Abstract: This work presents a new optimization software library, which includes a number of Data Envelopment Analysis (DEA) models. All the DEA models have been implemented in the AMPL algebraic modeling language for mathematical programming and solved using either Gurobi Optimizer v5.5 or MINOS v5.51 (for those models having general nonlinear objectives). In order to verify the correctness of the results, real world datasets have been examined and solved using the well-known commercial DEA Solver Pro v8.0c by SAITECH, Inc. We also present a literature review of both commercial and non-commercial DEA software packages and discuss about the benefits and drawbacks of our proposed DEA model library.

Keywords: Optimization Software, AMPL, Data Envelopment Analysis

1. INTRODUCTION

DEA constitutes a non-parametric approach based on linear optimization models and methods, in order to construct a non-parametric surface over the available data. Performance - or efficiency - measurement of each DMU can be computed, regarding its position on this surface. DEA evaluates the relative performance of Decision Making Units (DMUs) (e.g., universities, hospitals) and can easily handle multiple input and output of each DMU. The efficient DMUs define an efficient production possibility surface or frontier.

A large number of DEA models has been reported in the scientific literature during the last 30 years [6]. DEA models have been applied for the measurement of the relative performance in many situations, such as in the food production industry [17], e-health providers [18], and many more diverse types of DMUs. The majority of these DEA models belong into four main categories: i) radial, ii) non-radial and oriented, iii) non-radial and non-oriented, and iv) composite of radial and non-radial. The classification between radial and non-radial models depends whether we are interested in a proportionate change of input/output or not. The classification between oriented and non-oriented models depends whether we are mainly interested rather in input reduction or output expansion, or not (e.g., input reduction and output expansion at the same time).

The proposed DEA software package provides a number of DEA models. Apart from the basic DEA models either with constant returns-to-scale (CCR) or variable returns-to-scale (BCC) characteristics, our DEA model library features implementations of DEA models with increasing (IRS) or decreasing (DRS) returns-to-scale characteristics. There are also included other implementations of DEA models, e.g., to handle slacks-based measure (SBM) of efficiency. Moreover, the proposed DEA software package may also solve either radial or non-radial super efficiency models. Furthermore, the software package can solve either input or output oriented versions of the above DEA models.

Section 2 presents a literature review concerning either commercial or non-commercial DEA software packages as also the motivation of this research work. Following, in section 3 the mathematical formulations of all the DEA models that are currently implemented in our optimization software library are presented. It should be noted that, the formulations of these DEA models have the same notation as in [7]. Implementation details and an example of a DEA model in AMPL is given in section 4. Finally, in section 5 we discuss about some future research directions.

2. LITERATURE REVIEW AND MOTIVATION

2.1. Commercial DEA software packages

There are several commercial software packages for the solution of Data Envelopment Analysis problems. Apart from their price, all these DEA software packages mainly differ on the number of DEA implemented models, in the variety of the produced reports, etc. A list of some of the state-of-the-art commercial DEA software packages is briefly presented in this subsection. Although, this list is no extensive, it describes some of the
most well known DEA applications. For a recent, more analytical, review of DEA software packages, both commercial and non-commercial, the readers can refer to the work by Barr in [4].

DEA Solver Pro by SAITECH Inc.\textsuperscript{1} is a user-friendly, spreadsheet based, DEA software package which includes a large number of popular DEA models (e.g., dynamic DEA, network DEA). A detailed report of the features of the latest version of DEA Solver Pro can be found in [15]. Furthermore, KonSi DEA Analysis\textsuperscript{2} is also a well-known DEA software package which handles up to 250 input and output variables and produces a large number of reports including efficiency scores table, histogram, graph, and others. KonSi DEA 3D Visualization is a tool that provides 3D visualization of efficient frontier for CRS and VRS models with one input and two output parameters.

PIM-DEASOFT\textsuperscript{3} is a DEA software package developed under supervision of Dr. A. Emrouznejad and Prof. E. Thanassoulis, which has several useful features such as extensive data handling facilities for very large data sets, ability to assess productivity change over time, graphical output facilities, and also permits data importing and exporting to Excel. Banxia Frontier Analyst\textsuperscript{4} provide several DEA models with “undesirable”, categorical and other types of variables, and also it allows user to undertake Malmquist productivity analysis. Moreover this application features a fully threaded calculation engine in order to be multi-processor capable and thus split an analysis into multiple parts.

