1-1-2012

An Introduction to Technology Forecasting with a TFDEA Excel Add-in

Dong-Joon Lim
Portland State University

Timothy R. Anderson
Portland State University, tim.anderson@pdx.edu

Citation Details

An Introduction to Technology Forecasting with a TFDEA Excel Add-in

Dong-Joon Lim, Timothy R. Anderson

Dept. of Engineering and Technology Management, Portland State University, Portland, OR - USA

Abstract--**This paper describes an Excel add-in program that can run TFDEA (Technology Forecasting using Data Envelopment Anlaysis) within the spreadsheet. It utilizes freely available statistical software R and its packages developed by Statconn. This add-in allows access to both user-friendly tools for data manipulation in spreadsheet available and the power and precision of the results via R.**

I. INTRODUCTION

This guide introduces add-in software to run TFDEA in Microsoft Excel. As described in the original paper [1], TFDEA explains advancement of technology by capturing dynamics of efficiencies over time. Hence, it requires series of efficiency measurement based on DEA (Data Envelopment Analysis). Even if Microsoft Excel provides built-in optimization interface called 'Excel solver,' researchers who want to use it as an advanced optimization tools have found that it is inadequate either to solve large scale linear programming or automate with additional functions [2]. On the other hand, R is a powerful statistical tool extensible to various usages with superior numerical precision [3] however it is inconvenient to manipulate large data set especially for users who are not familiar with script based programming. TFDEA add-in software has been developed to bridge this gap by integrating user friendliness of spreadsheet with high computing capacity of specialized R packages and therefore provide convenient user interface for TFDEA.

II. FRAMEWORK

It may be helpful for users to have a general understanding on the flow of software in order to prevent error message beforehand. Fig. 1 illustrates this overall process where middleware is placed between Excel and R to make data set in the spreadsheet processed through DEA package. First of all, data to be analyzed should be placed on the spreadsheet in a typical table format where row indicates different data set and column contains different categories (attributes). Here, data sets that have been analyzed for previous publications applying TFDEA [4–7] from PSU (Portland State University) ETM (Engineering and Technology Management) department have been uploaded to the TFDEA database that they can be imported using dialogue box for revisit purpose. Next, users can select suitable parameters to run TFDEA by activating dialogue box (see section 3 for detailed information). R scripts for TFDEA have been compiled in the Excel macros that R is to be called when any calculation is needed. Lastly, when all calculations have been completed, results are parsed back through middleware to put them together into summarized form pre-designed in the Excel macro.
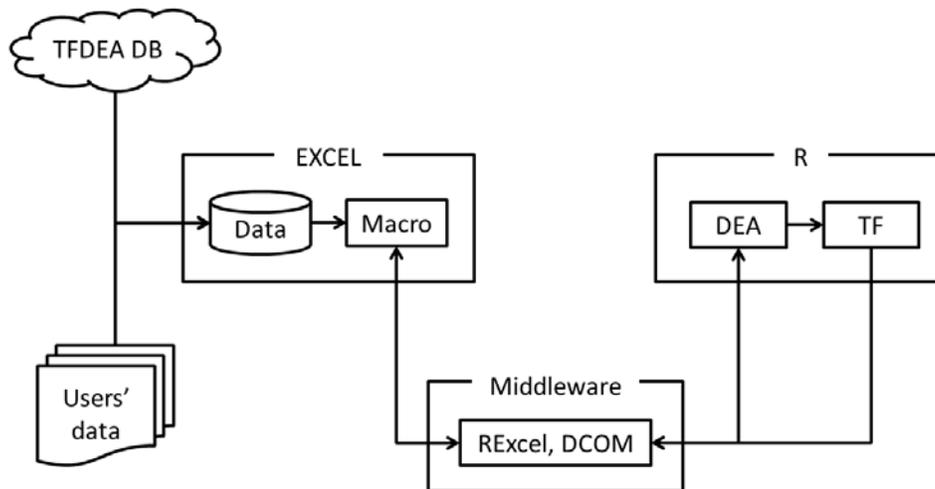

Figure 1 Overall process

III. INTERFACE

A. *Data process*

Data can be imported either from MySQL database or text files. Following Fig. 2 shows the dialogue box for MySQL importing function. Account information for TFDEA database is pre-coded to access previous published data set. Users can also connect other database to import their own data set (note that proper ODBC driver should be installed.) Data import from text files has been developed to read out input data file of Xpress software. Hence, a general text file can be imported using built-in Excel functions (i.e. Data → Get External Data).

