse	86,531	165,995	434	1,754,930
		Random		
Bordeaux	115,539	345,218	650	2,225,174
Lille	220,282	62,451	618	168,398
Lyon	51,771	134,719	322	494,442
Nancy	64,407	73,934	574	904,920
Orsay	105,075	89,048	684	1,141,004
Rennes	141,222	57,987	714	1,205,530
Sophia	39,918	57,533	568	1,198,338
Toulouse	71,738	165,995	434	922,932
		Load		
Bordeaux	127,089	345,218	650	3,675,094
Lille	238,159	62,451	618	327,408
Lyon	57,708	134,719	322	926,028
Nancy	74,616	73,934	574	1,176,234
Orsay	125,703	89,048	684	1,922,154
Rennes	152,832	57,987	714	2,328,890
Sophia	41,063	57,533	568	1,475,640
Toulouse	86,057	165,995	434	1,667,222
		T_s		
Bordeaux	127,018	345,218	650	2,238,318
Lille	236,793	62,451	618	299,846
Lyon	57,299	134,719	322	538,154
Nancy	90,771	73,934	574	1,297,252
Orsay	130,825	89,048	684	1,384,922
Rennes	152,226	57,987	714	1,392,750
Sophia	46,332	57,533	568	1,271,836
Toulouse	85,250	165,995	434	876,026
		Exponential		
Bordeaux	119,779	345,218	650	1,574,410
Lille	237,688	62,451	618	122,680
Lyon	56,349	134,719	322	612,766
Nancy	92,168	73,934	574	1,168,646
Orsay	127,303	89,048	684	1,387,566
Rennes	152,141	57,987	714	1,770,858
Sophia	48,360	57,533	568	1,847,484
Toulouse	86,203	165,995	434	671,122
		Gamma		
Bordeaux	67,374	345,218	650	1,141,048
Lille	159,213	62,451	618	884
Lyon	31,532	134,719	322	131,106
Nancy	18,833	73,934	574	156,116
Orsay	61,581	89,048	684	623,515
Rennes	116,158	57,987	714	644,109
Sophia	20,017	57,533	568	510,400
Toulouse	39,395	165,995	434	153,326

$$cdf(T_s) = \frac{\gamma(\kappa, T_s/\theta)}{\Gamma(\kappa)}$$

Hence, given a *threshold* value:

$$\begin{cases} \text{if } cdf(T_s) \geqslant threshold \text{ then leave resources } Idle \\ \text{if } cdf(T_s) < threshold \text{ then switch resources } Off \end{cases}$$

4. Arranging policies at a glance

Arranging policies establish the arrangement of jobs for their execution. A job can be moved from a set of resources to another, or a planned job execution can even be moved in time in order to take advantages of resources that are already switched on.

- *Do Nothing* (*DN*): Neither does this policy move jobs in time nor from one resource to another; jobs are executed as defined in the agenda. This is the current behaviour in Grid'5000. The combination of this arranging policy with the energy policy *Always On* in a simulation offers the current Grid'5000 behaviour, and includes results of energy consumption.
- *Simple Aggregation of Jobs* (*SA*): This policy attempts to find resources available (*Idle*) for new jobs. In this way, if a job is assigned to a set of resources which are *Off* and some resources are already switched on and available, we can save the time and

Table 2
Summary of DEA results for CRS, VRS, and scale efficiency.

