Teaching data envelopment analysis in production operations management through an undergraduate research project based on real-world data

Xiangrong Liu

Department of Management,
Ricciardi College of Business,
Bridgewater State University,
Harrington Hall, 95 Grove Street,
Bridgewater, MA 02324, USA
Email: xliu@bridgew.edu

Abstract: This teaching note reflects the recent experience in teaching an undergraduate production operations management course through the implementation of a real-world-problem-based research project. In this project, students investigated a multi-million dollar effort *Access* database compiled by the Environmental Protection Agency (EPA) and applied a data envelopment analysis (DEA) model to benchmark production efficiency among the coal-fired power plants within the states that they selected in the USA. This research project did much to developing students' quantitative problem-solving skills by weaving the research project throughout the whole semester, engaging students in illustrating DEA models with a real industrial production problem, and enhancing students' awareness of sustainability in production operations management.

Keywords: data envelopment analysis; DEA; linear programming; production operations management; power generation; pollution emission; undergraduate research project; real-world data; sustainability; operations management education.

Reference to this paper should be made as follows: Liu, X. (2015) 'Teaching data envelopment analysis in production operations management through an undergraduate research project based on real-world data', *Int. J. Information and Operations Management Education*, Vol. 6, No. 1, pp.14–23.

Biographical notes: Xiangrong Liu is currently an Assistant Professor of Operations Management at Bridgewater State University, MA, USA. She earned her PhD in Business Administration from Drexel University, PA, USA. In addition, she received a Certified Supply Chain Professional (CSCP) certification from APICS, the Association for Operations Management. Her research interests focus on supply chain management modelling revolving around sustainability issues such as remanufacturing, pollution-emission controls, and renewable energy. She has contributed a book chapter and published several journals including *Information Systems Education Journal* and *Council Undergraduate Research Quarterly*.

This paper is a revised and expanded version of a paper entitled 'An undergraduate research project in operations management – a real world application of data envelopment analysis to benchmark power production efficiency' presented at Northeast Decision Sciences Institute 2014 Annual Conference, Philadelphia, PA, 26–29 March 2014.

Copyright © 2015 Inderscience Enterprises Ltd.

1 Introduction

Production operations management is an advanced course for operations management concentrations/majors in almost all business schools. Presumed knowledge, such as forecasting, scheduling, quality management and process management are usually covered in a prerequisite fundamental operations management course, while the focus of this production operations management course is the concepts, practices and analysis of production processes in the manufacturing system. The general contents include aggregate planning, capacity planning, material resource planning (MRP) and enterprise resource planning (ERP), lean production and just-in-time, and inventory management. To deeply understand this content, students require strong analytical skills and a solid mathematics background, especially in the management sciences and operations research field. One typical example of these analytical applications is to use linear programming models to find the best mix of products. Such quantitative modelling tools are essential for students to master and learn to apply so that they can conduct analysis in their careers. Unfortunately, due to the lack of interest and less than strong math skills among students, professors are afraid of including the operations research/management science course in the current curriculum in most liberal arts schools. Further, it is unlikely that students will get this training outside of such a course. Due to the importance of operations research and mathematical modelling techniques to students' success, there is a great need to include these techniques, particularly the basic linear programming methods in this production operations management course.

Furthermore, how to enhance students' critical thinking and improve their analytical skills through teaching quantitative modelling in a production operations management course is the next question. Business education includes two major components: a classroom component and a student project component (Aserkar, 2013). The effectiveness of the classroom component has traditionally relied heavily on the delivery of concepts, models, and content by the instructor. In his empirical study on the impact of seminar characteristics, mentor characteristics, and students' perception of mentors measure by average grade, median grade and passing percentage, Steenhuis (2011) demonstrated that a solely lecture format did little to enhance student critical thinking and analytical skills. In other words, the traditional lecture sometime fails to stimulate students' interests and get students engaged. Gnanlet and Khanin (2013) state the best way to avoid the 'competence traps' due to lack of critical-thinking and problem-solving skills is to get them adjusted to specific knowledge domains.

