

Evaluación de la eficiencia relativa de los grupos de investigación de las facultades de ingeniería de la Universidad Industrial de Santander, mediante el Análisis Envoltente de Datos (DEA), y su posterior clasificación.

Nathalia Acevedo Rojas, Johanna Pico Mendoza

Trabajo de grado para optar el título de Ingeniera Industrial

Director

Gabriel Yáñez Canal

Doctor en ciencias especialidad en matemática educativa

Universidad Industrial de Santander

Facultad de Ingenierías Fisicomecánicas

Escuela de Estudios Industriales y Empresariales

Bucaramanga

2018

*A Dios y a mi Iglesia por su infinita bondad y misericordia.
A mami por todo y por tanto que me ha aportado hasta hoy
A papi por todo el apoyo brindado
A mis hermanos: Javier, Leonardo y Mauri
porque siempre han estado ahí "incomodándome"
A la negrita, mi hermana, por siempre Estar
A Natha, por acompañarme en esta travesía llamada "trabajo de grado"*

Johanna

*A mi familia.
Nathalia.*

Agradecimientos

A Dios, pues “en Él están la sabiduría y el poder, y el consejo y el entendimiento son suyos”
Job 12: 13

A nuestras familias, por su apoyo incondicional.

Al profesor Gabriel, por ser tan maravillosa persona, por enseñarnos y guiarnos. Creemos firmemente que no existe docente en la universidad que hubiera podido dirigir este proyecto mejor que él.

A todas las personas que de una u otra forma nos han aportado durante esta etapa.

Tabla de Contenido

Introducción	15
1. Planteamiento del problema	19
2. Justificación	20
3. Objetivos.....	22
3.1 Objetivo general.....	22
3.2 Objetivos específicos	22
4. Revisión de la literatura	23
4.1 Análisis bibliométrico.....	23
4.2 Análisis preliminar de la literatura.....	27
4.2.1 Acerca de la eficiencia.	27
4.2.2 Desarrollo de los métodos para estimar la eficiencia.....	28
5. Marco de antecedentes.....	31
5.1 Colombia.....	34
5.2 Universidad Industrial de Santander.....	37
6. Marco teórico.....	40
6.1 Análisis Envolvente de Datos	41
6.2 Escala de producción	44
6.3 Modelo CCR.....	45
6.3.1 Conjunto de posibilidades de producción.....	45
6.3.2 Modelo Fraccional.	46
6.3.3 Modelo Lineal.....	48
6.3.4 Modelo Dual.	49
6.3.5 Modelo orientado a las salidas.....	50
6.3.6 Solución óptima CCR.....	51
6.4 Modelo BCC.....	54
6.4.1 Conjunto de posibilidades de producción.....	55
6.4.2 Modelo lineal y envolvente.....	56
6.4.3 Solución óptima BCC.....	57
6.5 Otros modelos DEA.....	59

6.5.1 Súper Eficiencia	60
6.5.2 Eficiencia cruzada.....	60
6.5.3 Líder Global.....	61
7. Metodología.....	62
7.1 Especificación del modelo	62
7.1.1 Selección de unidades de decisión (UD).....	62
7.1.2 Orientación del modelo.....	64
7.1.3 Selección de variables de entrada y salida.....	65
7.2 Recolección de la información.....	68
8. Resultados y su discusión	71
8.1 Análisis descriptivo de la información	71
8.1.1 Análisis por Escuelas.....	72
8.1.2 Análisis por grupos de investigación.....	79
8.2 Tratamiento de la información.....	88
8.3 Correlación bivariada.....	90
8.4 Análisis Clúster de las variables	95
8.5 Corrida del modelo	100
9. Análisis de resultados	103
9.1 Valores objetivo y mejoras potenciales	106
9.2 Clasificación de los Grupos de Investigación.....	113
9.2.1 Clasificación según Súper Eficiencia.....	113
9.2.2 Clasificación según Eficiencia Cruzada.....	116
9.2.3 Súper Eficiencia vs Eficiencia Cruzada.....	119
9.2.3.1 Análisis Clúster según puntuación de Eficiencia.....	123
9.2.4 Clasificación por Escuelas.....	128
9.2.5 DEA vs Colciencias.....	131
10. Conclusiones.....	137
11. Recomendaciones	141
Referencias Bibliográficas.....	143

Lista de tablas

Tabla 1 Cumplimiento de objetivos	18
Tabla 2 Modelo BCC orientado a las entradas y orientado a las salidas.	57
Tabla 3 Modelo BCC de dos etapas.....	58
Tabla 4 Programa matemático de Eficiencia Cruzada y Súper Eficiencia	61
Tabla 5 Grupos de investigación en la UIS por áreas del conocimiento	64
Tabla 6 Grupos de investigación en Santander por área de conocimiento	64
Tabla 7 Variables para la formulación del modelo DEA.....	70
Tabla 8 Cantidad de grupos de investigación por Escuela	72
Tabla 9 Grupos de Investigación por Escuela - Facultad Ingenierías Fisicomecánicas	73
Tabla 10 Grupos de Investigación por Escuela - Facultad Ingenierías Físicoquímicas	74
Tabla 11 Publicación según el tipo de revistas	77
Tabla 12 Estadísticos descriptivos de las variables	87
Tabla 13 Ponderación según el tipo de revista para cada publicación.....	89
Tabla 14 Puntaje asignado según el máximo título obtenido.	90
Tabla 15 Resultados del modelo DEA CCR-O.....	104
Tabla 16 Conjunto de referencia para los GI no eficientes.....	105
Tabla 17 Valores objetivo.....	108
Tabla 18 Mejoras potenciales	109
Tabla 19 Mejora potencial: Incremento en la producción de artículos.....	112
Tabla 20 Mejora potencial: Incremento en la dirección de Trabajos de grado.....	112
Tabla 21 Resultado según método de Súper Eficiencia.....	115
Tabla 22 Matriz Eficiencia Cruzada	117
Tabla 23 Resultados según método de Eficiencia Cruzada	118
Tabla 24 Análisis Clúster según resultados de CCR-O y Eficiencia Cruzada.....	123
Tabla 25 Agrupación de los grupos Eficientes	127
Tabla 26 Agrupación de los grupos Ineficientes.....	128
Tabla 27 Clasificación por Escuela (Facultad de Ingenierías Fisicomecánicas)	129
Tabla 28 Clasificación por Escuela (Facultad de Ingenierías Físicoquímicas)	130
Tabla 29 Resultados de la clasificación mediante DEA y Colciencias	132
Tabla 30 Categorías de clasificación modelos DEA.....	134
Tabla 31 Diferencias entre clasificaciones.	135