		B	Li	Ly	N	O	R	S	T	σ	$\bar{x}$
Alwz. Off	crste	1.000	1.000	1.000	0.516	0.583	0.581	0.303	0.908	**0.255**	0.736
	vrste	1.000	1.000	1.000	0.667	0.650	0.670	0.567	0.938	0.177	0.812
	scale	1.000	1.000	1.000	0.773	0.897	0.868	0.535	0.968	0.151	0.880
Random	crste	1.000	1.000	1.000	0.427	0.521	0.581	0.284	0.889	0.273	0.713
	vrste	1.000	1.000	1.000	0.600	0.608	0.671	0.567	0.906	0.187	*0.794*
	scale	1.000	1.000	1.000	0.712	0.858	0.866	0.500	0.981	0.168	0.865
Load	crste	1.000	1.000	1.000	0.464	0.561	0.581	0.297	0.904	0.264	0.726
	vrste	1.000	1.000	1.000	0.675	0.634	0.670	0.601	0.937	**0.172**	0.815
	scale	1.000	1.000	1.000	0.687	0.885	0.868	0.495	0.965	0.171	0.862
T_s	crste	1.000	1.000	1.000	0.502	0.581	0.582	0.294	0.936	0.261	0.737
	vrste	1.000	1.000	1.000	0.657	0.648	0.670	0.567	0.937	0.178	0.810
	scale	1.000	1.000	1.000	0.763	0.896	0.868	0.519	0.999	0.159	**0.881**
Exp.	crste	1.000	1.000	0.944	0.511	0.572	0.581	0.307	1.000	0.259	**0.739**
	vrste	1.000	1.000	1.000	0.663	0.640	0.670	0.567	1.000	0.185	**0.817**
	scale	1.000	1.000	0.944	0.771	0.893	0.868	0.541	1.000	**0.148**	0.877
Gamma	crste	1.000	1.000	1.000	0.406	0.465	0.653	0.257	1.000	*0.295*	0.723
	vrste	1.000	1.000	1.000	0.667	*0.573*	0.726	0.567	1.000	*0.189*	**0.817**
	scale	1.000	1.000	1.000	0.608	0.812	0.899	0.453	1.000	*0.197*	*0.847*
$\sigma(vrse)$		**0.000**	**0.000**	**0.000**	0.025	0.027	0.021	0.013	*0.035*		
$\overline{vrste}$		**1.000**	**1.000**	**1.000**	0.655	0.626	0.680	*0.573*	0.953		**0.811**

the energy needed for them to be switched on. Notice that this policy does not change start or stop times, and hence is transparent to users.

5. Methodology

In order to compare energy efficiency between the locations of the Grid'5000, a software simulator has been developed. Grid'5000 Toolbox[1] replays the progress of the real grid regarding the operation of jobs and resources. Grid'5000 Toolbox is able to compute energy consumption of Grid'5000, and enables the user to establish several parameters including: (a) simulation start-time, (b) simulation stop-time, (c) location, (d) energy policy, and (e) arranging policy. These parameters can be set up through the *Configuration* tab as shown in Fig. 1.

The simulator operation is based on an agenda where jobs are registered, and on a list of resources representing the real resources at the sites. The simulator queries the agenda from simulation start-time to simulation stop-time. Each query is related to current simulation time (the moment in past-time the software is replaying), and hence the agenda seeks jobs and events that occur at given current time. Once the agenda returns new events, the simulator processes them and changes the states of the resources as would be needed for execution in the real world, whilst taking into account the policies selected in order to manage resources and jobs. The energy consumed is computed step by step by means of the information on energy consumption of each resource and on the resource states detailed in the resource list. The results of simulation executions are stored on a spreadsheet where researchers can find details about consumption, the number of times the resources are shut down and booted up, the comparison between minimal energy consumable and current energy consumed, etc. Results are also shown in the *Statistics* tab in a more visual way (see Fig. 2). A battery of tests has been performed in order to compute energy-saving results based on:

- One period of 12 months. From 1st January to 31st December 2008.
- Two arranging policies, *Do Nothing* and *Simple Aggregation of Jobs*.
- The seven energy policies listed in Section 3.
- Various values of several parameters as follows:
  1. *Load* policy. Load threshold parameter from 0.0 to 1 in steps of 0.3. A total of four scenarios.
  2. *Exponential* and *Gamma*. Threshold probability parameter from 0.0 to 1 in steps of 0.3, and window size from 2^0 to 2^8. Hence there are 36 different scenarios for each policy.