Next, data can be exported to either MOSEL or GLPK (GNU Linear Programming Kit) format. MOSEL is the name of language that Xpress uses and GLPK is the name of package that can be used in GUSEK (GLPK Under Scite Extended Kit) to solve optimization problems. Since MOSEL model for TFDEA is working under the license only, it is required to access PSU VPN (Virtual Private Network) connection. On the other hand, GUSEK is free optimization software downloadable from http://gusek.sourceforge.net/gusek.html#2. Since this guide focuses on R implementation, interested readers can be referred to corresponding official websites.
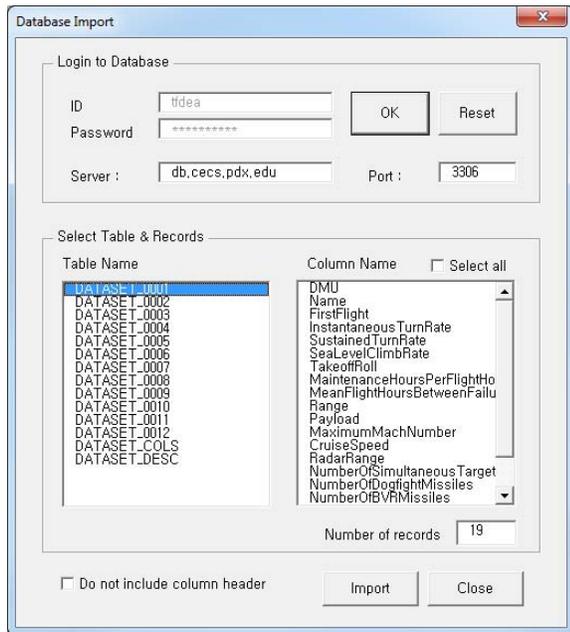

Figure 2 Data import from MySQL

B. *API (Application Programming Interface)*

Version 1.0 of TFDEA allows analysis to be done by Xpress or R through the API. Again, the focus of this guide can dispense with Xpress section. It should be noted that activating Xpress buttons without running Xpress software with authorized network connection could generate unintended error messages or freezing the dialogue box.

Fig. 3 shows the dialog box for R-TFDEA function. The upper half of the control boxes are designed to select data set from the spreadsheet to TFDEA model (Names, Dates, Input(s), and Output(s)). When users check 'Constant 1' box in either Input(s) or Output(s) section, TFDEA model will use dummy variable using constant 1. This is typically when done all of the metrics identified are best classified as inputs or as outputs leaving the other side with nothing.