Lista de figuras

Figura 1. Publicaciones por año.....	26
Figura 2. Publicaciones por país.....	26
Figura 3. Frontera envolvente modelo CCR.....	43
Figura 4. Frontera envolvente CCR y BCC.....	55
Figura 5. Profesores vinculados a grupos de investigación por Escuela.....	75
Figura 6. Estudiantes vinculados a grupos de investigación por Escuela.....	76
Figura 7. Producción de artículos por Escuela en 2016.....	77
Figura 8. Docentes vinculados a cada grupo de investigación.....	80
Figura 9. Estudiantes vinculados a cada grupo de investigación.....	81
Figura 10. Dinero asignado a los grupos de investigación.....	82
Figura 11. Años de conformación de cada grupo de investigación.....	83
Figura 12. Producción de artículos por grupo de investigación en 2016.....	84
Figura 13. Trabajos de grado dirigidos a 2016.....	85
Figura 14. Dispersión cantidad de Artículos vs Presupuesto.....	91
Figura 15. Dispersión Tesis vs Presupuesto.....	92
Figura 16. Dispersión Artículos vs Años de conformación.....	92
Figura 17. Dispersión Tesis vs Años de conformación.....	93
Figura 18. Dispersión Artículos vs Estudiantes.....	93
Figura 19. Dispersión Tesis vs Estudiantes.....	94
Figura 20. Dispersión Artículos vs Docentes.....	94
Figura 21. Dispersión Tesis vs Docentes.....	95
Figura 22 Dendrograma - Clúster Insumos.....	97
Figura 23. Dendrograma - Clúster Productos.....	99
Figura 24. Distribución de los resultados de eficiencia de los GI.....	102
Figura 25. Frecuencia en conjuntos de referencia.....	106
Figura 26. Dispersión Eficiencia Cruzada Vs. Súper Eficiencia.....	120
Figura 27. Dispersión Eficiencia Cruzada Vs. Súper Eficiencia sin puntos extremos.....	121
Figura 28. Dispersión CCR-O vs. Eficiencia Cruzada.....	122
Figura 29. Dispersión CCR-O vs. Súper Eficiencia.....	122
Figura 30. Dendrograma - Clúster CCR-O vs. Eficiencia Cruzada.....	124
Figura 31. Dendrograma - Súper Eficiencia vs. Clúster Eficiencia Cruzada.....	125
Figura 32. Dendrograma - Clúster entidades Eficientes.....	126
Figura 33. Dendrograma - Clúster entidades Ineficientes.....	127

Lista de Apéndices

**(Ver apéndices adjuntos en el CD y pueden visualizarlos en la Base de Datos de la
Biblioteca UIS)**

Apéndice A. Variables de entrada y salida utilizadas en trabajos similares

Apéndice B. Clasificación de los grupos de investigación según Colciencias

Apéndice C. Valores observados y ponderados

Apéndice D. Conjunto de referencia y valores λ_j

Apéndice E. Valores objetivo

Apéndice F. Matriz Eficiencia Cruzada

Apéndice G. Artículo de resultados

RESUMEN

TÍTULO: EVALUACIÓN DE LA EFICIENCIA RELATIVA DE LOS GRUPOS DE INVESTIGACIÓN DE LAS FACULTADES DE INGENIERÍA DE LA UNIVERSIDAD INDUSTRIAL DE SANTANDER, MEDIANTE EL ANÁLISIS ENVOLVENTE DE DATOS (DEA), Y SU POSTERIOR CLASIFICACIÓN*

AUTOR: ACEVEDO ROJAS, Nathalia
PICO MENDOZA, Johanna**

PALABRAS CLAVE: ANÁLISIS ENVOLVENTE DE DATOS, DATA ENVELOPMENT ANALYSIS, DEA, EFICIENCIA, GRUPOS DE INVESTIGACIÓN

DESCRIPCIÓN:

En la presente investigación se aplica la metodología Análisis Envolverte de Datos como herramienta para la medición de la eficiencia de los grupos de investigación de las Facultades de Ingeniería de la Universidad Industrial de Santander. Este método no paramétrico utiliza la programación matemática para calcular la eficiencia relativa de varias unidades a partir de la razón entre las entradas y las salidas de un proceso.

Con el fin de identificar las entidades eficientes, es decir, aquellas que hacen un adecuado uso de los insumos que disponen, se definen en primera instancia el conjunto de variables que esquematizan la tecnología de producción de los Grupos de Investigación de la institución en cuestión. A continuación, se aplica el modelo DEA CCR (Charnes, Cooper & Rhodes), con orientación a las salidas, para el cálculo de la eficiencia.

Posteriormente, se exhibe la puntuación o grado de eficiencia obtenido por cada entidad observada, identificando de este modo los grupos que no alcanzaron el índice unitario de eficiencia, así como las mejoras potenciales necesarias para que puedan lograrlo. Teniendo en cuenta que los modelos DEA no permiten identificar cuál de los eficientes es el *mejor*, se presenta la clasificación de las entidades según los métodos Súper Eficiencia y Eficiencia Cruzada, los cuales se basan en los resultados arrojados por el DEA.

Finalmente, se hace una comparación entre la clasificación obtenida según los métodos Súper Eficiencia y Eficiencia Cruzada con la publicada por Colciencias en las convocatorias de 2015 y 2017, con el propósito de conocer si hay diferencia entre los métodos basados en el DEA y el propuesto por Colciencias.

* Proyecto de Grado

** Universidad Industrial de Santander, Facultad de Ingenierías Fisicomecánicas, Escuela de Estudios Industriales y Empresariales, Ingeniería Industrial. Director: Gabriel Yáñez Canal.