From the 162 setups run, the best energy savers have been selected of each policy. From computed results, we select the following inputs and outputs to measure relative efficiency between locations:

- Inputs:
  1. The number of resources at the location. This parameter remains unchanged between simulations. Resources are the entities that run jobs.
  2. The number of times resources have been switched off and booted during the simulation. Each energy policy shows different behaviour when a job finishes, and therefore this input changes between each energy policy simulated.
- Outputs:
  1. The energy saved, in kW h, using a given energy policy. This is the amount of energy that the location would save if a given energy policy were applied.
  2. The number of jobs deployed at each location.

The following table shows the summary of inputs and outputs for each energy policy for which the DEA methodology is computed using, Coelli software (Coelli, 1996) due to its simplicity usage. Results are compared with those produced by other tools, such as Benchmarking library in R language (Bogetoft & Otto, 2010).

6. Input-orientated DEA results

The results computed are input orientated since firms are able to modify their inputs, and hence our study is focused on reducing inputs while maintaining the level of outputs (see Table 1).

---

[1] This software can be downloaded and executed from the web of the Idinfor research group (Idinfor, 2011).

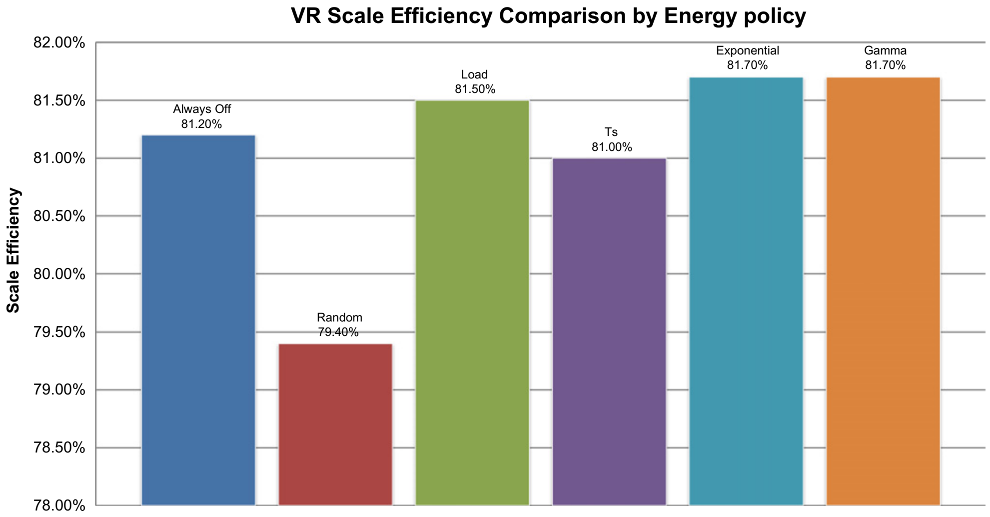


Fig. 3. Comparison of energy policies for VR scale technical efficiency.

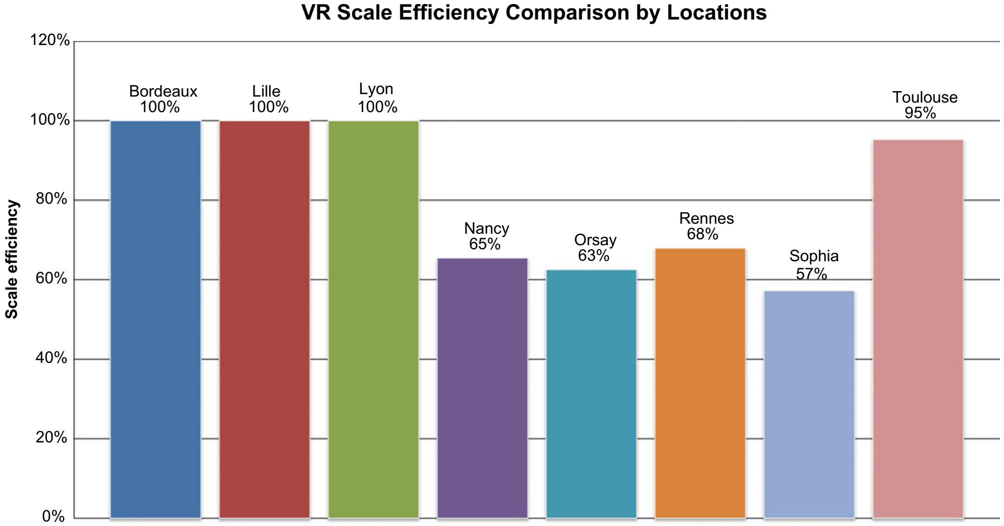


Fig. 4. Comparison of locations VR scale technical efficiency.