The second major component, the student project, is the experiential learning component, the goal of which is to extend theory into reality, from textbook concepts to real life organisation applications. However, the student project component sometimes deviates from the classroom component, or the classroom component is unable to provide students with sufficient technical support for the student project component (Abdelhakim et al., 2012). Therefore, integrating lecture and project components into a single-semester undergraduate course could minimise the weakness of using each component individually. However, this integration requires careful design and implementation, especially the selection of the topic, itself crucial to the achievement of the goal. For instance, the focus of this study is on a topic about the analysis of production efficiency of coal-fired power plants, which has recently received much attention. According to EPA data, more than 58 coal-burning power plants partially or entirely shut down after Obama's regulation and the EPA's first established mercury rule in 2012 (Martinson, 2015). The regulation and ruling have had significant impact on many states' economic development, particularly on those states that depend heavily on coal, such as West Virginia.

To put it in context, coal-fired power plants are widely spread over the world. According to 2013 data, 39.1% of electricity was generated by coal in the USA (EIA, 2014). Because of the popularity of coal-fired power plants, this was not an unfamiliar topic to students. With little research required to fully understand the production process of electricity generation through burning coal, students were able to associate this production process with the basic concepts of operations, production operations, and operations management covered at the beginning of the course. Students were easily able to connect operation in coal-fired power plants with the other operations as the transformation processes from inputs to outputs. Identifying inputs and outputs of this particular process required a close look by students in this research project.

Students became further attracted by this topic because of the decades-old debate over environmental issues associated with coal-fired power plants and especially the recent concerns with the pollution emissions stimulated by the release of the new limits for hazardous air pollution emissions established in 2012. As known to those in the field, the production process through coal-fired power plants generates tremendous amounts of emissions. For instance, fossil fuel-fired power plants account for 70% of the nation's SO_2 emissions, 13% of NO emissions, and 40% of man-made CO_2 emissions (EPA, 2014a). In addition, the combustion of coal contributes to smog, acid rain, and haze. Public health is under high risk if the power generated from coal-fired power plants continues to increase. The negative impact of coal-fired power plants is tremendous, which makes the evaluation of the efficiency of electricity production crucial, especially when facing the decision of shutting down coal-fired power. This is, therefore, a fitting problem for students to tackle due to the complexity which requires the appreciation of real critical thinking in the industry. Like stakeholders, students are eager to answer key questions: for each state, which coal fire power plants could be potentially shut down due to low efficiency? How could these inefficient coal-fired power plants avoid being shut down by benchmarking the best (efficient) ones to improve their efficiency?

Naturally, students' training in business automatically leads them to evaluate the performance of the plants by calculating profits or costs. Since pollution emissions cannot simply be assessed in terms of a dollar value, however, students have to employ non-parametric methods such as data envelopment analysis (DEA) (Charnes et al., 1978) to conduct comparisons. The basic idea in a DEA model comes from the calculation of multifactor productivity, which is the ratio of several outputs over several inputs. However, in order to compare the performance of a group of decision making units (DMU), DEA is a series optimisation model: Each optimisation model selects the best weights for each individual DMU to get the maximum efficiency ratio (weighted outputs over weighted inputs). At this best scenario, the optimal efficiency ratio can get to 1, indicating that this DMU is efficient; otherwise, it is not. All the efficient DMUs form a frontier, which the other DMUs can benchmark. Converting a DEA model to a linear programming model provides students an easier way to solve the problem and allows for more chances to practice linear programming. Although the modelling of DEA is usually deemed to be complicated and abstract, students in this introductory undergraduate research project can utilise the model in the industrial setting by themselves and explore how essential these analytical tools are in terms of assisting decision-making in the real business world.

This paper showcases the systematic and seamless integration of the student project component into the production operations management course. Firstly, students studied the production process of coal-fired power generation. Based on that, they applied DEA models to compare the efficiency of those power plants in a particular state by taking pollution emissions into consideration. After the comparison, students selected and recommended the best performer(s) for the other coal-fired power plants to benchmark. In the end, students studied the current policy and regulations in the state they selected.

The main purposes of introducing this undergraduate research project include:

1 Helping students analyse the traditional operation process and its implication in a real business world through analysing inputs and outputs.

2 Helping students understand the evaluation of efficiency of operations process and DEA modelling.

3 Helping students utilise fundamental linear programming models and advanced ones such as DEA models and solving the problems using Excel.

4 Helping students realise the sustainability issues involved in this problem. Not only can students simply realise the pollution emissions generated from this process and evaluate the impact of these emissions, they also need to incorporate them into the modelling, decision-making and policy-making.

5 Helping students experience handling a large database, searching and evaluating the other data sources, and more generally, becoming familiar with the complexity of a large database.