The lower half of the dialogue box is for setting the particular options in the TFDEA model. In the 'Define Frontier' section, users are supposed to select Static or Dynamic frontier type, which determines whether effective dates should be taken into account in calculating Rate of Change (RoC) and making a forecast. If Dynamic frontier type was selected, 2nd goal would have to be defined for resolving multiple optima [8]. Next, the target frontier should be selected to separate data set into learning period and forecasting period in order to validate the model. In the 'Return to Scale' section, one of scale techniques should be chosen. NIRS indicates Non-Increasing Return to Scale which is equivalent to DRS (Decreasing Return to Scale) whereas NDRS (Non-Decreasing Return to Scale) means IRS (Increasing Return to Scale). Lastly, orientation should be selected to define primary goal of DEA model. Input oriented model measures efficiency in a way that captures how many input(s) could have saved producing the same level of output(s) whereas output oriented model is keen to capture how many output(s) could have produced using the same amount of input(s).
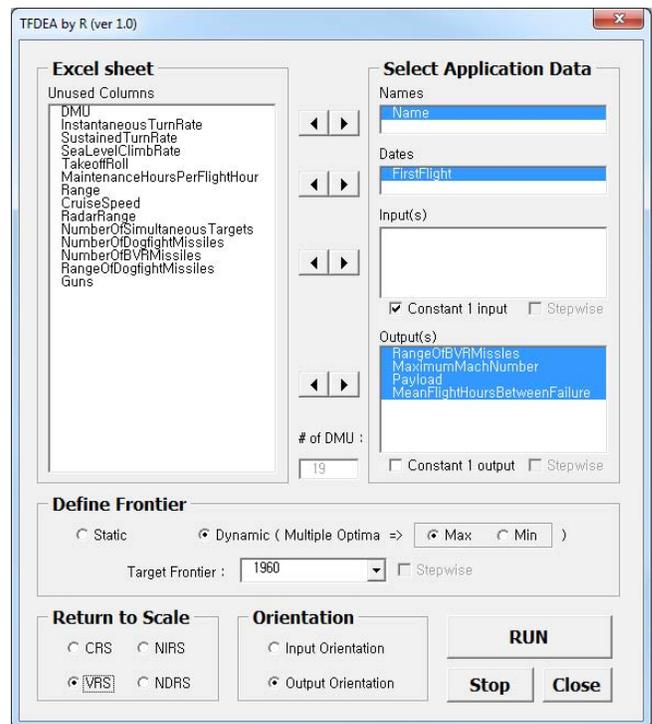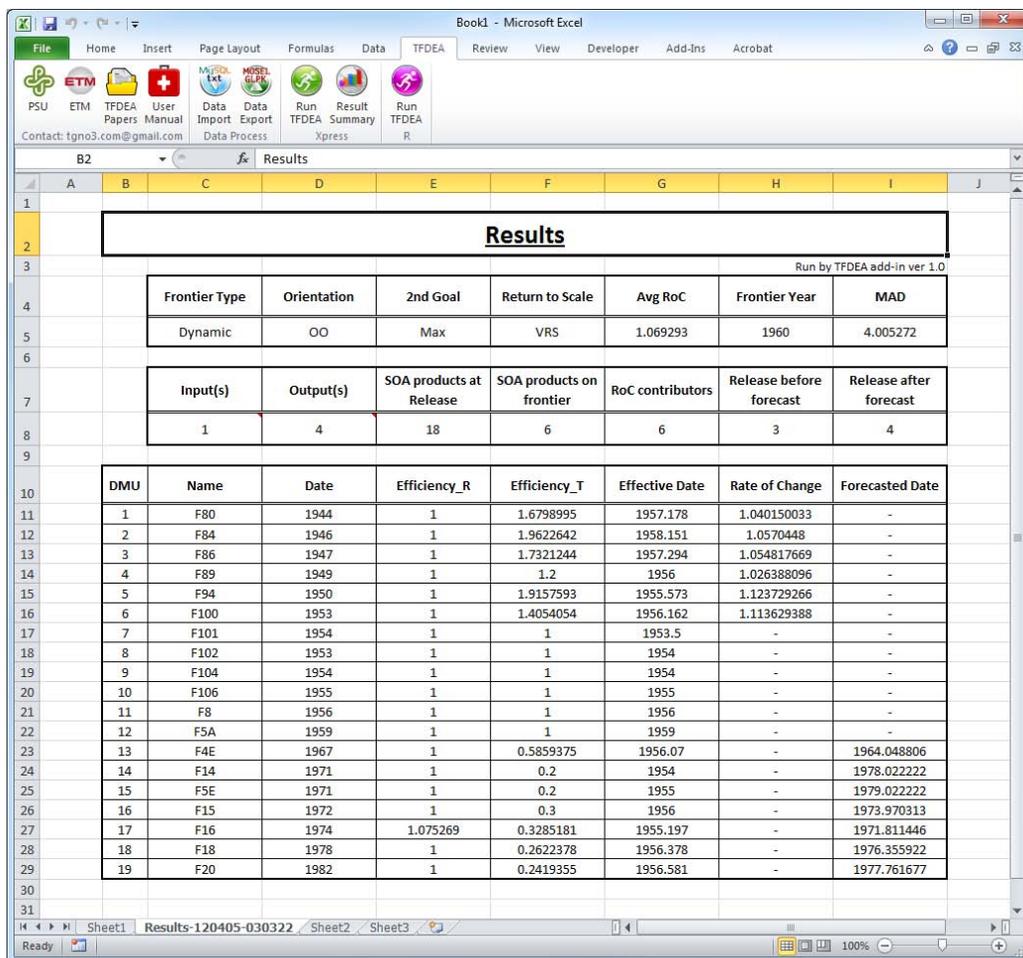