Table 2 shows the results generated by the DEA tool (Coelli, 1996) for an input-orientated DEA with 2 inputs, 2 outputs and 8 firms (locations[2]), and these are grouped by energy policy. CRSTE (constant returns-to-scale technical efficiency), VRSTE (variable returns-to-scale technical efficiency) and Scale (scale efficiency) results are shown. Mean and standard deviation are computed for each energy policy and each location.

Results in Table 2 and Fig. 3 show that the most efficient energy policies are those of *Exponential* and *Gamma* (Sections 3.6 and 3.7) in terms of VRSTE ($\bar{x} = 0.817$), followed by the *Load* and *Always Off* energy policies ($\bar{x} = 0.815$). On the other hand, the overall results of *Random* policy show this to be the least efficient ($\bar{x} = 0.754$). In terms of dispersion, the least dispersion is reached using the *Load* policy ($\sigma = 0.172$), which indicates that this policy works homogeneously for any of the policies. Fig. 3 shows a graphical comparison of scale efficiency per energy policy.

In the analysis of locations, it can be observed that Bordeaux, Lille and Lyon are the most efficient locations (VRSTE equals 1.000 for these policies), followed by Toulouse, and that the least efficient locations are Sophia and Orsay, followed by Nancy and Rennes. In terms of dispersion, Bordeaux, Lille and Lyon have the most homogeneous behaviour between policies, followed by Sophia, with Toulouse being the location whose performance is the most dispersed between policies, followed by Sophia, Rennes and Nancy. Fig. 4 shows this graphical comparison of VRSTE per locations.

As a consequence of these analyses, corrections on inputs and outputs can be carried out. Table 3 shows peers per location, including weights and corrections proposed per location/policy. Notice that the type of correction (increase or decrease) remains the same within each location, which constitutes further confirmation of the validity of these corrections. For example, the proposed corrective actions for Nancy are: increase the number of jobs deployed, decrease the number of resources (as they are underused) and decreasing the number of power cycles (since the policies are not working as efficiently as those in other locations).

By taking into account that certain locations are underused, the system manager could better balance the workload through the relocation of jobs from efficient locations to underused

---

[2] B, Li, Ly, N, O, R, S, and T stand for Bordeaux, Lille, Lyon, Nancy, Orsay, Rennes, Sophia, and Toulouse, respectively.

Table 3
Peers per location and per energy policy and correction proposals.