The implementation of the project and experiences gained from the project are discussed in the following sections. Section 2 introduces the primary dataset; Section 3 explains the class setting and the implementation of the research project; and Section 4 concludes the paper with the feedback from students and the discussion of limitations.

2 Data

The database that students used to conduct analysis was originally compiled by the U.S. Environmental Protection Agency (EPA) (http://www.epa.gov/ttn/atw/utility/eu_icr_parti_partii.mdb). Based on this multi-million dollar database, EPA established the first ever national standards limiting emissions of hazardous air pollutants, such as mercury from coal and oil-fired power plants in 2012. More than an estimated 1,400 coal and oil-fired electric generating units (EGUs) at 600 power plants were impacted by these standards (EPA, 2014b).

This large database provides 40 data tables detailing the features of all coal-fired power plant components such as boilers, controls, and ports, as well as multiple sample data for the emissions collected from different ports between 2004 and 2006 in the USA. The database is extremely valuable in terms of the large amounts of information, which also deems it highly complex. Understanding the structure of the database and the relationship among different data tables is the first step before students can extract the information. It requires students to apply the knowledge of *Access* learned in the course of information system and computer applications: introduction course. This is a valuable link across courses that helped students gain competence in using the *Access* database, which serves as an important tool to store business data in the professional world.

Additionally, and perhaps at the outset somewhat counter-intuitively, the ‘messiness’ of the database allowed students a valuable learning experience. Students needed to sift through the data, clean it up and deal with issues such as missing data, inconsistent measurements, or different levels of data (aggregate or individual). For example, the pollution level of a certain region could be reported at different time periods in a year, either for different generation units or different boilers. Students had to pick a consistent measurement when they compared multiple plants in a state. Further, students needed to handle missing data in such a complicated database, either ignoring the data record completely or replacing the data with the average number from different generator units at the same level. For details of how to deal with data, students to the teaching case designed by Donalds and Liu (2014).

Despite this disarray, students felt very excited about this opportunity of using real data. Through trial-and-error, referring to scholarly papers, Q&A in the classroom, and group discussion, students utilised multiple statistical tools, such as data averaging and outlier identifying to clean and interpret data. For example, from the simple statistical analysis, the group working on the project of Pennsylvania found out that the major dominating coal type used is coal refuse, as a lower rank coal, which determines a lower efficiency in the end.

Although the EPA database provided lots of information, some students might design models that need other input/output variables. As a result, students had to explore and gather secondary data, such as the industry and trend data to support their research. This process provided some opportunities for students to develop critical thinking skills and the capability of synthesising information because students needed to assess the credibility of sources and the validity of data.

3 Implementation

The whole project was scaffold into three phases. At the beginning of the semester, a Powerpoint presentation slide showed the main tasks, expectations and deadlines for each phase so that it was easier to guide students through the research step by step. A detailed timeline is included in Table 1.

Table 1 Timeline

Phase	*Week*	*Classroom component*	*Experiential component*
Phase I	Week 1	Syllabus	Project introduction/group formation
	Week 2	Introduction to POM/ productivity/efficiency	
Phase II	Week 3–4	Linear programming	Project draft 1: Background introduction, descriptive statistics
	Week 5–7	DEA model	Workshop of literature searching
			Project draft 2: Input and output variables analysis, literature review

Table 1 Timeline (continued)

Phase	*Week*	*Classroom component*	*Experiential component*
Phase III	Week 8		Presentation at the university's sustainability day
	Week 9	JIT and lean operations	
	Week 10	Aggregate planning	
	Week 11–12	Capacity planning	Project draft 3: Models and results
	Week 13–14	Material requirement planning	
	Week 15	Review	Presentation at undergraduate research symposium
	Week 16	Final exam	Final paper due

3.1 Phase I

Phase I began with the introduction of the research project to students. According to the EPA Database, ten states with the largest number of coal-fired power plants (Georgia, Indiana, Michigan, Missouri, North Carolina, Ohio, Pennsylvania, Tennessee, Utah, and Virginia) were identified. Students were free to choose to work individually or collaboratively in a group of two. Each group only chose one state that they were interested in investigating. Along with some basic concepts of operations and the way of calculating productivity covered in the lectures, students started to discover the basic background of the industry, the production process, and the current regulations of emissions in the specific state they select.