DEA Frontier\textsuperscript{5} developed by Prof. J. Zhu is a spreadsheet add-in for the solution of DEA models. This application includes, among others, many of the standard DEA approaches such as cross efficiency, bootstrapping, two-stage network, or DEA-based supply chain model. Roughly speaking, this application solves more than 150 different DEA models and it is described in [23]. Finally, DEA Online Software (DEAOS.com\textsuperscript{6}) is web-based software provided by Behin-Cara Co. Ltd. This service is capable for the solution of either single periodic or multiple periodic models, for a maximum of 500 units and unlimited indices in every project.

2.2. Non-commercial DEA software packages

Apart from the previously mentioned commercial (and usually expensive) DEA applications, several research efforts have been also made in order to provide free, quality, DEA software packages. The majority of these DEA software packages is built using a general-purpose optimization modeling language (either open source or not). However, there is no extra cost in order to obtain these software packages by the respective developers. Such general-purpose optimization modeling languages provide user with flexibility to modify the models and easily develop new, more complicated DEA models. Furthermore, a user is also able to select an open source solver in conjunction with the modeling language.

FEAR\textsuperscript{7} (Frontier Efficiency Analysis with R), is a free DEA software library which can be linked to the general-purpose statistical package R, and was developed by Prof. P. W. Wilson. In order to use FEAR, one should first download and install the R language. The set of included routines are highly flexible, thus allowing measurement of technical, allocative, and overall efficiency. Furthermore, there are publicly available versions for either Linux, Mac OS, or Windows systems. A short paper introducing FEAR can be found in [22].

The modeling abilities of the General Algebraic Modeling System (GAMS) have been already utilized by several authors, e.g., [14, 9, 12], in order to provide DEA software implementations\textsuperscript{8}. Moreover, Green [11] presented an implementation of the Charnes-Cooper-Rhodes [5] based on a Modeling Language for Mathematical Programming (AMPL). Also, Emrouznejad had presented the implementation of DEA models using macros of the SAS programming language [8]. Kleine and Sebastian have also developed a free web-based DEA-tool\textsuperscript{9}. This tool allows user to process various DEA models such as additive models, range adjusted models, super-efficiency models, Russell-Measure, and others [13]. Open Source DEA (OSDEA\textsuperscript{10}) consists a free and open source DEA solver library that can be used in java programs. User can write their own models, by modifying the java code of the OSDEA Library.

\textsuperscript{1} http://www.saitech-inc.com/Products/Prod-DSP.asp
\textsuperscript{2} http://www.dea-analysis.com
\textsuperscript{3} http://www.deasoftware.co.uk
\textsuperscript{4} http://www.banxia.com/frontier
\textsuperscript{5} http://www.deafrontier.net/software.html
\textsuperscript{6} https://www.deaos.com
\textsuperscript{7} http://www.clemson.edu/economics/faculty/wilson/Software/FEAR/fear.html
\textsuperscript{8} http://www.gams.com/modlib/libhtml/dea.htm
\textsuperscript{9} http://www.dea.uni-hohenheim.de
\textsuperscript{10} http://www.opensourcedea.org
2.3. Motivation

Although there is a plethora of successful commercial DEA software packages, we believe that more research efforts are still required in order to develop more complete non-commercial software packages. Thus, our goal was to develop a library with a large number of state-of-the-art DEA models. Since the proposed DEA software library is built using a general-purpose optimization modeling language such as AMPL, researchers will find it convenient that with little effort, modifications can be easily made and provide new DEA models. The benefits of the proposed DEA software library are first that it utilizes the rich features of the AMPL modeling language, second that the code of the DEA models can be easily extended, and third that the majority of the Operational Research (OR) scientists are more familiar to general-purpose optimization modeling languages, rather than other programming languages. Finally, the variety of the DEA models that are already implemented also consists a strong point of the proposed DEA software package.

3. DEA MODELS

3.1. Input-oriented Charnes-Cooper-Rhodes model

This is the input-oriented version (dual problem) of the most basic DEA model, presented in [5]:

\[
\begin{align*}
\min \ z &= \theta \\
\theta, \lambda \in & \\
\text{s.t.} \quad & \theta x_o - X \lambda \geq 0 \\
& Y \lambda \geq y_o \\
& \lambda \geq 0 \\
\end{align*}
\]

(\text{CCR - I})

The efficiency score is denoted by a scalar $\theta$.

