ISSN: 2455-6653

Use of Main Components and Data Envelopment Analysis to Evaluate the Efficiency of Water Operators

José Sánchez
Uriel Rivera-Ortega
(UPAEP 2018)

Abstract

The purpose of this work is to analyze the efficiency of Water Operator Organizations of México from the premise that each operator organization is in different working conditions and evaluated from different performance indicators, so obtaining efficiency for each organism could ambiguously arise in this regard, this manuscript paper provides from analysis of generalized variables and their relationship with factors of operation those variables that influence the efficiency process. This document is based on a free database from the Program of Indicators of Management of Operative Organizations (PIGOO in spanish), which initially involves 139 municipal water systems throughout the Mexican Republic. The result highlights the influence of certain variables in the performance of the operation of water systems, according to this, the usefulness of this work is expected to contribute by defining elements that improve the competitiveness of the aforementioned organizations.

Keywords : Data Envelopment Analysis (DEA) , Principal Components Analysis (ACP) , Water Operators

0 Introduction

Efficiency is presented as an indicator of the competitiveness of organizations. There are various studies that have displayed methodologies by analyzing multiple criteria, in order to improve strategic competitiveness [1]. These techniques as an efficiency measurement method try to evaluate the performance of an entity with respect to an optimal value. Although it is possible to identify the best governmental practices by comparing economic elements, the optimization process of the organizations is guided by more than one measure of performance, which is a problem when there is not explicit relationship between performance measurements. The solution for this type of problems is not concentrated on obtaining a single optimal solution, but on creating a set of favorable solutions called efficient frontier [2]. To find this efficient frontier a nonparametric technique of interest is the Analysis Data Envelopment (DEA) which shows the maximum relation of the products (outputs) by giving their inputs (inputs).

In addition to the problem of conflict between measurements of performance, there is also the problem related to the amount of data associated with a number of variables performance ; the number of variables in the DEA study should ensure that the number of alternatives is sufficiently discriminated to obtain the efficient frontier [3]; although for the first problem DEA technique resolves the conflict between performance measurements. The second problem is approached from the technique called Principal Components Analysis (PCA), which aims to reduce data by grouping relatively homogeneous variables. This transformation of a wide set of variables with high correlation leads towards a reduced set of variables that explains most of the variations in the data [4].

1 Theoretical framework

The structured theory about the exploration and understanding of the competitiveness factors, as well as the relationship between the variables and the determination of the efficient frontier, are concentrated in the Principal Component Analysis and DEA techniques which will be below.

1. 1 Principal Component Analysis

In the most investigations and study cases, the most frequent is to take as much information as possible and collect the largest number of variables involved and, consequently, a quantity of data of a different category. It is possible that in a collection of research data, these are interrelated from the different variables in which they have been included, this condition presages variability in a study, so it is necessary to reduce the number of variables under the theoretical justification that variables with strong correlation are actually measuring the same concepts but from different approaches.

The PCA consists of concentrating the information contained in an original set of variables to take it to another set of variables, always in smaller quantity than the original ones, therefore if there is a set of K original variables, the information is transformed into a set of W components being W <K [4], each of these W components are translated into factors as a result of a linear combination of the K variables. The utility of the PCA is to take advantage of the fact that the resulting factors reflect the variability of the original groups. In PCA the first main component is the axis that passes through the center of the data and minimizes the distance of each data point to the same axis explaining the behavior of a group of data, the rest of the components arise as other axes that have the purpose of explaining what the first main component could not do, however one characteristic to be fulfilled by the following main components is that they must be orthogonal (independent) and pass through the centroid of the data. The variability explained by each axis gives place (originates) to the concept of eigenvalue which can also be expressed as a percentage of the total variation.
The mathematical model of the ACP is defined as follows [4]:

$$X_{ij} = a_{1j}F_{i1} + a_{2j}F_{i2} + \ldots + a_{Kj}F_{iK} \qquad (1.1)$$

Where X_{ij} is the value of the j-th variable in the i-th case resulting from the product of a_{1j}, a_{2j}, …, a_{Kj}, as vector of constants and each of the factors F.