Figure 3 API dialogue box

Figure 4 Results of TFDEA using R

Once application data and parameter settings are assigned, TFDEA can be performed by hitting the 'RUN' button. Analysis time varies with the number of variables, products, and certain parameter settings. Since the Excel macro calls R through middleware to run the DEA formula, it could take several more seconds to activate R and StatconnDCOM if they are not running.

Fig. 4 shows the sample results. A new sheet is generated with the name of 'Results-YY/MM/DD-HH/MM/SS' next to the sheet where application data and parameters have been defined.

IV. ILLUSTRATIVE EXAMPLES

Fig. 5 shows the results of TFDEA using Fighter Jet data published in 2006 [4]. The first table summarizes parameters which have been used for the analysis with average rate of change (Avg RoC: 1.069293) and mean absolute deviation (MAD: 4.005272). The first two columns in the second table show how many input and output variables have been used for the DEA model. The names of variables are put into the

comment box that shows up when the mouse rolls over the correspondent cells. Third and fourth columns indicate how many DMUs (Decision Making Unit i.e. products) used to be or still are state-of-the-art (SOA) denoted by efficiency score of 1. Fifth column indicates how many DMUs have been taken into account for calculating the average rate of change. Here, it can be noticed that TFDEA makes forecasts based on technology advancements from 1944 to 1953. The last two columns categorize forecasts into two groups; 'release before the forecast' and 'release after the forecast.' The third table shows detailed results of this analysis including efficiencies at release date, efficiencies at target frontier, and effective dates. Detailed information or definitions of these parameters can be found in original paper [1].

It should be addressed here that the results are not exactly the same as previous paper. This is because the issue of multiple optima has not been resolved in R based DEA platform yet. Without secondary objective, the results can be seen equal to another paper highlighting this issue [8]. This problem is discussed in future works section in detail.

Results

Run by TFDEA add-in ver 1.0

Frontier Type	Orientation	2nd Goal	Return to Scale	Avg RoC	Frontier Year	MAD
Dynamic	OO	Max	VRS	1.069293	1960	4.005272

Input(s)	Output(s)	SOA products at Release	SOA products on frontier	RoC contributors	Release before forecast	Release after forecast
1	4	18	6	6	3	4

DMU	Name	Date	Efficiency_R	Efficiency_T	Effective Date	Rate of Change	Forecasted Date
1	F80	1944	1	1.6798995	1957.178	1.040150033	-
2	F84	1946	1	1.9622642	1958.151	1.0570448	-
3	F86	1947	1	1.7321244	1957.294	1.054817669	-
4	F89	1949	1	1.2	1956	1.026388096	-
5	F94	1950	1	1.9157593	1955.573	1.123729266	-
6	F100	1953	1	1.4054054	1956.162	1.113629388	-
7	F101	1954	1	1	1953.5	-	-
8	F102	1953	1	1	1954	-	-
9	F104	1954	1	1	1954	-	-
10	F106	1955	1	1	1955	-	-
11	F8	1956	1	1	1956	-	-
12	F5A	1959	1	1	1959	-	-
13	F4E	1967	1	0.5859375	1956.07	-	1964.048806
14	F14	1971	1	0.2	1954	-	1978.022222
15	F5E	1971	1	0.2	1955	-	1979.022222
16	F15	1972	1	0.3	1956	-	1973.970313
17	F16	1974	1.075269	0.3285181	1955.197	-	1971.811446
18	F18	1978	1	0.2622378	1956.378	-	1976.355922
19	F20	1982	1	0.2419355	1956.581	-	1977.761677

Figure 5 Results of Fighter Jet

V. COMMON MISTAKES

A. Inappropriate Return to Scale (RTS) or Orientation

Except under CRS (Constant Return to Scale), input efficiency of DMU is not equal to output efficiency [9]. As a consequence, a different choice of orientation usually affects the results of TFDEA. When VRS (Variable Return to Scale) is applied, the model tends to recognize more DMU as SOAs, which is likely to scale up overall efficiency. This could oversee technology advancements over time thereby, ultimately, ending up with unrealistic forecasts both in practice and software. Relating returns to scale to technology forecasting is an on-going research area.