Policy	Peers	Corrections		
		Jobs	Resources	Bootings
Bordeaux				
Alwz. Off	B (1.000)	↔	↔	↔
Random	B (1.000)	↔	↔	↔
Load	B (1.000)	↔	↔	↔
T_s	B (1.000)	↔	↔	↔
Exp.	B (1.000)	↔	↔	↔
Gamma	B (1.000)	↔	↔	↔
Summary	Bordeaux	↔	↔	↔
Lille				
Alwz. Off	Li (1.000)	↔	↔	↔
Random	Li (1.000)	↔	↔	↔
Load	Li (1.000)	↔	↔	↔
T_s	Li (1.000)	↔	↔	↔
Exp.	Li (1.000)	↔	↔	↔
Gamma	Li (1.000)	↔	↔	↔
Summary	Lille	↔	↔	↔
Lyon				
Alwz. Off	Ly (1.000)	↔	↔	↔
Random	Ly (1.000)	↔	↔	↔
Load	Ly (1.000)	↔	↔	↔
T_s	Ly (1.000)	↔	↔	↔
Exp.	Ly (1.000)	↔	↔	↔
Gamma	Ly (1.000)	↔	↔	↔
summary	Lyon	↔	↔	↔
Nancy				
Alwz.Off	Li (0.206) Ly (0.794)	▲	▼	▼
Random	Li (0.075) Ly (0.925)	▲	▼	▼
Load	Li (0.221) Ly (0.779)	▲	▼	▼
T_s	Li (0.186) Ly (0.814)	▲	▼	▼
Exp.	Li (0.198) Ly (0.802)	▲	▼	▼
Gamma	Li (0.207) Ly (0.793)	▲	▼	▼
Summary	Lille and Lyon	▲	▼	▼
Orsay				
Alwz. Off	Li (0.415) Ly (0.585)	▲	▼	▼
Random	Li (0.316) Ly (0.684)	▲	▼	▼
Load	Li (0.377) Ly (0.623)	▲	▼	▼
T_s	Li (0.410) Ly (0.590)	▲	▼	▼
Exp.	Li (0.391) Ly (0.609)	▲	▼	▼
Gamma	Li (0.235) Ly (0.765)	▲	▼	▼
Summary	Lille and Lyon	▲	▼	▼
Rennes				
Alwz.Off	Li (0.527) Ly (0.473)	▲	▼	▼
Random	Li (0.531) Ly (0.469)	▲	▼	▼
Load	Li (0.527) Ly (0.473)	▲	▼	▼
T_s	Li (0.529) Ly (0.471)	▲	▼	▼
Exp.	Li (0.528) Ly (0.472)	▲	▼	▼
Gamma	Li (0.663) Ly (0.337)	▲	▼	▼
summary	Lille and Lyon	▲	▼	▼
Sophia				
Alwz. Off	Ly (1.000)	▲	▼	▼
Random	Ly (1.000)	▲	▼	▼
Load	Li (0.065) Ly (1.000)	▲	▼	▼
T_s	Ly (1.000)	▲	▼	▼
Exp.	Ly (1.000)	▲	▼	▼
Gamma	Ly (1.000)	▲	▼	▼
Summary	Lyon	▲	▼	▼
Toulouse				
Alwz. Off	B (0.179) Li (0.089) Ly (0.732)	▲	▼	▼
Random	B (0.167) Li (0.055) Ly (0.777)	▲	▼	▼
Load	B (0.179) Li (0.088) Ly (0.733)	▲	▼	▼
T_s	B (0.178) Li (0.086) Ly (0.735)	▲	▼	▼
Exp.	T (1.000)	↔	↔	↔
Gamma	T (1.000)	↔	↔	↔
Summary	Bordeaux, Lille, Lyon, Toulouse	▲	▼	▼

locations. The system manager could also unplug a number of resources at underused locations, in the search for a threshold which guarantees both satisfaction of users and energy saving objectives.

6.1. Detailed analysis of Always Off energy policy technical efficiency

Sophia is selected to illustrate this energy policy. Sophia is the least efficient location in general, and also the least efficient

Table 4
Corrections proposed for Sophia under the *Always Off* energy policy.

Results for firm: Sophia
Technical efficiency = 0.567
Scale efficiency = 0.535 (irs)

Projection summary

Variable		Original value	Radial movement	Slack movement	Projected value
Output	Saved energy	48,848	0	8867	57,715
Output	# Jobs	57,533	0	77,186	134,719
Input	# Resources	568	−246	0	322
Input	# Bootings	2,337,336	−1,012,296	−397,567	927,472

Listing of peers

Peer	Lambda weight
Lyon	1.000

Table 5
Corrections proposed for Orsay under the *Random* energy policy.

Results for firm: Orsay
Technical efficiency = 0.608
Scale efficiency = 0.858 (irs)

Projection summary

Variable		Original value	Radial movement	Slack movement	Projected value
Output	Saved energy	105,075	0	0	105,075
Output	# Jobs	89,048	0	22,811	111,859
Input	# Resources	684	−268	0	415
Input	# Bootings	1,141,004	−447,675	−302,021	391,307

Listing of peers

Peer	Lambda weight
Lille	0.316
Lyon	0.684

Table 6
Corrections proposed for Nancy under the *Load* energy policy.