3.2. Output-oriented Charnes-Cooper-Rhodes model

The output-oriented version, is formulated as follows:

\[
\begin{align*}
\max \ z &= \eta \\
\eta, \mu \in & \\
\text{s.t.} \quad & x_o - X \mu \geq 0 \\
& \eta y_o - Y \mu \leq 0 \\
& \mu \geq 0 \\
\end{align*}
\]

(\text{CCR - O})

3.3. Input-oriented Banker-Charnes-Cooper model

The input-oriented version of the classic DEA model, presented in [2], differs from the CCR-I model only in the addition of the convexity condition ($e$ is the unity vector):

\[
\begin{align*}
\min \ z &= \theta \\
\theta, \lambda \in & \\
\text{s.t.} \quad & \theta x_o - X \lambda \geq 0 \\
& Y \lambda \geq y_o \\
& e \lambda = 1 \\
& \lambda \geq 0 \\
\end{align*}
\]

(\text{BCC - I})

3.4. Output-oriented Banker-Charnes-Cooper model

The output-oriented version, is formulated as follows:

\[
\begin{align*}
\max \ z &= \eta \\
\eta, \lambda \in & \\
\text{s.t.} \quad & X \lambda \leq x_o \\
& \eta y_o - Y \lambda \leq 0 \\
& e \lambda = 1 \\
& \lambda \geq 0 \\
\end{align*}
\]

(\text{BCC - O})
3.5. Input-oriented Slacks-based measure model

The formulation of the input-oriented Slacks-based measure model, presented in [19], is as follows:

\[
\begin{align*}
\min_{\lambda, s^-} & \quad 1 - \frac{1}{m} \sum_{i=1}^{m} \frac{s^-_i}{x_{io}} \\
\text{s.t.} & \quad x_o = X\lambda + s^-_i \\
& \quad y_o \leq Y\lambda \\
& \quad \lambda, s^- \geq 0
\end{align*}
\]

(SBM - I - C)

3.6. Output-oriented Slacks-based measure model

The formulation of the output-oriented Slacks-based measure model, is as follows:

\[
\begin{align*}
\min_{\lambda, s^+} & \quad 1 + \frac{1}{m} \sum_{i=1}^{m} \frac{s^+_i}{x_{io}} \\
\text{s.t.} & \quad x_o \geq X\lambda \\
& \quad y_o \geq Y\lambda - s^+ \\
& \quad \lambda, s^+ \geq 0
\end{align*}
\]

(SBM - O - C)

3.7. Free Disposal Hull model

The Free Disposal Hull (FDH) model described in [21], aims to evaluate the DMUs based only on their actually observed input and output values. A mixed-integer programming formulation of the FDH model follows:

\[
\begin{align*}
\min_{\theta, \lambda} & \quad \theta \\
\text{s.t.} & \quad \theta x_o - X\lambda \geq 0 \\
& \quad y_o - Y\lambda \leq 0 \\
& \quad e\lambda = 1 \\
& \quad \lambda_j \in \{0, 1\}
\end{align*}
\]

(FDH)

3.8. Input-oriented increasing returns-to-scale model

The formulation of the input-oriented increasing returns-to-scale (IRS) model, described in [3], is as follows:

\[
\begin{align*}
\min_{\theta, \lambda} & \quad \theta \\
\text{s.t.} & \quad \theta x_o - X\lambda \geq 0 \\
& \quad Y\lambda \geq y_o \\
& \quad e\lambda \geq 1 \\
& \quad \lambda \geq 0
\end{align*}
\]

(IRS - I)

3.9. Output-oriented increasing returns-to-scale model

The formulation of the output-oriented increasing returns-to-scale (IRS) model, is as follows:

\[
\begin{align*}
\max_{\eta, \lambda} & \quad \eta \\
\text{s.t.} & \quad X\lambda \leq x_o \\
& \quad \eta y_o - Y\lambda \leq 0 \\
& \quad e\lambda \geq 1 \\
& \quad \lambda \geq 0
\end{align*}
\]

(IRS - O)
3.10. Input-oriented decreasing returns-to-scale model

The formulation of the input-oriented decreasing returns-to-scale (DRS) model, described in [3], is as follows:

\[
\begin{align*}
\min_{\theta, \lambda} & \quad 0 \\
\text{s.t.} & \quad \theta x_0 - X\lambda \geq 0 \\
& \quad Y\lambda \geq y_o \\
& \quad 0 \leq e\lambda \leq 1 \\
& \quad \lambda \geq 0 \\
\end{align*}
\]  

(\text{DRS} - I)