This model expresses that the information of the variables is explained entirely by the "K" factors. De la Garza & González [5] worked with the proportion of the variability of each variable by the factors, this is the reason that in PCA the initial value of all the variables is equal to 1. The components are chosen according to the highest variance, a form to maximize the variance is increasing the coefficients a_{Kj} maintaining the orthogonality of the transformation, which requires that the vector a_{1j}, a_{2j}, ..., a_{Kj} be equal to 1, that is [4]:

$$\sum_{i=1}^{K} a^2{}_{ij} = 1 \qquad (1.2)$$

1. 2 Data Envelopment Analysis (DEA)

The objective of a Data Envelopment Analysis is the formation of enveloping faces that define the efficient and inefficient units. In this case it is as important to know the " best " as to know the distance (actions) that separates the efficient Decision Measurement Units (DMU) of those who are not. This will be a guideline for the progress of the entities who wish to achieve the called benchmarks (optimal entities). This situation presents an implicit problem in relation to the "goals" that less efficient DMUs must reach because common sense explains the impossibility of competing with benchmarks for which the efficiency is too great to look for a comparison, in terms of González and Álvarez [6] DEA analyzes could be reconfigured to seek a single-stage procedure to achieve closer efficiency goals.

While a virtue of the DEA is obtaining efficiency in the use of multiple units in its function of inputs or products, also presents some criticism in that it does not contemplate influences on the process productive, which generates uncertainty in the final results, in studies by Drake and Howcroft [7] it is mentioned that the DEA is capable of working better if the number of observations is close to twice the sum of Inputs and Outputs, which would indicate that in studies with small samples should be added relatively too many categories, this results in a complexity to identify the optimal DMU, however the DEA models create from iterations a progress in the proposal to choose the efficient DMU [8].

2 Methodology

This work was carried out through an investigation of institutional references related to competitiveness factors that allude to the efficiency of municipal entities. From the database of PIGOO, an explanatory study was documented, since the appropriate variables for the efficiency study were determined. Initially, 139 municipal water operators were considered from all over the republic. However, under a discrimination process, the study contemplated 51 organisms, the reason for excluding certain organisms consisted of not having complete information in all the study variables from 2010 to 2015; the research process is used for the integration of statistical techniques in order to model the influence of social variables in the performance of municipal water operators.

2.1 Selection of variables with ACP

The study begins with the collection of the information corresponding to the variables involved with the performance of the water operators in the different municipalities of the country; showing that isn't competition in the service offered regarding to the distribution of the drinking water and drainage network, because each municipality and/or the metropolitan area is represented by a single water operator.
Under this principal condition, each municipality will be considered as a DMU object of efficiency study under the following indicators:

ISSN: 2455-6653

1	Drinking water coverage reported%	15	Micromeasurement%
2	Sewage coverage reported%	16	User registry%
3	Consumption (liters / hours / day)	17	Losses by network length (m^3 / km)
4	Cost between produced volume ($ / m^3)	18	Losses per intake (m^3 / intake)
5	Equipment (liters / hours / day)	19	Claims per thousand intakes (units)
6	Commercial efficiency%	20	Networks and facilities%
7	Collection efficiency%	21	Rehabilitation of household outlets%
8	Physical efficiency 1%	22	Rehabilitation of pipe%
9	Physical efficiency 2%	23	Cost-Rate Ratio
10	Overall efficiency%	24	Working relationship %
11	Employees dedicated to leak control	25	Outlets with continuous service%
12	Employees per thousand intakes (units)	26	Users supplied with pipes%
13	Hours with service in the sampling area	27	Users with payment on time%
14	Macromeasure%	28	Volume treated%

Table 1. List of indicators taken from the PIGOO database ($=mexican pesos).

With this number of variables which in turn include data for 51 utilities for a period of 5 years, the ACP technique is used to determine whether more than one variable is strongly correlated with a different variables that are part of the study, that in technical terms one can consider multicollinearity. In these conditions it is convenient to consider the benefit of grouping these correlated variables considering that they are actually measuring the same issue.