ABSTRACT

TITLE: RELATIVE EFFICIENCY EVALUATION OF ENGINEERING FACULTIES' RESEARCH GROUPS OF THE INDUSTRIAL UNIVERSITY OF SANTANDER BY DATA ENVELOPMENT ANALYSIS, AND ITS SUBSEQUENT CLASSIFICATION*

AUTHOR: ACEVEDO ROJAS, Nathalia
PICO MENDOZA, Johanna**

KEYWORDS: DATA ENVELOPMENT ANALYSIS, DEA, EFFICIENCY, RESEARCH GROUPS

DESCRIPTION:

In the present research is applied the methodology Data Envelopment Analysis as a tool to evaluate the efficiency of engineering faculties' research groups of the Industrial University of Santander. This no parametric method uses the mathematical programming to calculate the relative efficiency of many entities based on the ratio between inputs and output of a process.

With the purpose of identifying the efficient entities, that is, those who make an adequate use of the inputs at their disposal, it's defined firstly the set of variables that schematize the production technology of the research groups of the institution at hand. Following is applied the method DEA CCR (Charnes, Cooper & Rhodes), in its orientation to outputs, to calculate the efficiency.

Afterward, its exhibit the efficiency punctuation obtained by every observed entity, identifying in this way the groups that don't reached the unitary efficiency index, as well as the potential improvements they need. Finally, and taking into account that the traditional DEA methods don't identify which of the efficient entities is the best, is presented the ranking of these entities by Super Efficiency and Cross Efficiency methods, which are based on DEA results.

Finally, it's made a comparison between Super Efficiency and Cross Efficiency with Colciencias ranking, in the 2015 and 2017 call, with the purpose of know if there exist differences between DEA methods and the Colciencias' one.

* Bachelor Thesis.

** Industrial University of Santander, Physical and Mechanical Engineering Faculty, Industrial and Business Studies Department, Industrial Engineering. Direct by Gabriel Yáñez Canal

Introducción

El reto primario de las universidades es trabajar en pro del desarrollo de la sociedad y para lograrlo utilizan la investigación como herramienta fundamental. El buen manejo de dicha herramienta requiere que se fortalezca la cultura investigativa y la actividad grupal, de modo tal que un estudio que revele las razones organizativas que influyen en la eficiencia de la investigación, no solo no está de más, sino que es necesario, pues permitiría identificar los factores claves para impulsar el desempeño de los grupos de investigación y facilitar la definición de criterios para evaluar su alcance.

Ahora bien, de acuerdo con Colciencias “un grupo existe siempre y cuando demuestre producción de resultados tangibles y verificables, fruto de proyectos y otras actividades de investigación convenientemente expresadas en un plan de acción (proyectos) debidamente formalizado”. Entonces, podría decirse que estos grupos enfocan sus actividades a la generación de conocimiento, el cual servirá como punto de partida de futuras investigaciones que van a producir más conocimiento; lo que quiere decir que en el caso de estas entidades dicho conocimiento resulta ser tanto un insumo como un producto, y por consiguiente la eficiencia en su producción y la calidad de la misma deben ser aspectos medibles.

Es por lo anterior que el presente trabajo busca ofrecer un panorama general de los grupos de investigación de las facultades de ingeniería de la Universidad Industrial de Santander, al realizar la medición de su eficiencia relativa mediante el Análisis Envolvente de Datos (DEA por sus siglas en inglés) y su posterior clasificación.

Una vez obtenida una medida de eficiencia para cada grupo, se procede a compararlas y a identificar los factores críticos que determinan su desempeño. Se identifican los grupos de

investigación más eficientes, se analizan las mejoras potenciales requeridas en los grupos no eficientes y se determina cuánto y en qué factores una unidad que no alcanzó la eficiencia debe mejorar para alcanzarla.

Se opta por el modelo DEA, pues como lo refiere Murias (2003) “es la técnica para la evaluación de la eficiencia más utilizada en el sector educativo” –aseveración que pudo comprobarse una vez realizado el análisis bibliométrico que se presenta en los apartados posteriores–. Además de ello, es una metodología que cuenta con múltiples ventajas en lo que refiere al manejo de los datos, como lo son el no requerir asignar *pesos* a cada entrada y salida según el criterio personal de los expertos en el tema, y tampoco establecer una forma específica para la función de producción, como sí se hace necesario en los enfoques de regresión estadística. Pero aún más importante, es que el Análisis Envolvente de Datos permite considerar múltiples entradas y salidas con distintas unidades de medida, lo que resulta ser bastante adecuado para el trabajo que se propone.

Cabe resaltar que el análisis del presente trabajo se limita a los grupos de investigación del área de las ingenierías ya que el DEA requiere que las unidades productivas (llamadas Unidades de Decisión por la terminología del método) cumplan con la característica de homogeneidad, por lo cual dicha facultad resulta ser ideal para este estudio y además cuenta con suficientes entidades para realizar el respectivo análisis, pues está conformada por un amplio número de grupos de investigación con aspectos comunes.

El estudio se inicia con una revisión de la literatura (Apartado 4) con el fin de identificar de qué manera se ha abordado el tema de la estimación de la eficiencia a nivel global y adicionalmente para conocer cómo ha venido evolucionando la estimación de la misma. Posteriormente, en el Capítulo 5, se lista y profundiza un poco a cerca de las diferentes aplicaciones que se han hecho de la metodología DEA en el sector educativo a nivel mundial, nacional e institucional.

En el Apartado 6 se ahonda en el aspecto teórico del Análisis Envolvente de Datos y se desglosan los dos modelos que con mayor frecuencia son utilizados, tal es el caso de aquellos propuestos por Charnes, Cooper & Rhodes (CCR) y Banker, Charnes & Cooper (BCC). Ahí mismo, se mencionan métodos que se basan en el DEA para clasificar las entidades evaluadas (Súper Eficiencia y Eficiencia Cruzada). En el siguiente Capítulo, se describe el proceso que se llevó a cabo para la aplicación del DEA, es decir, la metodología aplicada. A continuación, en los Capítulos 8 y 9, se presenta la discusión de los resultados y el análisis de los mismos respectivamente.

Finalmente, se presentan las conclusiones de este estudio, cuya principal resulta ser la obtención de doce entidades eficientes de un total de 31 evaluadas, es decir, que el 38,7% fue catalogado como eficiente. Por otro lado, la elección de la clasificación (ranking) se deja a consideración de los interesados.

