$$
\begin{aligned}
&Max \quad \theta_0 - \varepsilon \vec{1} s^+ - \varepsilon \vec{1} s^- \\
&s.t. \\
&Y\lambda - s^+ = \theta_0 Y_0 \\
&X\lambda + s^- = X_0 \\
&\sum_i \lambda_i = 1 \\
&\lambda, s^+, s^- \geq 0
\end{aligned} \tag{5}
$$

The BCC model seeks to maximize the proportion of the outputs of the firm under evaluation, j_0, based on a weighted combination of the inputs and outputs of the other firms, which outperform firm j_0-th.

Theta lies between 0 and 1 and represents the technical inefficiency of the firm under evaluation. If the optimal theta for a given unit is equal to one, then it is considered radially efficient and radially inefficient otherwise.

When the unit under evaluation, DMU_0, is rated as inefficient, the solution to the dual problem provides a number of DMUs—the peer group or reference set—which are rated as efficient with the weights of DMU_0. These units are those whose corresponding values of lambda are non-zero for the unit under evaluation. Moreover, the optimal solution of the model provides a virtual unit on the frontier constructed as a linear combination of the units in the reference set. The unit being evaluated should be transformed into that virtual DMU in order to become efficient. This is made by a radial reduction of the inputs or expansion of the outputs—for an input or output orientation respectively—by means of the optimal value of theta. Therefore, theta is the expansion ratio that the outputs should be expanded to make the analyzed unit efficient. If we desire the unit under analysis to be more efficient, we must multiply the output (or input, depending on the orientation) vector by theta as well as add the slacks to the radial expansion (or contraction) of the outputs (inputs).

Selection of Inputs and Outputs

Some excellent examples and characteristics on a wide variety of different compressors appear in [4]. We used some of the examples provided on that Web page, including the measurement of several variables that can be used for an efficiency analysis.

To analyze the efficiency of different compressors, we should take into account a series of variables or characteristics they possess (speed, compress ratio). In our study, we considered just one input, which is given by the size (in bytes) of the file to be compressed. This input has the same size for all units (as we have used the same file for the analysis of all compressors). Hence, as the DEA methodology is unit invariant, the inclusion of this input is effectively the same as taking an input equal to one for all units.

As we only have one input and as it is the same for all units, the CCR and the BCC models provide the same results; a scale efficiency analysis cannot be carried out in this case.

Table 1. TE related to text files.

Compressor	TE score	Compressor	TE score
BIX1.00b7	1	SUPERPACK2.0	0.86
EXPv1	1	AKT320.62b	0.84
LZOP1.00w	1	HPA1.8	0.84
PAR1.49b	1	AKT0.5b	0.84
PKZIP2.6.02win95	1	HYPER2.6	0.78
XPA32.1.02	1	STUFFIT5.5->LZH	0.78
JAR32.1.02	0.99	LHA2.55b	0.77
WINACE2.11	0.98	CAR1.50	0.76
ZZIP0.32b+TAR	0.98	MSXIE1.40	0.76
LZAP0.20.0+TAR	0.98	SAR1.0	0.76
WINZIP7.0	0.97	PUT3.47	0.76
JAR16.1.02	0.95	PAK2.51	0.75
LHA2.67(Win32)	0.93	JAM-Aug1996+TAR	0.74
QUANTUM0.97	0.93	ARQ3.2	0.73
STUFFIT5.5->CAB	0.92	ZIP-ARCHIV2.0	0.71
YAC1.02	0.91	DWC5.10	0.71
HUFFMANC.E.II0.21d	0.91	ARX1.0	0.71
PPMC3h+TAR	0.91	SBX1.4	0.71
DEEPFREEZE1.05a	0.91	ICE1.02c+TAR	0.71
STUFFIT5.5->TGZ	0.91	PAK-DD1.0a	0.71
7-ZIP2.00	0.91	OPAQUE1.0	0.71
STUFFIT5.5->ZIP	0.90	KBOOM1.1	0.69
ARJZ0.15	0.90	PKPAK3.61	0.68
UCv2.37beta	0.90	SQUISH1.0+TAR	0.68
PKZIP2.50	0.90	PKARC3.5	0.68
LHARK0.4d	0.90	ICOMP3.00.061	0.68
UC2v3.0PRO	0.90	MDCD1.0	0.64
LIMIT1.2	0.89	NSK5.0	0.63
ZET0.10b	0.89	SQUASH1.21	0.63
CODEC3.21	0.88	ARC6.02	0.62
GZIP1.2.4+TAR	0.88	ASD0.1.3	0.62
PKZIP2.04g	0.88	LARC3.33	0.62
QUARK1.0beta	0.88	SPLINT2.1+TAR	0.60
AIN2.32	0.88	BLINK2.50	0.52
SQZ1.08.4a	0.88	NULIB3.24+TAR	0.51
BSArc2.0R1.1	0.87	GAS2.0+TAR	0.48
OOP2.3	0.87	NPACK1.0+TAR	0.47
SKY1.15	0.87		

We have also considered three different outputs—the speed (in bytes per second) at compressing the input files, the speed at extracting files, and the compress ratio. The first two can represent a positive service or output (the benefit obtained by speed of the compressor) that we want to maximize. The third output is also important as it indicates how much the input file can be compressed. The compression ratio has been calculated as the percentage the input file has been reduced. For example, if a file of size 100 has been reduced to a file of size 25, we will say the compression ratio is 75%. Therefore, the higher the compression rate the better as we will have a greater compressed file.