Due to the number of variables used, it was not possible to plot the point cloud that would visually reflect the behavior of the data, however, for the ACP model, this requirement is not necessary in the development of the study. In this way, the first step of the ACP will be the choice of the main components under the premise that at least 90% of the variability of the data is explained. This value is statistically typical value in studies of this nature [9]

The separation of the variables was carried out in two types: Inputs and Outputs. This classification does not concern with the development of the ACP but with the efficiency process of the DEA. So the variables that will be worked in ACP are:

1	Cost between produced volume ($ / m^3)	13	Losses by network length (m^3 / km)
2	Equipment (liters / hours / day)	14	Losses per intake (m^3 / intake)
3	Commercial efficiency%	15	Networks and facilities%
4	Collection efficiency%	16	Rehabilitation of household outlets%
5	Physical efficiency 1%	17	Rehabilitation of pipe%
6	Physical efficiency 2%	18	Cost-Rate Ratio
7	Employees dedicated to leak control	19	Working relationship %
8	Employees per thousand intakes (units)	20	Outlets with continuous service%
9	Hours with service in the sampling area	21	Users supplied with pipes%
10	Macromeasure%	22	Users with payment on time%
11	Micromeasurement%	23	Volume treated%
12	User registry%		

Table 2. List of indicators of inputs variables for the ACP study ($=mexican pesos).

When the components have been selected, they are presented in the form of a correlation matrix. This step allows confirming the relationship between variables, in this case variables type input. The following table shows a triangular formation with its diagonal with values of 1 because the same concepts exist in the columns and rows axis, so the variable C in the row and column crossing will have correlation value of 1 since it is perfectly correlated to be the same concept. Values above 0.75 indicate strong correlation and because the correlation offers the same result regardless of the order of the even variables, the table presents a lower triangular matrix. To simplify the presentation, each variable has been represented by the variable number taken from table 2.

Table 3 shows weak correlation between the chosen variables since few absolute values within the table are close to 1. This condition is fundamental at the moment in which the ACP model decides how many components should be generated to explain 90% of the total variability of the data collected.

Variables	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20	21	22	23
1	1.000																						
2	-0.199	1.000																					
3	0.068	0.153	1.000																				
4	0.255	0.113	0.447	1.000																			
5	0.102	-0.132	-0.084	0.166	1.000																		
6	0.026	-0.131	-0.058	0.079	0.407	1.000																	
7	-0.184	-0.223	0.140	-0.093	-0.105	0.070	1.000																
8	0.160	0.115	0.171	-0.009	-0.259	-0.211	0.161	1.000															
9	-0.020	-0.293	-0.087	-0.134	0.003	-0.073	0.215	0.173	1.000														
10	0.382	0.196	0.172	0.315	0.096	0.046	-0.101	-0.097	-0.131	1.000													
11	0.350	-0.135	0.049	0.245	0.224	0.184	-0.146	-0.114	-0.207	0.283	1.000												
12	0.258	0.187	-0.144	0.059	-0.058	0.076	-0.361	0.067	-0.276	-0.012	0.339	1.000											
13	-0.131	-0.073	-0.096	-0.184	0.283	-0.037	0.028	-0.108	0.054	-0.240	0.031	-0.006	1.000										
14	-0.083	0.061	0.061	-0.022	0.238	-0.176	0.093	0.355	0.020	-0.173	0.085	0.062	0.234	1.000									
15	0.057	-0.112	-0.015	-0.072	0.074	0.107	0.172	-0.022	0.140	0.224	0.083	0.048	0.091	-0.115	1.000								
16	-0.033	-0.200	0.142	0.084	0.160	0.124	0.014	0.218	0.090	-0.066	0.138	0.215	0.225	0.145	0.075	1.000							
17	-0.053	0.169	0.057	-0.050	-0.208	0.033	0.019	0.134	0.029	0.061	-0.024	0.060	-0.154	0.094	0.005	-0.079	1.000						
18	-0.052	0.118	0.007	-0.106	-0.112	-0.004	0.247	0.111	0.142	-0.069	-0.394	-0.110	0.113	-0.031	0.096	-0.121	0.036	1.000					
19	-0.216	-0.007	0.156	-0.116	-0.050	-0.106	0.239	0.213	0.253	-0.142	-0.252	-0.136	-0.035	0.044	-0.115	0.019	-0.014	0.031	1.000				
20	0.238	0.149	0.177	0.375	0.139	0.160	-0.424	-0.197	-0.536	0.229	0.503	0.326	-0.123	-0.044	0.002	-0.174	-0.013	-0.233	-0.160	1.000			
21	0.008	0.037	-0.125	-0.106	0.159	0.009	0.065	0.389	0.149	0.005	-0.260	0.010	-0.124	0.263	-0.058	0.186	0.082	-0.086	0.094	-0.195	1.000		
22	-0.168	0.000	0.135	0.427	-0.008	0.243	0.092	-0.122	0.079	-0.143	0.023	-0.017	-0.229	-0.158	-0.180	0.144	-0.032	-0.177	0.016	0.041	-0.046	1.000	
23	0.116	0.114	0.065	0.130	0.081	-0.055	-0.323	-0.134	-0.379	0.250	0.050	0.194	-0.238	-0.203	-0.095	0.219	0.002	-0.270	-0.029	0.204	0.075	-0.016	1.000

Table 3. Correlation matrix between variables inputs.