Cumplimiento de objetivos

Tabla 1
Cumplimiento de objetivos

Objetivos específicos	Cumplimiento
Revisar la bibliografía correspondiente a la aplicación del Análisis Envolvente de Datos en el ámbito académico, como herramienta para la evaluación de la eficiencia.	Apartado 4 Apartado 5
Definir las variables de entrada y salida a incluir en el estudio y describir la metodología a utilizar en la recolección de información necesaria para la construcción de las mismas.	Apartado 7 Apartado 8
Especificar la formulación del modelo DEA más apropiado para el análisis de la eficiencia de los grupos de investigación de las facultades de ingenierías.	Apartado 8.5
Aplicar el modelo de Análisis Envolvente de Datos y, con base en los resultados, realizar una clasificación de los grupos de investigación.	Apartado 9
Determinar las prácticas de aquellos grupos que no alcanzaron la frontera de eficiencia y compararlas con las de aquellos que sí la alcanzaron.	Apartado 9
Cuantificar las mejoras potenciales que deben realizar las unidades que no alcanzaron la frontera de eficiencia, para que puedan lograr un 100% de eficiencia productiva.	Apartado 9
Elaborar un artículo publicable.	Apéndice G

1. Planteamiento del problema

En cuanto a la investigación en las instituciones de educación superior, el Ministerio de Educación (MEN) afirma que:

Tan importante como una planta física adecuada, un excelente programa académico y un buen grupo de profesores es pues la existencia de grupos y líneas de investigación para constituirse como universidad. Por eso mismo, en los actuales procesos de acreditación que se están dando en la educación superior, el ítem de la investigación resulta ser uno de los más importantes.

Así mismo, Arteaga y Cruz (1999) exponen que “una de las funciones esenciales de las universidades modernas es la construcción del nuevo conocimiento. Dichas instituciones promueven en tal sentido espacios de investigación, donde los diversos grupos académicos participan en la generación de nuevo conocimiento”.

Ahora bien, de acuerdo con cifras publicadas por la Red de Indicadores de Ciencia y Tecnología¹ –Iberoamericana e Interamericana–, en Colombia, para el año 2015, los porcentajes de investigadores por sector de ocupación Gobierno, Empresas, Educación Superior y Organizaciones privadas sin ánimo de lucro representan el 1.64%, 1.19%, 89.65% y 7.51%, respectivamente, demostrándose así la importante participación del sector educativo, específicamente de las universidades, en el ámbito investigativo de nuestro país.

Si bien la universidad dentro de sus políticas de investigación busca fortalecer y estimular la actividad investigativa a través de diversas estrategias, ¿cómo establecer qué tan eficiente está siendo la institución en este sentido?, es por lo anterior que surge el interés por conocer qué tan

¹ Red de Indicadores de Ciencia y Tecnología (RICYT), es la institución encargada de promover, en un marco de cooperación internacional, el desarrollo de instrumentos e indicadores para la medición y análisis de la ciencia y la tecnología en Iberoamérica

eficiente están siendo los grupos de investigación, pues refleja –de cierta forma– el grado de excelencia de la Universidad Industrial de Santander como ente educativo, si se comparte la noción de que “la calidad de la educación superior está íntimamente asociada con la práctica de la investigación” (Restrepo, s.f).

2. Justificación

Según la Organización para la cooperación y el Desarrollo Económico (OCDE, 2006) “la investigación es el trabajo creativo llevado a cabo de forma sistemática para incrementar el volumen de los conocimientos humanos, culturales y sociales y el uso de esos conocimientos para derivar nuevas aplicaciones”. Por otro lado, los grupos de investigación se conocen como el “conjunto de personas que interactúan para investigar y generar productos de conocimiento en uno o varios temas, de acuerdo con un plan de trabajo de corto, mediano o largo plazo (tendiente a la solución de un problema)” (Colciencias, 2014). Según los resultados de la medición –publicada por Colciencias en el año 2015– de grupos de investigación, se evidencia un incremento de la productividad y la calidad científica del país. Es decir, la calidad de los grupos de investigación y de los investigadores en Colombia va en aumento, así como la participación regional en la producción científica. Por tal razón, se considera importante dedicar esfuerzos a fortalecer los grupos de investigación con los que cuenta la universidad, reforzando su autonomía y llevando a cabo un seguimiento estricto de los productos que resultan de su operación. Si adicional a esto se tiene en cuenta que la intención de la Universidad Industrial de Santander para el año 2018 es fortalecer su carácter público, aportando al desarrollo político, cultural, social y económico del

país, como resultado de un proceso de generación y adecuación de conocimiento en el cual la investigación constituye el eje articulador de sus funciones misionales, se ratifica la importancia de realizar un análisis de los grupos de investigación con base en las políticas de investigación de la institución, las cuales se encuentran articuladas alrededor de cuatro lineamientos: investigación orientada por programas; fortalecimiento de la actividad investigativa; articulación con el entorno y apropiación social del conocimiento.

Es por lo anterior que surgió el interés por realizar la medición de la eficiencia de los grupos de investigación de la Universidad Industrial de Santander. Para ello se considera adecuado utilizar el Análisis Envolvente de Datos (Data Envelopment Analysis, DEA), un método que calcula eficiencia relativa de varias unidades similares a partir de la razón entre las entradas y las salidas de un proceso.

3. Objetivos

3.1 Objetivo general

Evaluar la eficiencia relativa de los grupos de investigación de las facultades de ingeniería de la Universidad Industrial de Santander, mediante el Análisis Envolvente de Datos (DEA).

3.2 Objetivos específicos

- Revisar la bibliografía correspondiente a la aplicación del Análisis Envolvente de Datos en el ámbito académico, como herramienta para la evaluación de la eficiencia.
- Definir las variables de entrada y salida a incluir en el estudio y describir la metodología a utilizar en la recolección de información necesaria para la construcción de las mismas.
- Especificar la formulación del modelo DEA más apropiado para el análisis de la eficiencia de los grupos de investigación de las facultades de ingenierías.
- Aplicar el modelo de Análisis Envolvente de Datos y, con base en los resultados, realizar una clasificación de los grupos de investigación.
- Determinar las prácticas de aquellos grupos que no alcanzaron la frontera de eficiencia y compararlas con las de aquellos que sí la alcanzaron.
- Cuantificar las mejoras potenciales que deben realizar las unidades que no alcanzaron la frontera de eficiencia, para que puedan lograr un 100% de eficiencia productiva.
- Elaborar un artículo publicable.