3.11. Output-oriented decreasing returns-to-scale model

The formulation of the output-oriented decreasing returns-to-scale (DRS) model, is as follows:

\[
\begin{align*}
\max_{\eta, \lambda} & \quad \eta \\
\text{s.t.} & \quad X\lambda \leq x_0 \\
& \quad \eta y_o - Y\lambda \leq 0 \\
& \quad 0 \leq e\lambda \leq 1 \\
& \quad \lambda \geq 0 \\
\end{align*}
\]  

(\text{DRS} - O)

3.12. Input-oriented non-radial super efficiency model

The aim of the “super efficiency” DEA models is to provide discriminating characteristics among DMUs, in situations where a significant number of DMUs are efficient. Moreover, super efficiency DEA models using slacks based measure have several benefits, such as unit invariant efficiency scores. The input-oriented non-radial super efficiency model, under the constant returns-to-scale assumption presented in [20], is as follows:

\[
\begin{align*}
\min_{\phi, \lambda} & \quad 1 + \frac{1}{m} \sum_{i=1}^{m} \phi_i \\
\text{s.t.} & \quad \sum_{j=1, j \neq o}^{n} x_{ij}\lambda_j - x_{io}\phi_i \leq x_{io} \quad (i = 1, \ldots, m) \\
& \quad \sum_{j=1, j \neq o}^{n} y_{rj}\lambda_j \geq y_{ro} \quad (r = 1, \ldots, s) \\
& \quad \phi_i \geq 0 (\forall i), \lambda_j \geq 0 (\forall j) \\
\end{align*}
\]  

(SuperSBM - I - C)

3.13. Output-oriented non-radial super efficiency model

The formulation of the output-oriented non-radial super efficiency model, under the constant returns-to-scale assumption, is as follows:

\[
\begin{align*}
\min_{\psi, \lambda} & \quad \frac{1}{1 - \frac{1}{s} \sum_{r=1}^{s} \psi_r} \\
\text{s.t.} & \quad \sum_{j=1, j \neq o}^{n} x_{ij}\lambda_j \leq x_{io} \quad (i = 1, \ldots, m) \\
& \quad \sum_{j=1, j \neq o}^{n} y_{rj}\lambda_j + y_{ro}\psi_r \geq y_{ro} \quad (r = 1, \ldots, s) \\
& \quad \psi_r \geq 0 (\forall r), \lambda_j \geq 0 (\forall j) \\
\end{align*}
\]  

(SuperSBM - O - C)


The two previous sub-sections can be extended for variable returns-to-scale models. The first super efficiency DEA model with radial characteristics was introduced in [1]. The formulation of the input-oriented radial super efficiency model, under the variable returns-to-scale assumption, is as follows:
\begin{align*}
\min z &= \theta \\
\text{s.t. } & \theta x_{jo} \geq \sum_{j=1, j \neq o}^{n} \lambda_j x_{ij} \ (i = 1, \ldots, m) \quad \text{(SuperRadial – I – V)}\\
& y_{jo} \leq \sum_{j=1, j \neq o}^{n} \lambda_j y_{rj} \ (r = 1, \ldots, s)\\
& \sum_{j=1, j \neq o}^{n} \lambda_j = 1 \\
& \lambda_j \geq 0 (\forall j)
\end{align*}

### 3.15. Output-oriented radial super efficiency model

The formulation of the output-oriented radial super efficiency model, under the variable returns-to-scale assumption, is as follows:

\begin{align*}
\min z &= \frac{1}{\eta} \\
\text{s.t. } & x_{jo} \geq \sum_{j=1, j \neq o}^{n} \lambda_j x_{ij} \ (i = 1, \ldots, m) \quad \text{(SuperRadial – O – V)}\\
& \eta y_{jo} \leq \sum_{j=1, j \neq o}^{n} \lambda_j y_{rj} \ (r = 1, \ldots, s)\\
& \sum_{j=1, j \neq o}^{n} \lambda_j = 1 \\
& \lambda_j \geq 0 (\forall j)
\end{align*}

### 4. AMPL IMPLEMENTATION DETAILS

The proposed DEA software package features a variety of DEA models, implemented in AMPL modeling language as separate files with extension “\text{.mod}”. All the required data are separated from the corresponding model files and need to be stored in a spreadsheet file. Since a ranking of the DEA efficiency scores of \(k\) DMUs requires the solution of \(k\) Linear Problems (LPs), it is much more convenient to use AMPL script files. These sophisticated script files apart from utilizing the “model” and “data” commands in order to read the files having the model declarations and the data statements respectively, they can also run repetitive AMPL commands. Furthermore, these script files have extension “\text{.run}” and usually contain a variety of options, (e.g., selection of appropriate solver, precision, rounding) and/or contain the required commands for producing formatted solution reports (e.g., slack variables, \(\theta\), \(\lambda\) values).