The next step was to consider the data in the analysis of principal components. It was chosen to use the software Minitab due to its performance in statistical data analysis.

Table 4 shows that up to variable 15 that exceeds 90% of the explained variability in the accumulated variable (highlighted in yellow on the table). Variable 1 explains this variability in 15.2%, however the other variables (up to 23 according to table 2) they contribute individually no more than 2%, so to achieve an accumulated 90% of the explanation of variability it is necessary to consider most of the variables, which is confirmed with fig. 1.

	Variables Input																						
	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20	21	22	23
Eigenvalor	3.4856	2.1209	1.9569	1.8839	1.6571	1.4554	1.2563	1.2274	1.1398	1.0042	0.8857	0.8169	0.7725	0.5979	0.5308	0.45	0.3828	0.3604	0.2907	0.2532	0.2028	0.1464	0.1222
Proportion	0.152	0.092	0.085	0.082	0.072	0.063	0.055	0.053	0.05	0.044	0.039	0.036	0.034	0.026	0.023	0.02	0.017	0.016	0.013	0.011	0.009	0.006	0.005
Accumulated	0.152	0.244	0.329	0.411	0.483	0.546	0.601	0.654	0.704	0.747	0.786	0.821	0.855	0.881	0.904	0.924	0.94	0.956	0.968	0.98	0.988	0.995	1

Table 4. Cumulative correlation matrix of input variables.

ISSN: 2455-6653

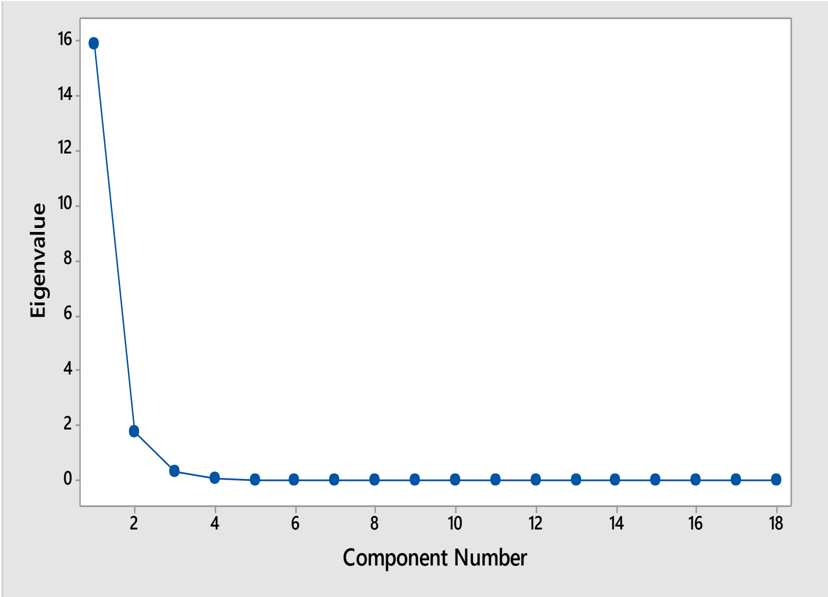

Fig. 1. Proportion of the variability explained by each of the input variables.

Table 5 highlights the values that will be the coefficients of the ACP model. They are obtained by selecting the highest absolute value by row. As shown in the mathematical model of the previous section, the product of these factor coefficients of the jth variable with each of the factors will be what determines the value of the jth variable in the ith case.