B. Zero average RoC (Rate of Change)

TFDEA makes forecast using following equation

$$t_k^f = \frac{\ln(E_k^{SE})}{\ln(RoC_{Avg}^{t_f})} + t_k^{eff} \quad \forall k \qquad (1)$$

where t_k^f is the forecasted year of technology k, E_k^{SE} the super-efficiency of technology k, $RoC_{Avg}^{t_f}$ the average rate of change until frontier year t_f, and t_k^{eff} the effective year of technology k. Zero average RoC, which happens mostly within small data set, makes the value of the log function negative infinity, thereby making forecast infinite. Researchers should therefore select proper parameters to prevent the model from counting only a few DMU to be considered or excluding key DMUs, which show significant technological advancements over time, from the SOA group. This could be an issue with the model being too complex in terms of the number of inputs and outputs given the number of products.

C. Infeasible super-efficiency

Under VRS, the super-efficiency model can be infeasible [10]. This may become a critical issue since TFDEA measures distance from target frontier to future DMU (after frontier date) using super-efficiency scores to make a forecast. Hence, it should be carefully verified, when one wants to use

VRS, the difference of super-efficiency scores between input-orientation and output-orientation in order to capture the feasibility. (Note that, it is possible that neither an input nor an output super-efficiency score can be computed for certain units.) This is also an issue of ongoing research.

VI. FUTURE WORKS

A. Multiple-optima

This add-in is developed using the package 'Benchmarking' provided by R. The 'Benchmarking' package. For optimization module it depends upon the 'lpsolve API' which has been developed under LGPL (Lesser General Public License). In other words, 'Benchmarking' package is wrapping the MLP (mixed linear programming) solver module to provide a simple DEA interface through an R command. However, the problem of multiple-optima can be resolved by manipulating optimization problem at a basic level, which cannot be reached through the Benchmarking package, with a secondary objective function. One of future works, therefore, is to develop purely standalone computational model of TFDEA using 'lpsolve API' and wrap it up with VBA (Visual Basic for Applications) in Excel rather than relying on DEA command from simplified R package.

B. Non-radial model

Current TFDEA uses only radial DEA model to measure the (super) efficiency. However, non-radial model would be an interesting option that can incorporate subject judgment into the DEA model. Most of all, non-radial model has been preferred as a solution for zero-output problem in traditional DEA model as well as infeasibility problem of super-efficiency [11]. Since 'Benchmarking' package doesn't allow extension of DEA model, this work also requires element level of linear programming.

VII. USAGE PERMISSION

Licensor hereby grants any users a non-exclusive, non-transferable, non-sub-licensable free of charge license to download, install and use this add-in program on an unlimited number of computers for non-commercial, educational use. (Re)Distribution of the software is prohibited within any organization and only permissible from official TFDEA website. Since this add-in is built on R and Statconn environment, users should also visit their official websites to download the program and abide by licenses of them [12]. All publications based on analyses performed using this software directly or indirectly must include the citation of this PICMET paper.

VIII. CONCLUSION

Despite its strength compared to stochastic frontier model, TFDEA hasn't been widely used due to its computational complexity. This software has been developed in order to lessen this burden from the programming side so that researchers from various fields can focus more on data gathering and interpretation of the results to which their expertise is worth being contributed.

REFERENCES

[1] O. L. Inman, "Technology forecasting using data envelopment analysis," Portland State University, 2004.

[2] R. M. Heiberger and E. Neuwirth, *R Through Excel: A Spreadsheet Interface for Statistics, Data Analysis, and Graphics*, 1st ed. New York: Springer, 2009, p. 368.