Results for firm: Nancy
Technical efficiency = 0.675
Scale efficiency = 0.687 (irs)

Projection summary

Variable		Original value	Radial movement	Slack movement	Projected value
Output	Saved energy	74,616	0	22,948	97,564
Output	# Jobs	73,934	0	44,823	118,757
Input	# Resources	574	−186	0	387
Input	# Bootings	1,176,234	−382,423	0	793,810

Listing of peers:

Peer	Lambda weight
Lille	0.221
Lyon	0.779

performing under the *Always Off* energy policy. The corrective actions recommended for this location and policy are detailed in Table 4. This location presents a CRS technical efficiency of 0.303 and a VRS technical efficiency of 0.567, and hence in order to achieve overall efficiency and to belong to the efficient frontier it must reduce input and increase output. This means that number of bootings and shuttings should be reduced by in 1.4 million (−60%), and, most importantly 246 resources (−43%) should be removed. In addition, these measures have to be followed by an increase of 77,186 (+134%) in the number of jobs run at this location and a reduction of 8867 kW h (−18%) in energy consumption.

The peer for this location is Lyon, which belongs to the segment of the production frontier where Sophia has to tend. Within these new dimensions, Sophia will make the most of its resources and will become efficient in the means of production. The other non-efficient locations should be corrected in a similar way.

6.2. Detailed analysis of Random energy policy technical efficiency

Orsay is selected to illustrate this energy policy although it is not the least efficient location for this energy policy. The corrective actions recommended for this location and policy are detailed in Table 5. Orsay presents a CRS technical efficiency of 0.303 and a VRS technical efficiency of 0.521, and hence in order to achieve overall efficiency and to belong to the efficient frontier it must reduce input and increase output. This means that the number of bootings and shuttings in must be reduced by 749,696 (−65%), and most importantly, 268 resources (−39%) should be removed.

Table 7
Corrections proposed for Toulouse under the T_s energy policy.

Results for firm: Toulouse
Technical efficiency = 0.937
Scale efficiency = 0.999 (irs)

Projection summary

Variable		Original value	Radial movement	Slack movement	Projected value
Output	Saved energy	85,250	0	0	85,250
Output	# Jobs	165,995	0	0	165,995
Input	# Resources	434	−27	0	406
Input	# Bootings	876,026	−55,393	0	820,632

Listing of peers

Peer	Lambda weight
Lille	0.086
Lyon	0.735
Bordeaux	0.178

Table 8
Corrections proposed for Rennes under the *Exponential* energy policy.

Results for firm: Rennes
Technical efficiency = 0.670
Scale efficiency = 0.868 (irs)

Projection summary

Variable		Original value	Radial movement	Slack movement	Projected value
Output	Saved energy	152,141	0	0	152,141
Output	# Jobs	57,987	0	38,556	96,543
Input	# Resources	714	−235	0	478
Input	# Bootings	1,770,858	−584,429	−832,549	353,879

Listing of peers

Peer	Lambda weight
Lille	0.528
Lyon	0.472

Table 9
Corrections proposed for Nancy under the *Gamma* energy policy.

Results for firm: Nancy
Technical efficiency = 0.667
Scale efficiency = 0.608 (irs)

Projection summary

Variable		Original value	Radial movement	Slack movement	Projected value
Output	Saved energy	18,832	0	39,073	57,906
Output	# Jobs	73,934	0	45,857	119,791
Input	# Resources	574	−190	0	383
Input	# Bootings	156,116	−51,909	0	104,206

Listing of peers

Peer	Lambda weight
Lyon	0.793
Lille	0.207

In addition, these measures have to be followed by an increase of 22,811 (+25%) in jobs run at this location.

The peers for this location are Lyon and Lille, which both belong to the segment of the production frontier where Orsay has to tend. Within these new dimensions, Orsay will make the most of its resources and will become efficient in the means of production. The other non-efficient locations should be corrected in a similar way.