4. Revisión de la literatura

Pautasso (2013) afirma que para realizar una adecuada revisión de la literatura se requiere que el investigador realice una búsqueda y evaluación de información relevante. Sin embargo, la identificación de dicha información relevante no es algo tan sencillo si se tiene en cuenta que, adicional a la información científica ya existente, continuamente aparecen nuevas publicaciones (Gutiérrez y Maz, 2004).

Luego de haber determinado con exactitud que el asunto de interés que se desea abordar en el presente trabajo es la medición de la eficiencia de los grupos de investigación de la universidad, se ha procedido a realizar una revisión y análisis de la literatura con el fin de recopilar información relevante y de impacto que enriquezca este estudio.

4.1 Análisis bibliométrico

El objeto de estudio de esta investigación es evaluar la eficiencia relativa de los grupos de investigación de las facultades de ingeniería de la Universidad Industrial de Santander, mediante el Análisis Envolvente de Datos (DEA), con el propósito de conocer el grado de eficiencia en la que se encuentra cada uno de éstos. Para lo anterior se hace necesario realizar un análisis de la literatura en bases de datos de carácter científico con el fin de analizar de qué manera se ha abordado el tema a nivel global, conocer los parámetros utilizados a la hora de seleccionar el modelo más apropiado, las variables a considerar en el estudio, la recolección de los datos, así como las diferentes técnicas estadísticas adicionales utilizadas por los investigadores para complementar el estudio.

Para la presente investigación se tendrán en cuenta fuentes bibliográficas que hagan referencia a la medición de la eficiencia en el sector educativo (universidades, departamentos universitarios, grupos de investigación, colegios, escuelas, entre otros) por medio de la metodología Análisis Envoltante de Datos a nivel mundial y que hayan sido publicadas durante los últimos veinte años (1998-2017). Para la realización de la revisión de la literatura se llevaron a cabo las siguientes etapas: identificación, selección y análisis de la bibliografía encontrada.

La ecuación de búsqueda presentada a continuación tiene como objetivo identificar aquellas fuentes que aporten a la presente investigación. La búsqueda de las fuentes de interés se llevó a cabo por medio de la base de datos *Scopus*.

(TITLE-ABS-KEY ("data envelopment analysis") OR TITLE-ABS-KEY (dea) AND TITLE-ABS-KEY (efficienc*) AND TITLE-ABS-KEY (assess*) OR TITLE-ABS-KEY (measur*) AND TITLE-ABS-KEY ("higher education") OR TITLE-ABS-KEY (education) OR TITLE-ABS-KEY (research))

Para la identificación de los documentos pertinentes se tuvieron en cuentas los siguientes criterios de inclusión y exclusión:

Inclusión:

1. Se toman los documentos publicados en los últimos veinte años, es decir, en el periodo de 1998-2017.
2. Solo se tendrán en cuenta documentos de tipo artículos, libros y capítulos de libros.

Exclusión

1. No se tendrán en cuenta documentos publicados en idiomas diferentes al inglés y el español.

2. Se excluirán tópicos que no estén relacionados a la medición de la eficiencia en el sector educativo.

Con la ecuación de búsqueda propuesta se encontró un total de 1.954 documentos comprendidos entre los años 1983-2018. Al aplicar el primer criterio de inclusión la cantidad de dichos documentos se redujo a 1.860, posteriormente, al aplicar el segundo criterio esta cantidad disminuye a 1.500. Finalmente, aplicando los dos criterios de exclusión a la búsqueda se obtuvo un total de 1.430 documentos.

La selección de los documentos se hizo de acuerdo al aporte que éstos hacen al tema objeto de estudio, para ello se ordenaron los resultados teniendo en cuenta la cantidad de citas de cada una de éstas, es decir, con base en el prestigio que tales documentos tienen en las bases de datos científicas.

En cuanto al análisis de la bibliografía encontrada, y tomando como base los resultados obtenidos a través de la base de datos *Scopus*, se encuentra que (ver Figura 1) las publicaciones por año mantuvieron un comportamiento casi constante durante el periodo 2001-2005 y a partir del año 2010 éstas han venido en aumento, es decir, que la aplicación de la metodología DEA para medir la eficiencia en el sector educativo ha ganado auge en los últimos años.



Figura 1. Publicaciones por año. Adaptado de Scopus (2017).

En la Figura 2, se evidencian los países con más aporte (Estados Unidos 18,76%, China 10,7% y Taiwán 10,28%). Por otro lado, se puede apreciar que la participación de los países latinoamericanos no es tan representativa, siendo Brasil (2,94%) el país latino que mayor participación ha tenido. En cuanto a Colombia se evidencia un aporte no muy llamativo, solo un 0,28%, esto es 4 artículos.

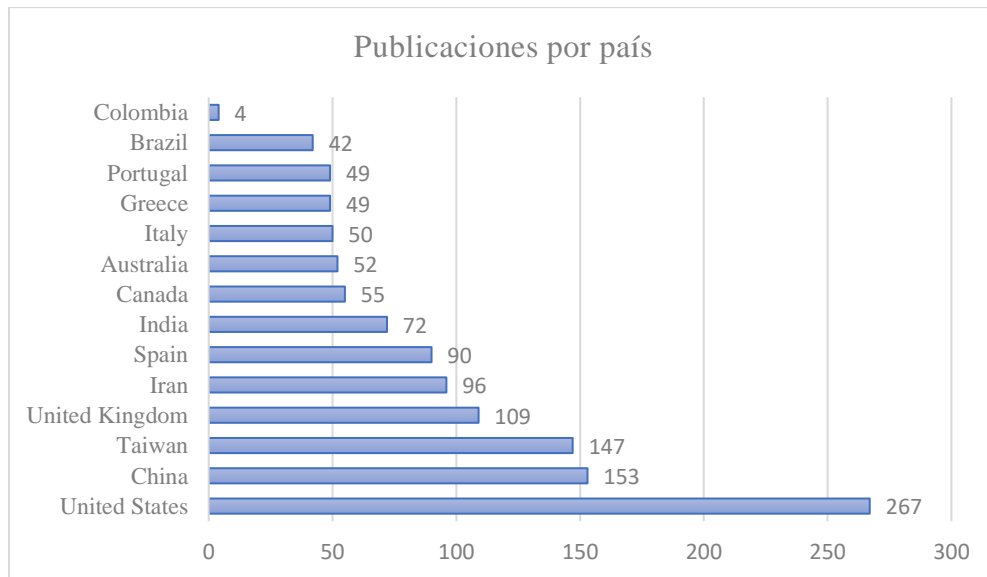


Figura 2. Publicaciones por país. Adaptado de Scopus (2017).