Variable/ ACP	PC1	PC2	PC3	PC4	PC5	PC6	PC7	PC8	PC9	PC10	PC11	PC12	PC13
1	0.236	0.035	0.172	-0.016	**0.434**	-0.166	-0.2	-0.099	-0.196	-0.311	-0.13	-0.063	0.247
2	0.084	**-0.407**	-0.02	-0.15	-0.097	0.253	0.363	0.131	0.11	-0.114	0.08	0.274	-0.238
3	0.082	-0.238	0.17	**0.385**	0.09	0.341	0.017	-0.205	0.195	0.112	-0.017	-0.062	0.175
4	0.276	-0.102	0.139	**0.421**	0.012	0.19	0.024	-0.05	-0.055	-0.273	0.17	0.088	-0.02
5	0.123	0.432	0.169	0.043	-0.085	0.06	**0.469**	-0.118	-0.211	-0.052	-0.057	0.131	0.102
6	0.122	0.312	-0.04	0.221	-0.058	-0.092	0.299	**0.46**	0.019	-0.025	-0.325	-0.103	0.161
7	-0.289	0.064	0.025	0.316	0.178	0.107	0.025	0.093	0.059	0.245	-0.135	**-0.436**	-0.163
8	-0.166	-0.261	**0.476**	-0.034	0.168	0.01	-0.143	0.09	-0.002	-0.122	-0.152	-0.094	-0.102
9	-0.31	0.143	0.092	0.189	0.171	-0.189	-0.143	0.007	-0.123	-0.111	0.175	**0.561**	0.087
10	0.258	-0.107	0.009	0.105	**0.423**	-0.124	0.308	-0.155	-0.005	0.101	0.156	0.126	-0.061
11	**0.346**	0.225	0.138	0.008	0.125	0.132	-0.285	0.094	-0.096	0.274	-0.007	0.022	-0.046
12	0.244	-0.021	0.193	-0.294	-0.061	-0.037	-0.224	**0.332**	0.272	-0.178	-0.244	0.175	-0.086
13	-0.109	0.363	0.057	-0.227	-0.059	**0.402**	0.035	-0.157	0.213	-0.048	0.136	0.137	0.158
14	-0.097	0.037	**0.448**	-0.177	-0.069	0.37	0.066	0.03	-0.278	0.129	0.122	-0.095	-0.064
15	-0.003	0.202	-0.031	0.008	0.444	-0.037	0.119	0.125	0.39	0.212	0.029	0.185	**-0.455**
16	0.006	0.199	0.431	0.093	-0.158	-0.109	-0.068	-0.032	**0.543**	-0.034	0.143	-0.051	0.128
17	-0.024	-0.244	0.059	-0.049	0.1	-0.022	0.082	0.506	-0.028	0.408	0.259	0.124	**0.554**
18	-0.219	-0.07	-0.151	-0.025	0.261	0.229	0.214	0.135	0.204	**-0.485**	-0.189	-0.104	0.255
19	-0.21	-0.135	0.108	0.158	-0.126	-0.009	0.012	-0.227	0.016	0.268	**-0.647**	0.432	0.091
20	**0.417**	-0.045	-0.036	-0.037	-0.023	0.214	-0.016	0.056	-0.128	0.09	-0.287	0.037	-0.079
21	-0.124	-0.068	**0.408**	-0.076	-0.061	-0.348	0.347	0.077	-0.204	-0.077	0.002	-0.102	-0.217
22	0.061	-0.013	-0.019	**0.486**	-0.367	-0.048	-0.131	0.27	-0.003	-0.174	0.155	0.098	-0.183
23	0.254	-0.155	0.059	-0.065	-0.182	**-0.359**	0.186	-0.305	0.303	0.114	0.013	-0.148	0.198

Table 5. Values for Factors.

3.2 DEA model

The projection of the original data in the main components created is the prelude to the application of the DEA technique. The DEA study consists of finding the organisms that are optimal in terms of efficiency, that is, a line of action for maximization is proposed of the efficiency of those entities that belong to the so-called efficient frontier [10].

ISSN: 2455-6653

Table 6 shows the list of entities that will work as DMUs for the DEA study, the inputs will be the main components obtained in the previous section and the outputs will be the indicators that are attached to what the clients of the operating agencies of the water perceived as the service offered.