3.2 Phase II

The classroom components in Phase II focused on linear programming and DEA modelling, which serves as the mathematical foundations for this research project and paves the road for students to conduct analysis. In the research part, students identified input and output variables in the project by themselves and provided the rationales for their decisions. Correspondingly, students needed to extract the dataset from the database, or locate the other sources for additional data if needed. Additionally, a librarian and I organised a workshop so that students would be able to learn how to conduct a literature survey. For many students, this was their first experience with academic databases. The literature survey helped students not only come into contact with cutting-edge research, but also support the innovativeness and originality of their own research.

3.3 Phase III

Phase III allowed students one and half months to model the data, analyse them, and complete the research project. The lecture component moved to cover the other traditional materials in this course (capacity planning, MRP and inventory management). Some class time made time for Q&A to discuss difficulties and brainstorm solutions. Students were able to learn from the instructor and peer students how to set up and solve their models. At the same time, they still had to complete their work independently since coal-fired power plants in different states have different datasets and face different economic, social and cultural environments.

3.4 The deliverables

The deliverables required were two presentations and one final research paper:

1 A short group presentation required in the middle of the semester with only the introduction completed. The presentation took place on the university's sustainability day. Students demonstrated the involvement in sustainability at the college of business.

2 A final project group presentation at the end of the semester took place at the university's mid-year undergraduate research symposium. As a high impact practice, participating in the undergraduate research symposium provided students a chance to exchange ideas and observe the research from other colleges (Shanahan et al., 2015).

3 A final paper was due by the end of the semester. Therefore, students had some time to incorporate the feedback from the symposium into their paper.

4 Feedback

At the end of the semester, students filled out a survey that had ten questions and a Likert scale of options for each. Thirteen out of 25 students anonymously provided feedback. Questions asked about the effectiveness of this research project with respect to goal achievement (questions 2, 3, 6 and 8), career motivation (questions 4 and 7), research preparation (questions 5 and 9), appreciation of complexity (question 1) and the power plant topic (question 10).

The summary of the class feedback is sorted by average score and listed in Table 2. Almost all the students (92%) confirmed that they gained a better understanding of production efficiency (question 2) and no less than 77% students confirmed that they had a better understanding of DEA models (question 3). Students did not realise how important the database skills (MS Access) were in terms of contribution to their future career (question 7). The lack of motivation potentially led to the widely distributed competence level of using MS Access (question 6). Instead the agreement on the contribution of mathematical modelling skills towards to career development was much higher at 77 % (question 4). Through the whole semester project, students were confident with their skills in conducting undergraduate research in the Operations Management field (77% in question 5). Students gained experience in teamwork and collaboration while dealing with unstructured problems (questions 8 and 9). However, with the challenges of modelling real world decision problems (indicated in question 1), it is not surprising that the attitude towards the topic of power plants varied among students (question 10), a finding mirrored in Beliën et al.'s (2013) study in which the researchers attempted to implement a game to teach integer programming.

Table 2 Students' evaluation

		Strongly agree (5)	*Agree (4)*	*Neither agree nor disagree (3)*	*Disagree (2)*	*Strongly disagree (1)*	*Weighted score*
1	I have a better appreciation of the challenges of modelling real world decision problems.	31%	54%	15%	0%	0%	4.15
2	I have a better understanding of production efficiency.	23%	69%	8%	0%	0%	4.15
3	I have a better understanding of DEA models.	23%	54%	23%	0%	0%	4.00
4	Mathematical modelling skills will be useful for my future career.	23%	54%	23%	0%	0%	4.00
5	I am better prepared to conduct operations management research.	15%	62%	23%	0%	0%	3.92
6	I gained competence using MS Access software.	31%	38%	23%	0%	8%	3.85
7	Database skills will be useful for my future career.	31%	38%	15%	8%	8%	3.77
8	I gained valuable experience working in a team.	15%	38%	46%	0%	0%	3.69
9	I gained valuable experience conducting research that involved an unstructured problem.	8%	54%	38%	0%	0%	3.69
10	I liked the research topic of power plant efficiency.	0%	46%	46%	8%	0%	3.38

In the end, students felt proud of themselves when they were able to talk about the coal-fired power plants production process and the environmental impact of the plants to a wider public audience. For the state they selected, students were able to list all the coal-fired power plants, the history of their development and the special regulations. They had complete ownership over their work. They also felt comfortable about their ability to provide reasonable solutions based on the abstract models they developed. Finally they commented that the course was 'the most difficult one' they had seen in their whole education in business school. The biggest barrier was the fact that there was not a single straightforward right answer to the research problem. The ambiguity and uncertainty in this type of real world scenario is different from the neat structure of most textbook exercise problems and pre-scripted case studies. Additionally, students faced the following other challenges:

1 understanding the large database of world data and collecting data from a secondary source

2 translating complicated and unstructured managerial problems into abstract DEA models

3 dealing with technical modelling issues, such as including non-desirable outputs, which is not covered in the textbook and is an excellent extension of the traditional model.