4.2 Análisis preliminar de la literatura

4.2.1 Acerca de la eficiencia. Hoy en día es común escuchar el término *eficiencia*, pues ha venido ganando popularidad a lo largo de los años. Sin embargo, éste suele confundirse con el de *productividad* y al momento de definir estos dos términos la tendencia es a la unificación de los mismos. En la literatura se encuentra que diversos autores definen los dos términos sin resaltar ningún tipo de diferencia entre ellos; tal es el caso de Sengupta y Cooper, Seiford & Tone (como se citó en Daraio & Simar, 2007) quienes definen productividad y eficiencia como la relación entre las salidas (outputs) y las entradas (inputs).

Si bien es cierto que puede generarse confusión al momento de definir *productividad* y *eficiencia*, “esto es desafortunado porque no son precisamente lo mismo” (Coelli, Rao, O’Donnell & Battese, 2005, p.3), por lo que se hace necesario exponer la definición de los dos términos. Vicent (como se citó en Daraio & Simar, 2007) define *productividad* como la razón entre un producto y los componentes que lo hacen posible. Para Lovell (1993) es “la relación entre la producción y la entrada de una unidad productiva” (p.3). En el caso de *eficiencia*, cabe resaltar que fue gracias al surgimiento de la microeconomía moderna (Marshall, 1898) que se desarrolló formalmente el concepto. Así, Farrell (1957) afirma que “cuando se habla de eficiencia, generalmente hace referencia al éxito que se tiene en producir lo más posible un producto con un conjunto dado de insumos” (p.254). Lovell (1993) se refiere a ésta como “la comparación entre los valores observados de los óptimos de las salidas y las entradas de una unidad de producción” (p.3).

Con el paso del tiempo, el concepto de *eficiencia*, ha sufrido transformaciones y clasificaciones, como los son: eficiencia técnica, eficiencia de escala y eficiencia asignativa. Koopmans (1951) distingue la *eficiencia técnica* afirmando que una unidad productiva es técnicamente eficiente si,

y sólo si, aumentar cualquier salida o disminuir cualquier entrada es posible solo si se disminuye alguna otra salida o se aumenta alguna otra entrada. Posteriormente Farrell (1957) plantea la necesidad de manejar la anterior definición de *eficiencia técnica* estableciendo que ésta depende de la mejor práctica observada en el grupo de comparación.

Álvarez (2013) afirma que la *eficiencia de escala* se da “cuando se está produciendo en una escala de tamaño óptima, que es la que le permite maximizar el beneficio” (p.20), es decir, que la eficiencia de una unidad productiva podría mejorarse manteniendo la misma combinación de entradas y cambiando su escala; entonces dicha unidad será eficiente en escala cuando el tamaño de sus operaciones sea óptimo –obteniendo la máxima productividad–, de modo que cualquier cambio en la escala de operación disminuiría su eficiencia.

Ahora, en cuanto a la *eficiencia asignativa*, Daraio & Simar (2007) afirman que “mide el éxito de una empresa al elegir un conjunto óptimo de insumos con un conjunto dado de precios de los insumos” (p.15). Estos datos también son fundamentales en la medición de la eficiencia pues permite desarrollar medidas de desempeño para realizar una selección óptima de la proporción de insumos a utilizar con el fin de producir determinada cantidad de salidas al mínimo costo.

4.2.2 Desarrollo de los métodos para estimar la eficiencia. Podría decirse que el inicio de la búsqueda de técnicas para medir la eficiencia de unidades productivas se dio por Frederick Taylor en el siglo XIX, como lo afirma Kanigel (1997) al referirse a esta época:

Frederick Winslow Taylor, con tan sólo veinticinco años cuando comenzó sus experimentos, fue el primer experto en eficiencia del mundo. Su búsqueda incesante de “la mejor manera” – la frase que fue asociada a él y a todo el movimiento de eficiencia– influyó en Henry Ford y Lenin, y motivó al gurú de la administración Peter F. Drucker para que lo incluyera junto con Freud y Darwin, como uno de los tres personajes más influyentes del mundo moderno. (p.1)

Ahora bien, el impacto del movimiento de la eficiencia en el siglo XX se dio a tal punto que, a mediados del mismo, suscitó un cambio de gran relevancia en las técnicas existentes hasta ese entonces, con autores como Debreu (1951) y su coeficiente de utilización de recursos que resulta siendo la primer medida de la eficiencia productiva “que busca la máxima reducción factible de las variables de entrada, o bien el incremento máximo de las salidas; con un inconveniente, y es que hace posible la aparición de holguras en los insumos o excedentes de productos” (Daraio & Simar, 2007, p.14). Y el aporte hecho por Koopmans (1951) a la definición de la eficiencia que se menciona en el apartado anterior.

Pero no fue hasta que Farrell (1957) aplica de forma empírica los trabajos realizados por Koopmans (1951) y Debreu (1951), que se vislumbra la otra cara de las técnicas para medir la eficiencia que hasta ese entonces se utilizaban en tal propósito –basados principalmente en estimaciones con Mínimos Cuadrados Ordinarios– (Díez, 2007).

En base a tales estudios, Farrell concibe la eficiencia productiva como la combinación de la eficiencia técnica y la eficiencia asignativa. La noción original se sitúa en un enfoque orientado a las entradas, es decir, en la búsqueda de la mínima utilización de insumos que produzcan determinada cantidad de productos. El autor ilustra el escenario en torno a unidades productivas que utilizan dos entradas (x_1 y x_2) para obtener una única salida (y), de modo que fuera posible la construcción de un gráfico de dos dimensiones, presumiendo rendimientos a escala constante y el conocimiento total de la función producción, que le permitieran representar la información del proceso productivo con una sola isocuanta² y así distinguir el componente técnico y asignativo de la eficiencia. La función de producción se estimó a partir de la observación de una muestra de empresas de la industria en cuestión, consideradas plenamente eficientes, para así dar forma a la

² Representación gráfica que identifica las distintas combinaciones de las entradas que una entidad eficiente podría usar para producir determinada cantidad de salida.

isocuanta. Finalmente, aquellas unidades productivas que estuvieran ubicadas sobre la curva de la isocuanta serían consideradas como eficientes, y en contraste las que se ubicaran por debajo de ella como ineficientes; esto último resultaba ser una idea inusitada, pues hasta ese momento no se admitía la posibilidad de catalogar una operación como ineficiente.