Water Operator	DMU	Water Operator	DMU	Water Operator	DMU	Water Operator	DMU	Water Operator	DMU	Input	Output
Aguascalientes, Aguascalientes	1	Cuauhtémoc, Chihuahua	11	Huauchinango, Puebla	21	Naucalpan, México	31	Tampico, Tamaulipas	41	PC1	Drinking water coverage reported (%)
Cancún, Quintana Roo	2	Culiacán, Sinaloa	12	Izucar de Matamoros, Puebla	22	Nicolás Romero, México	32	Tecate, Baja California	42	PC2	Sewagen coverage reported (%)
Cd. Juárez, Chihuahua	3	Delicias, Chihuahua	13	La Piedad, Michoacán	23	Pachuca, Hidalgo	33	Tijuana, Baja California	43	PC3	Consumption (l/h/d)
Cd. Mante, Tamaulipas	4	Durango, Durango	14	León, Guanajuato	24	Piedras Negras, Coahuila	34	Tlalnepantla, México	44	PC4	
Cd. Valles, San Luis Potosí	5	Ensenada, Baja California	15	Matehuala, San Luis Potosí	25	Puebla, Puebla	35	Torreón, Coahuila	45	PC5	
Celaya, Guanajuato	6	Fresnillo, Zacatecas	16	Mazatlán, Sinaloa	26	Saltillo, Coahuila	36	Tulum, Quintana Roo	46	PC6	
Chilpancingo, Guerrero	7	González, Tamaulipas	17	Mexicali, Baja California	27	San Juan del Río, Querétaro	37	Tuxpam, Veracruz	47	PC7	
Ciudad de México, Distrito Federal	8	Guadalajara, Jalisco	18	Monclova-Frontera, Coahuila	28	San Martín Texmelucan, Puebla	38	Tuxtla Gutiérrez, Chiapas	48	PC8	
Ciudad Guzmán, Jalisco	9	Guanajuato, Guanajuato	19	Monte Escobedo, Zacatecas	29	Santa María del Tule, Oaxaca	39	Veracruz, Veracruz	49	PC9	
Córdoba, Veracruz	10	Hermosillo, Sonora	20	Monterrey, Nuevo León	30	Silao, Guanajuato	40	Xalapa, Veracruz	50		
								Zacatecas, Zacatecas	51		

Table 6. DMU´s, Inputs and Outputs.

To follow the case study, we proceeded with the study of the outputs. Knowing the premise that an omission of an important input or output results in biased conclusions for an AED analysis. The use of correlation analysis and main components prior to an efficiency study was vital for the validation of DEA results. Table 7 shows the results of the linear combination that produced the ACP and that are established as inputs, in the same table the output values that correspond to each DMU appear.