Because most students in this class were business students without too much engineering background, a sharp learning curve was expected at the beginning of the research project, similar to other industry projects as mentioned in the literature (Hillon et al., 2012; Sawhney et al., 2013). With sufficient preparation to bridge the gap between theory and real problems and continuous support, students were able to overcome these difficulties.

This type of success has also been mentioned in Torres and Prybutok's (2014) study, where undergraduate research projects required students to develop an industry analysis of an economic sector of Mexico City.

Due to the small class and limited number of the students taking the evaluation in the course, the findings might be different if the same project is implemented in different production operations management classes with a larger class size. Therefore, it would be very interesting to compare the results between the replicated one and the current one.

References

Abdelhakim, M.N.A., Carmichael, J.N. and Ahmad, S.S. (2012) 'Quality evaluation of university web portals: a student perspective', *International Journal of Information and Operations Management Education*, Vol. 4, No. 3, pp.229–243.

Aserkar, R. (2013) 'Learning through industry interface projects', *INFORMS Transactions on Education*, Vol. 13, No. 2, pp.93–101.

Beliën, J., Colpaert, J., Boeck, L.D., Eyckmans, J. and Leirens, W. (2013) 'Teaching integer programming starting from an energy supply game', *INFORMS Transactions on Education*, Vol. 13, No. 3, pp.129–137.

Charnes, A., Cooper, W. and Rhodes, E. (1978) 'Measuring the efficiency of decision making units', *European Journal of Operational Research*, Vol. 2, No. 6, pp.429–444.

Donalds, K. and Liu, X. (2014) 'Cleaning data helps clean the air', *Information Systems Education Journal*, Vol. 12, No. 5, pp.4–15.

Gnanlet, A. and Khanin, D. (2013) 'Problem-solving and incompetence traps in service operations course', *International Journal of Information and Operations Management Education*, Vol. 5, No. 3, pp.276–295.

Hillon, M.E., Cai-Hillon, Y. and Brammer, D. (2012) 'A brief guide to student projects with industry', *INFORMS Transactions on Education*, Vol. 13, No. 1, pp.10–16.

Martinson, E. (2015) *The Fall of Coal*, Politico [online] http://www.politico.com/story/2015/04/coal-power-plants-epa-regulations-117011.html (accessed 19 April 2015).

Sawhney, R., Maleki, S., Wilck, J. and Hashemian, P. (2013) 'Center for Productivity Innovation's student project with industry program at the University of Tennessee, Department of Industrial and Systems Engineering', *INFORMS Transactions on Education*, Vol. 13, No. 2, pp.83–92.

Shanahan, J., Liu, X., Manak, J., Miller, S., Tan, J. and Yu, C. (2015) 'Research-informed practice, practice-informed research: the integral role of undergraduate research in professional disciplines', *Council Undergraduate Research Quarterly*, Vol. 35, No. 4, pp.6–16.

Steenhuis, H. (2011) 'Mentors, their preparation and their effectiveness in an undergraduate operations management course', *International Journal of Information and Operations Management Education*, Vol.4, No. 1, pp.1–21.

Torres, J.L.N. and Prybutok, V.R. (2014) 'Using undergraduate student research as a pedagogical tool in a predominately undergraduate institution in Mexico', *International Journal of Information and Operations Management Education*, Vol. 5, No. 4, pp.311–323.

U.S. Energy Information Agency (EIA) (2014) *Short-Term Energy Outlook* [online] http://www.eia.gov/forecasts/steo/report/electricity.cfm (accessed 30 December 2014).

U.S. Environmental Protection Agency (EPA) (2014a) *Air Emissions* [online] http://www.epa.gov/cleanenergy/energy-and-you/affect/air-emissions.html (accessed 30 December 2014).

U.S. Environmental Protection Agency (EPA) (2014b) *Clean Power Plants* [online] http://www.epa.gov/mats/powerplants.html (accessed 30 December 2014).