Más tarde, Farrell también examinó la extensión del modelo a más de dos entradas y salidas, al igual que un planteamiento con rendimientos de escala variable.

Ahora bien, el verdadero problema en la práctica resulta ser la estimación de la isocuanta, es decir, conseguir la función de producción. Para ello, Farrell enuncia dos caminos a seguir para obtener el valor estimado de ésta, ya sea utilizando métodos paramétricos o haciéndolo a través de métodos no paramétricos. Por supuesto, en esa época los investigadores se inclinaban por la primera opción ofrecida pues se encontraba en un ámbito más explorado la metodología estadística.

El análisis de la eficiencia productiva había cautivado el interés de M. J. Farrell por el desarrollo de un modelo que permitiera su medición, y es en 1957 con su trabajo *“The measurement of productive efficiency”* donde propone una herramienta para lograrlo; por una parte introduce la técnica de la frontera de eficiencia (isocuanta) para realizar la comparación de unidades productivas; y por otra, impulsa la investigación que busca resolver el problema de estimación de la función de producción a través de dos corrientes que serían de interés en años posteriores, los métodos paramétricos y no paramétricos.

Es cierto que el enfoque que presentó Farrell en ese entonces pasó desapercibido hasta el año 1978 cuando A. Charnes, W. W. Cooper y E. Rhodes (CCR), tomando como inspiración los aspectos abordados por Farrell, publican su trabajo conocido como *“Measuring the efficiency of decision making units”*, producto de la tesis de investigación de Rhodes que buscaba evaluar la

eficiencia relativa –y comparar las prácticas– de un conjunto de colegios perteneciente a un programa educativo de Estados Unidos para estudiantes desventajados; el reto era realizar tal estimación considerando múltiples entradas y salidas, y por supuesto sin contar con información acerca de precios (dado el contexto de la investigación).

Aunque la investigación realizada por CCR abarcó en gran medida los planteamientos realizados por Farrell, se distinguió al utilizar técnicas de programación lineal que resultaron en un modelo superior al propuesto por Farrell con la isocuanta, y que sería conocido como Análisis Envolvente de Datos. Posteriormente, surgieron otros modelos que contemplarían características distintas como los rendimientos variables a escala, propuesto en 1984 por Banker, Charnes & Cooper (BCC); un *Modelo Aditivo* (Charnes *et al.*, 1985) que lidia directamente con las holguras, es decir, el exceso en los insumos y/o la escasez de productos, el cual permite discriminar las entidades evaluadas entre eficientes e ineficientes por la existencia de dichas holguras. De igual forma, se introduce el modelo *Slack-Based Measure* (SBM), el cual es definido como “un producto de las ineficiencias de insumos y productos” (Tone, 2001) y resulta ser similar a los modelos CCR y BCC, solo que éste, al igual que en el *Modelo Aditivo*, trata las holguras de manera directa.

5. Marco de antecedentes

Ante el interés de las entidades tanto públicas como privadas por conocer qué tan bien están administrando sus recursos a través de los cuales producen determinados productos, nace la necesidad de medir el grado de eficiencia que éstas poseen. A lo largo de los años diferentes metodologías han venido surgiendo para tal propósito, entre ellos se encuentra el Análisis

Envolvente de Datos el cual según Murias (2003) “es la técnica no paramétrica más aplicada”, especialmente en el sector público.

A nivel internacional se han realizado diversos estudios con el objetivo de medir la eficiencia en el sector educativo, entre ellos encontramos a (Abbott & Doucouliagos, 2003; Johnes, 2006; Johnes & Yu, 2008; Katharaki & Katharakis, 2010; Tse & Yew, 2011 y Sagarra, Mar-Molinero & Agasisti, 2017) quienes comparten el objetivo de la investigación el cual busca medir la eficiencia de las universidades de Australia, Inglaterra, China, Grecia, Malasia y México respectivamente, a través de la metodología DEA.

Korhonen, Tainio & Wallenius (2001), llevaron a cabo un estudio para evaluar la actividad investigativa a partir del cual pudiera identificarse una de las mejores formas de utilizar los recursos económicos de los que disponían los grupos de investigación en la Escuela de Economía de Helsinki, Finlandia. El estudio evaluó 18 grupos de investigación pertenecientes a la Escuela de Economía de Helsinki, con la metodología DEA, manejando un modelo BCC y posteriormente en el análisis utilizan el sistema Pareto Race para comparar las puntuaciones de eficiencia de las unidades ineficientes con las unidades que tienen el mismo valor que la solución más preferida.

En España Pino, Solís, Delgado y Barea (2010) plantearon como objeto de su investigación, evaluar la eficiencia de los grupos de investigación de Andalucía-España, mediante DEA. Para la realización de dicho estudio, seleccionaron el modelo CCR orientado a las salidas y utilizaron el software *Cplex*³. Como Unidades de Toma de Decisión (DMU) eligieron un total de 86 grupos de investigación de las áreas de tecnologías de la información y la comunicación (TIC) pertenecientes a las universidades y al Consejo Superior de Investigaciones Científicas CSIC de Andalucía. En

³ Cplex es un software de IBM que sirve como solucionador de programación matemática de alto rendimiento para programación lineal, programación entera mixta y programación cuadrática.

cuanto a las variables, teniendo en cuenta que el número propuesto para la realización del estudio era muy extenso, aplicaron un análisis factorial para reducir la dimensión de las mismas, y regresión lineal y correlación canónica con el propósito de identificar si había relación entre ellas. Una vez realizado el anterior análisis la cantidad fue reducida a un total de tres variables de entrada (número de doctores activos, resto del personal de I+D+i⁴ y subvenciones recibidas en la convocatoria de ayudas a los grupos de investigación) y cinco variables de salida (ingresos por proyectos de financiación pública, publicación en revistas ISI, número de tesis doctoral, total otros proyectos y número de contratos total).