DMU/ ACP	PC1	PC2	PC3	PC4	PC5	PC6	PC7	PC8	PC9	Drinking water coverage reported (%)	Sewage coverage reported (%)	Consumption (l/h/d)
1	61.398	112.694	48.022	99.047	45.402	6236.301	30.805	60.727	10.450	99.264	98.336	180.866
2	40.697	82.689	64.479	88.000	22.461	5692.546	55.109	43.933	1.957	100.000	90.940	111.622
3	64.587	133.959	2.089	89.445	45.741	26.550	69.498	59.188	3.124	97.500	93.184	172.000
4	59.547	132.239	75.663	81.649	18.156	7314.125	57.447	59.955	6.114	98.000	88.000	171.694
5	75.549	102.425	62.007	119.417	2.347	3538.311	59.295	68.286	5.820	98.822	84.778	142.748
6	73.165	104.754	56.795	83.987	42.538	5289.254	59.535	60.298	10.400	99.888	99.678	147.396
7	0.891	76.269	25.262	75.677	3.488	2026.037	31.168	111.914	4.488	68.000	88.200	159.438
8	54.524	124.296	118.479	87.820	33.946	4.702	54.074	83.642	4.090	98.010	93.736	152.380
9	20.362	131.307	41.002	88.409	15.542	27.608	33.194	66.956	0.645	96.793	95.665	270.270
10	27.545	87.835	48.773	63.661	2.073	6419.834	68.680	64.182	1.809	92.600	92.000	152.842
11	75.796	87.211	54.016	89.855	45.443	34.754	58.660	68.226	1.391	97.000	94.000	166.018
12	72.280	94.977	141.890	104.129	21.507	3193.000	64.484	70.699	2.634	99.440	98.200	148.414
13	70.526	145.568	1.569	87.750	43.971	6408.202	60.913	74.489	2.200	100.000	100.000	220.998
14	56.629	165.841	1.533	111.449	42.984	28.098	50.970	91.587	4.095	98.620	97.624	157.530
15	73.943	76.866	18.179	110.450	49.053	2087.444	70.641	76.174	11.727	99.086	92.972	154.434
16	20.193	107.194	68.218	85.228	24.878	5043.023	56.608	60.883	0.506	97.194	84.312	144.354
17	56.465	85.964	17.067	108.116	8.372	1249.727	59.372	65.126	10.153	96.800	83.600	182.738
18	61.535	87.119	42.661	68.090	35.904	4997.502	64.681	62.803	7.420	97.976	97.686	146.308
19	75.798	79.855	48.084	89.948	41.877	3583.002	62.066	64.890	3.004	91.436	92.694	124.665
20	57.102	135.433	66.770	97.030	8.193	5931.974	59.566	62.285	3.419	98.000	94.278	199.786
21	54.976	54.536	66.529	71.563	0.738	6100.840	47.978	152.488	4.734	80.000	80.000	172.000
22	5.556	116.017	56.268	85.792	41.695	10.354	50.301	67.833	9.024	90.188	78.618	190.566
23	38.356	79.193	34.674	119.458	42.889	36.836	69.852	102.674	0.802	98.300	98.426	154.854
24	72.875	56.941	33.279	106.029	48.281	2181.785	63.453	51.277	3.704	98.926	98.926	93.398
25	73.139	84.421	13.113	106.848	40.892	1049.764	53.830	47.936	3.720	96.580	87.420	185.384
26	74.704	132.395	66.942	72.647	28.356	6608.573	60.731	58.849	4.122	97.200	92.600	192.954
27	74.390	116.520	22.732	76.573	46.137	34.654	71.432	93.699	2.286	99.738	95.306	239.896
28	56.422	106.909	23.416	100.355	30.381	1584.844	56.065	61.857	12.039	99.200	98.180	187.576
29	76.119	102.409	46.698	108.868	4.054	5763.957	64.466	59.615	2.312	99.644	98.852	174.390
30	36.642	105.972	3.910	78.997	42.949	1.778	58.626	73.565	2.997	98.116	95.266	205.826
31	30.923	78.356	2.132	39.587	1.570	2.155	60.424	75.027	11.268	80.340	80.050	155.452
32	7.408	153.920	91.750	50.153	11.924	8878.103	49.272	97.883	5.919	99.425	53.500	119.150
33	57.398	67.475	57.503	106.260	4.434	4133.919	55.833	57.446	4.272	98.804	94.866	155.000
34	59.094	124.002	72.521	98.299	1.923	3193.000	60.747	61.400	2.681	98.120	98.170	204.916
35	9.373	75.908	38.056	68.493	40.275	5500.652	54.138	56.523	4.284	96.038	94.366	114.860
36	61.615	70.712	42.329	85.201	2.936	3341.693	66.273	56.411	3.803	97.632	95.186	101.724
37	71.371	72.071	50.625	54.876	27.982	3803.797	59.645	63.930	3.351	97.360	81.866	110.810
38	76.288	72.293	71.972	78.007	44.498	7411.513	52.376	66.288	3.222	91.620	97.732	103.540
39	1.542	48.760	16.078	103.793	1.886	4411.869	66.642	81.703	12.336	100.000	100.000	101.580
40	41.059	67.183	63.376	83.185	21.002	4013.652	55.089	60.704	0.466	87.920	87.900	92.404
41	70.952	184.921	2.535	108.838	45.492	1016.977	59.459	59.712	6.461	99.000	95.600	439.806
42	74.766	89.703	16.711	72.966	50.026	29.476	32.762	64.437	9.118	99.374	96.244	193.366
43	74.559	71.199	19.926	95.444	47.200	2375.734	69.971	59.485	7.490	98.900	89.264	139.838
44	70.016	110.826	113.558	63.936	40.690	8816.956	53.457	67.045	2.731	80.200	80.200	124.485
45	58.076	114.927	83.658	93.835	45.707	30.654	55.504	60.865	5.087	98.900	96.860	136.880
46	27.978	94.072	79.010	52.533	41.536	4.410	58.845	56.875	9.082	94.986	20.620	128.990
47	27.525	114.159	107.003	104.272	2.837	8176.779	51.621	61.394	10.792	86.400	58.600	111.250
48	42.831	126.763	147.920	89.828	28.293	6.537	49.483	79.094	0.780	91.600	86.400	115.342
49	47.485	144.015	55.095	97.908	38.675	34.453	69.042	92.779	4.497	98.854	84.974	274.952
50	75.375	119.577	113.924	98.949	32.141	23.214	53.489	81.301	8.447	91.764	68.252	155.000
51	28.757	91.113	59.832	108.661	44.911	5811.979	56.676	75.102	17.388	99.460	96.140	114.558

Table 7. Inputs and ouputs for DEA.