Initially, the user needs to fill the data into a pre-formatted spreadsheet file, as it is depicted in figure 1. This spreadsheet file has a few pre-defined named ranges, using the “name manager” facility of MS Excel. The scope of these named ranges is to store the data of the AMPL set regarding the \(n\) DMUs, the matrices containing the input \(X\) and output \(Y\) variables, as also their cardinality numbers \(m\) and \(s\) respectively. These named ranges allow the interoperability of AMPL with the relational database tables in MS Excel, using the Microsoft Open Database Connectivity (ODBC) driver. More detailed instructions about the communication of the handler supports packages via the ODBC driver can be found in the book by Fourer et al. [10].

An example of the source code of the input-oriented Slacks-based measure model file, presented in the subsection 3.5 is given below.

```AMPL
# Model file for the input-oriented Slacks-based measure model
#
# Set Definition
#
set DMU ordered;
#
# Parameter Definition
#
param m >= 0;
param s >= 0;
param X[DMU,1..m] >= 0;
```
param Y{DMU,1..s} >= 0;
param iDMU symbolic default first(DMU);

#-------------------------------------------------
# Variable Definition
var lamda{DMU} >= 0;
var s_minus{1..m} >= 0;
#-------------------------------------------------
# Objective Function
minimize efficiency : 1 - (1/m)*sum{i in 1..m}s_minus[i] / X[iDMU,i];
#-------------------------------------------------
# Constraints
subject to input_constraint{j in 1..m}:
X[iDMU,j] = sum {n in DMU} X[n,j] * lamda[n] + s_minus[j];
subject to DMU_constraint{j in 1..s}:
Y[iDMU,j] <= sum {n in DMU} Y[n,j] * lamda[n];

An example of the source code of the script file that i) reads the above model file and the data stored in a spreadsheet file named “data.xls”, ii) selects the Gurobi solver, and iii) displays the efficiency scores (objective values) and the decision variables (i.e., slack variables and $\lambda$ values) is given below. The execution of this script file will solve sequentially $k$ LPs (12 in our example).

# Script file for the input-oriented Slacks-based measure model
model SBM-I-C.mod;
table m IN "ODBC" "data.xls" : [], m;
read table m;
table s IN "ODBC" "data.xls" : [], s;
read table s;
table DMU IN "ODBC" "data.xls" "DMU" : DMU <- [DMU];
read table DMU;
table X IN "ODBC" "data.xls" :
The solution report contains information regarding the efficiency score, the reference set \((\lambda)\), and the projection for each DMU. A part of this report concerning the third (C) DMU of our example, is shown below.

Gurobi 5.5.0: optimal solution; objective 0.8521654412
4 simplex iterations
\(\lambda\) :=
A 0.0294118
B 1.04706
s_{−} :=
1 4.51765
2 18.3941

As it can be seen from the above results, the efficiency score of the C DMU was 0.8521654412. The reference set gives us the opportunity to identify the efficient DMUs and the extent to which they operate similarly to a given inefficient DMU. The peers of the C DMU are those efficient DMUs that define the facet of the frontier against which the (inefficient) C DMU is projected. Thus, the reference set of this DMU comprises DMUs A and B. Furthermore, the input excesses \(s_{−}\) are also displayed in the above results.

In order to verify the correctness of the results, real world datasets have been examined and solved using the well-known commercial DEA Solver Pro v8.0c by SAITECH, Inc. In all these cases, the same results have been obtained, except minor rounding differences.

5. CONCLUSION AND FUTURE WORK

It is well-known that algebraic modeling language are ideal tools for rapid prototyping and optimization model development [11]. Thus, our proposed DEA software package utilizes the flexibility and convenience of the AMPL modeling language. A strong point of the proposed work is the variety of the implemented DEA models either with constant returns-to-scale, variable returns-to-scale characteristics, increasing or decreasing returns-to-scale characteristics, and also their input or output oriented versions.

Since this is a work in progress, new models are still being added to the proposed library. A future research direction is to improve some implementation issues regarding, e.g., any imbalance in data magnitudes, treat problems with nonpositive values (for ratio models) and others. Such guidelines for the preparation of data for DEA problems can be found in [16]. Furthermore, the proposed DEA software library will be soon publicly available by the authors.

References