6.3. Detailed analysis of Load energy policy technical efficiency

Nancy is selected to illustrate this energy policy although it is not the least efficient location for this energy policy. The corrective actions for this location and policy are detailed in Table 6. Nancy presents a CRS technical efficiency of 0.464 and a VRS technical efficiency of 0.675, and hence in order to achieve overall efficiency and to belong to the efficient frontier it must reduce input and increase output. This means that the number of bootings and shuttings must be reduced by 382,423 (−32%), and, most importantly 186 resources (−32%) should be removed. In addition, these measures have to be followed by an increase of 44,823 (+60%) in the jobs run at this location and a reduction of 22,948 kW h (+30%) in energy consumption.

The peers for this location are Lyon and Lille, which both belong to the segment of the production frontier where Nancy has to tend. Within these new dimensions, Nancy will make the most of its resources and will become efficient in the means of production. The other non-efficient locations should to be corrected in a similar way.

6.4. Detailed analysis of T_s energy policy technical efficiency

Toulouse is selected to illustrate this energy policy although it is not the least efficient location for this energy policy. The corrective actions recommended for this location and policy are detailed in Table 7. Toulouse presents a CRS technical efficiency of 0.936 and a VRS technical efficiency of 0.937, and hence in order to achieve overall efficiency and to belong to the efficient frontier it must reduce input but it has no needs of increasing output. This means that the number of bootings and shuttings must be reduced by 55,393 (−6%), and most importantly 27 resources (−6%) should be removed.

The peers for this location are Lyon, Lille and Bordeaux which belong to the segment of the production frontier where Toulouse has to tend. Within these new dimensions, Toulouse will make the most of its resources and will become efficient in the means of production. The other non-efficient locations should be corrected in a similar way.

6.5. Detailed analysis of Exponential energy policy technical efficiency

Rennes is selected to illustrate this energy policy although it is not the least efficient location for this energy policy. The corrective actions for this location and policy are detailed in Table 8. Rennes presents a CRS technical efficiency of 0.581 and a VRS technical efficiency of 0.670, and hence in order to achieve overall efficiency and to belong to the efficient frontier it must reduce input and increase the output ‘number of jobs’. This means that the number of bootings and shuttings must be reduced by 1.4 millions (−80%), and most importantly 235 resources (−32%) should be removed. In addition, these measures have to be followed by an increase of 38,556 (+66%) in the jobs run at this location.

The peers for this location are Lyon and Lille which belong to the segment of the production frontier where Rennes has to tend. Within these new dimensions, Rennes will make the most of its resources and will become efficient in the means of production. The other non-efficient locations should be corrected in a similar way.

6.6. Detailed analysis of Gamma energy policy technical efficiency

Nancy is selected to illustrate this energy policy although it is not the least efficient location for this energy policy. The corrective actions recommended for this location and policy are detailed in Table 9. Nancy presents a CRS technical efficiency of 0.406 and a VRS technical efficiency of 0.608, and hence in order to achieve overall efficiency and to belong to the efficient frontier it must reduce input and increase output. This means the number of bootings and shuttings must be reduced by 51,909 (−33%), and most importantly 190 resources (−33%) should be removed. In addition, these measures have to be followed by an increase of 45,857 (+62%) in the jobs run at this location and a reduction of 39,073 kW h (+207%) in energy consumption.

The peers for this location are Lyon and Lille which belong to the segment of the production frontier where Nancy has to tend. Within these new dimensions, Nancy will make the most of its resources and will become efficient in the means of production. The other non-efficient locations should be corrected in a similar way.

7. Conclusions

The hypothesis that DEA methodology can be useful for the analysis of technical efficiency in Grid computing environments has been proved. Data Envelopment Analysis enables Grid managers to detect which grid locations present the best and worst performance in terms of energy consumption and efficiency. This methodology also enables several energy policies to be analyzed with regard to their behaviour and the potential differences between running a certain policy at one particular location or another.

By means of DEA methodology, system managers are armed with knowledge of which locations are underused and hence decisions regarding the switching off of resources and the relocation of underused locations can be made in order to achieve a better utilization of the Grid infrastructure as a whole.