The DEA analysis was run in the Excel application and under the general mathematical model of DEA. Table 8 shows the results that in addition to the efficient DMUs include the slack values of the technique, the latter are particularly important when carrying out decisions on the future performance of water utilities. The values of 1.00 in the box of "Eff. Score" correspond to the DMUS whose frontier value in one of the indicators presents it as an efficient frontier, that is, the efficiency reaches the value of 1, for the case of the DMUs with a score lower than 1 indicates " remote" that is of a border value.

DMU	Eff. score	DMU	Eff. score	DMU	Eff. score
1	1.00	18	1.00	35	1.00
2	1.00	19	0.92	36	1.00
3	1.00	20	0.96	37	1.00
4	0.96	21	1.00	38	1.00
5	1.00	22	1.00	39	1.00
6	0.96	23	1.00	40	1.00
7	1.00	24	1.00	41	1.00
8	1.00	25	1.00	42	1.00
9	1.00	26	1.00	43	1.00
10	1.00	27	1.00	44	0.92
11	1.00	28	0.9990	45	1.00
12	0.89	29	1.00	46	1.00
13	1.00	30	1.00	47	1.00
14	1.00	31	1.00	48	1.00
15	0.94	32	1.00	49	0.92
16	1.00	33	1.00	50	0.81
17	1.00	34	1.00	51	0.88

Table 8. Result DEA.

3 Conclusions

In this manuscript it has been studied how the main components of the indicators of water operators contributed in the essence of the behavior of the input variables for the measurement of the efficiency, in fact the study was fulfilled in the sense of collating with greater knowledge the influence of variables that affect the global performance of the operators of water.

Due to the efficiency results shown in Table 8, where a total of 11 non-efficient DMUs are observed out of a total of 51 DMUs studied, it can be deduced that in the field of water operators, the existence of several performance indicators can position a water operator agency as efficient even when a particular variable does not comply. In this sense, this work should help organizations operator of water in subsequent studies in which critical variables are involved to standardize at a national level.

References

[1] Augusto, et al., (2005). A performance assessment approach to the Pursuit of Competitiveness Excellence: Methodology and a Framework for Implementation. JGC Vol. 13 no. 1 & 2.

[2] Sanchez, M. L., Villareal, G., & Cabrera, M. (2008). Optimización Multicriterio por Análisis Envolvente de Datos. Ingenierias, 52-59.

[3] Buitrago, O., Palacio, O., Britto, R., & Adarme, W. (2016). Propuesta metodológica para la selección de la configuración de centros de distribución inmóticos utiliando análisis envolvente de datos. Ingeniare, 480-492.

[4] Cáceres, R. Á. (1995). Estadística Multivariante y no paramétrica con SPSS. Madrid: Días de Santos.

ISSN: 2455-6653

[5] *De la Garza García, J., & González Cavazos, B. A. (2013). Análisis Estadístico Multivariante. México: Mc. Graw Hill.*

[6] *González and Álvarez (2001). From efficiency measurement to efficiency improvement: The choice of a relevant benchmark. European Journal of Operational Research. Vol. 133 Issue 3. Pages 512-520*

[7] *Drake and Howcroft (1994). Relative efficiency in the branch network of a UK bank: An empirical study. Omega Vol. 22 Issue1. Pages 83-90*

[8] *Mendhi , T., & Soroosh , N. (2009). A new integrated DEA model for finding most BCC-eficient DMU. Applied Mathematical Modelling, 597-604.*

[9] *Triola, M. (2009) Estadística. Pearson*

[10] *Joro, T., & Korhonen, P. (2015). Extension of Data Envelopment Analysis with Preference Information. International Series in Operations Research & Management Science Vol 218. Springer*