[3] P. Bogetoft and L. Otto, *Benchmarking with DEA, SFA, and R*, 1st ed. New York: Springer, 2010, p. 367.

[4] O. Inman, T. Anderson, and R. Harmon, "Predicting U.S. jet fighter aircraft introductions from 1944 to 1982: A dogfight between regression and TFDEA," *Technological Forecasting and Social Change*, vol. 73, no. 9, pp. 1178-1187, 2006.

[5] T. R. Anderson, K. Hollingsworth, and L. Inman, "Assessing the rate of change in the enterprise database system market over time using DEA," in *PICMET '01. Portland International Conference on Management of Engineering and Technology. Proceedings Vol.1: Book of Summaries (IEEE Cat. No.01CH37199)*, p. 203.

[6] T. Anderson, R. Fare, S. Grosskopf, L. Inman, and X. Song, "Further examination of Moore's law with data envelopment analysis," *Technological Forecasting and Social Change*, vol. 69, no. 5, pp. 465-477, 2002.

[7] J. Kim, T. Daim, and T. Anderson, "A look into the future of wireless mobile communication technologies," *Technology Analysis & Strategic Management*, vol. 22, no. 8, pp. 925-943, Nov. 2010.

[8] T. R. Anderson and L. Inman, "Resolving the issue of multiple optima in Technology Forecasting using Data Envelopment Analysis," *Technology Management in the Energy Smart World (PICMET), 2011 Proceedings of PICMET '11:*, pp. 1-5.

[9] W. W. Cooper, *Handbook on Data Envelopment Analysis*, vol. 2010. New York: Springer, 2010, p. 497.

[10] J. Zhu, "Super-efficiency and DEA sensitivity analysis," *European Journal of Operational Research*, vol. 129, no. 2, pp. 443-455, Mar. 2001.

[11] H. O. Fried, C. A. K. Lovell, and S. S. Schmidt, *The measurement of productive efficiency and productivity growth*, 1st ed. New York: Oxford University Press, 2008, p. 638.

[12] Baier Thomas and Neuwirth Erich, "Excel :: COM :: R," *Computational Statistics*, vol. 22, no. 1, pp. 91-108, 2007.

APPENDIX I.
INSTALLATION

TFDEA add-in requires following software installed:
- Microsoft Excel (2003, 2007, or 2010)
- R with 'Benchmarking' package
- Statconn tools (middleware)

The easiest way to install R and middleware is to download the most current version of 'RAndFriends' from Statconn website (http://rcom.univie.ac.at/download.html). This free package contains R and Statconn tools that users can easily install software needed at once. For the first time installers, 'Benchmarking' package should be installed in R as well (run R → Packages → Install package(s) → Select CRAN mirror server → Benchmarking).

If one wants to customize each module with specific settings, individual installation is also available. Most current version of R is downloadable from the official website (http://www.r-project.org/) for free. Once the first time installers have installed R, 'Benchmarking' package should also be installed. Likewise, most current versions of Statconn programs (StatconnDCOM, rcom, RExcel) are available on Statconn website for non-commercial free version.

One who wants to use database import function, proper version of MySQL ODBC drivers corresponding to the operating system in use should be installed (http://dev.mysql.com/downloads/connector/odbc/).

Finally, TFDEA add-in file (TFDEA_ver1.0.xlsm) is available for download from official TFDEA website. Users can open this file whenever they want to use TFDEA add-in functions. If one wants to set up this add-in as a default menu in Excel, he/she may want to follow these steps: Open TFDEA add-in file (TFDEA_ver1.0.xlsm) in Excel → File → Save As → Excel Add-In (*.xlam) → Save (to automatically designated folder) → File → Options → Add-Ins → Go → Browse → Tfdea_Ver1.0 → Check (activate checkbox) Tfdea_Ver1.0 → OK. By doing this, TFDEA ribbon tab will be presented next to the data tab whenever Excel is running (as seen in Fig.7). Deactivation can be simply done by unchecking Tfdea_Ver1.0 in Add-in setting.

Figure 6 TFDEA add-in interface