Acknowledgements

This research is partially supported by the projects of the Spanish Ministry of Economy and Competitiveness ARTEMISA (TIN2009-14378-C02-01) and Simon (TIC-8052) of the Andalusian Regional Ministry of Economy.

References

Afonso, A., Schuknecht, L., & Tanzi, V. (2010). Public sector efficiency: Evidence for new EU member states and emerging markets. *Applied Economics, 42*, 2147–2164.

Amirteimoori, A., & Emrouznejad, A. (2012). Optimal input/output reduction in production processes. *Decision Support Systems, 52*, 742–747.

Banker, R. D., Charnes, A., & Cooper, W. W. (1984). Some models for estimating technical and scale inefficiencies in data envelopment analysis. *Management Science, 30*, 1078–1092.

Barkhi, R., & Kao, Y.-C. (2010). Evaluating decision making performance in the GDSS environment using data envelopment analysis. *Decision Support Systems, 49*, 162–174.

Bogetoft, P., & Otto, L. (2010). *Benchmarking with DEA, SFA, and R* (Vol. 157). Springer.

Charnes, W., Cooper, W. W., & Rhodes, E. (1978). Measuring the efficiency of decision making units. *European Journal of Operational Research, 2*, 429–444.

Charnes, A., Cooper, W. W., & Rhodes, E. (1981). Evaluating program and managerial efficiency: An application of data envelopment analysis to program follow through. *Management Science, 27*, 668–697.

Chiang, K., & Hwang, S.-N. (2010). Efficiency measurement for network systems IT impact on firm performance. *Decision Support Systems, 48*, 437–446.

Coelli, T. (1996). A guide to DEAP version 2.1: A data envelopment analysis (computer) program by. *Frontiers A Journal of Women Studies, 96/08*, 1–49.

Eilat, H., Golany, B., & Shtub, A. (2008). R&D project evaluation: An integrated DEA and balanced scorecard approach. *Omega, 36*, 895–912.

Emrouznejad, A., Parker, B. R., & Tavares, G. (2008). Evaluation of research in efficiency and productivity: A survey and analysis of the first 30 years of scholarly literature in DEA. *Socio-Economic Planning Sciences, 42*, 151–157.

Gonzalez-Rodriguez, M., Velasco-Morente, F., & González-Abril, L. (2010). La eficiencia del sistema de protección social español en la reducción de la pobreza. *Papeles de población, 16*, 123–154.

Idinfor (2011). <http://madeira.lsi.us.es/GrupoIdinfor>.

Orgerie, A.-C., Lefèvre, L., Gelas, J.-P. (2008). Save watts in your grid: Green strategies for energy-aware framework in large scale distributed systems. In *2008 14th IEEE international conference on parallel and distributed systems* (pp. 171–178).

Pérez-Reyes, R., & Tovar, B. (2009). Measuring efficiency and productivity change (PTF) in the Peruvian electricity distribution companies after reforms. *Energy Policy, 37*, 2249–2261.

Pombo, C., & Taborda, R. (2006). Performance and efficiency in Colombia's power distribution system: Effects of the 1994 reform. *Energy Economics, 28*, 339–369.

Serrano-cinca, C., & Fuertes-calle, Y. (2005). Measuring DEA efficiency in Internet companies. *Decision Support Systems, 38*, 557–573.

Tovar, B., Javier Ramos-Real, F., & de Almeida, E. F. (2011). Firm size and productivity. Evidence from the electricity distribution industry in Brazil. *Energy Policy, 39*, 826–833.

Tseng, L.-Y., Chin, Y.-H., & Wang, S.-C. (2009). A minimized makespan scheduler with multiple factors for grid computing systems. *Expert Systems with Applications, 36*, 11118–11130.

Vaninsky, A. (2006). Efficiency of electric power generation in the United States: Analysis and forecast based on data envelopment analysis. *Energy Economics, 28*, 326–338.

Weyman-Jones, T. (1991). Productive efficiency in a regulated industry: The area electricity boards of England and Wales. *Energy Economics*, 116–122.