Entre otros trabajos realizados a nivel mundial cuyo objetivo es la medición de la eficiencia usando DEA, hay que destacar el realizado por Galache *et al.* (2010) en las universidades públicas de Andalucía, España, titulado “Análisis de la eficiencia de los grupos de investigación por ramas de conocimiento”. Los autores se plantearon la necesidad de evaluar la eficiencia técnica de los 2.073 grupos de investigación que hacen parte de las 9 universidades públicas de Andalucía. Para ello definen las variables de entrada en función de los recursos humanos requeridos en la actividad investigativa (Número de profesores con doctorados, sin doctorados, personal técnico de apoyo, becarios, personal auxiliar y colaboradores); y consideran dentro de las variables de salida los indicadores de producción comúnmente evaluados en este tipo de entidades, como lo son las publicaciones realizadas en revistas, libros y capítulos, congresos, tesis, patentes, número y cuantía de proyectos, patentes y contactos, y otras actividades. Los autores no sólo se quedan con la clasificación de eficiencia de los grupos de investigación, sino que van más allá y utilizan la información obtenida tanto de las unidades eficientes como de las ineficientes para, además de indicar las fortalezas y debilidades de cada uno junto con las posibilidades de mejora, realizar una

⁴ Investigación, desarrollo e innovación.

agrupación de unidades similares que les permite visualizar aspectos que caracterizan su clasificación a través de un análisis clúster.

En Argentina Alberto, Carignano y Ercole (2010), por medio de los modelos CCR, BCC y eficiencia cruzada, buscan analizar la eficiencia de las universidades públicas de ese país. Previo a la aplicación de la metodología, aplican un análisis clúster jerárquico para clasificar las instituciones en tres categorías (grandes, medianas y pequeñas) para luego sí llevar a cabo la medición de eficiencia de las mismas. En su investigación los autores destacan la importancia de la etapa de recolección de la información, pues afirman que los resultados a obtener dependen de la calidad de la información de que se disponga.

5.1 Colombia

A nivel nacional Arenas, Soto y Rivera (2004) y Giraldo (2007) realizan la misma investigación en diferentes periodos de tiempo con el objetivo de evaluar los grupos de investigación de la Universidad Tecnológica de Pereira según los indicadores de eficiencia de Colciencias versus su evaluación a partir del Análisis Envoltante de Datos. En las dos investigaciones los autores analizan la metodología utilizada por Colciencias para medir la eficiencia de la investigación científica en cuanto a la generación de nuevo conocimiento, formación de investigadores y extensión de los resultados de la investigación. Posteriormente, aplican DEA y comparan los resultados obtenidos por las dos metodologías encontrándose algunas diferencias en los resultados obtenidos en lo relacionado a puntuación y posición. Por su parte, Restrepo y Villegas (2007) clasifican los grupos de investigación de la facultad de ingeniería de la Universidad de Antioquia aplicando DEA a aquellos grupos que pertenecían a las categorías A, B o C de Colciencias en el año 2006 por medio de los modelos CCR orientado a las salidas, Eficiencia Cruzada y Súper

Eficiencia. Adicionalmente llevan a cabo un análisis de correlación entre las variables de entrada y salida seleccionadas.

En la Universidad Nacional de Colombia, Rojas (2010) realiza una clasificación de los grupos de investigación de la facultad de ingeniería de esa universidad mediante la estimación de la eficiencia técnica utilizando DEA; posteriormente realiza una comparación con la clasificación dada por Colciencias en la convocatoria del año 2008; esto último determina que la evaluación se realice con la información comprendida en el periodo de tiempo que utilizó esta última entidad. Dicha comparación determina que la selección de las variables de salida sea acorde a los productos que mide el índice utilizado por Colciencias al realizar la clasificación; de ese modo el autor realiza la evaluación de los 34 grupos de investigación correspondientes a la facultad de ingeniería según la producción de cada uno de ellos, medida en los índices de nuevo conocimiento, nuevo conocimiento tipo A, índice de formación y el índice de divulgación; considerando una única entrada al modelo, el tamaño de los grupos –medido con el número de integrantes adscritos sin considerar su título dentro del mismo–. Marlon Rojas utiliza el DEA orientado a maximizar las salidas del proceso, con el modelo CCR de rendimientos constantes a escala para distinguir entre unidades eficientes e ineficientes, y el método de eficiencia cruzada para realizar la posterior clasificación de los grupos de investigación. Finalmente, para contrastar la clasificación realizada por Colciencias de la forma A1, A, B, C y D el autor utiliza el análisis clúster a fin de realizar agrupaciones que se acomoden a esas categorías.

Visbal, Mendoza y Causado (2016) en su investigación titulada “Eficiencia en las instituciones de educación superior públicas colombianas: una aplicación de Análisis Envolvente de Datos” llevan a cabo un análisis de la eficiencia de cada universidad durante el año 2011 y adicional a ello realizan una descomposición de las ineficiencias globales encontradas en: ineficiencia técnica,

administrativa, de escala y de mezcla, con el fin de identificar las mejoras potenciales que deberían realizar aquellas instituciones que no lograron alcanzar la frontera de eficiencia. Para este estudio, los autores optaron por aplicar los modelos BCC, CCR y SBM con orientación a las salidas, con las siguientes variables de entrada (Docentes tiempo completo equivalente, Gastos en personal administrativo, Recursos financieros, Recursos físicos) y de salida (Matrícula de posgrado, Saber Pro, Revistas indexadas, Artículos en revistas indexadas). Producto de esta investigación los autores encontraron que, de las 32 instituciones analizadas, sólo 15 de éstas no presentan ineficiencia alguna, por otro lado, la cobertura en posgrados es un aspecto crítico en las universidades públicas del país, ya que éste debería aumentarse 75% en promedio, es decir, ofrecer nuevos programas especialmente de maestría y doctorado. Adicionalmente las revistas indexadas y los artículos de revistas indexadas son los aspectos a los que más deben prestar atención aquellas instituciones que fueron clasificadas como no eficientes, pues son los que más mejoras necesitan.

Posteriormente Visbal, Martínez & Guijarro (2017) realizan una investigación similar a la anterior titulada “Assessing the Efficiency of Public Universities through DEA. A Case Study” en donde analizan la eficiencia de las universidades públicas de Colombia durante el año 2012 a través del Análisis Envolvente de Datos. Para tal fin utilizan los modelos CCR, BCC y SBM orientados a las salidas, realizan una clasificación adicional de las universidades por medio del modelo de eficiencia cruzada y por último analizan los cambios de productividad en el periodo 2011-2012 a través de los índices de Malmquist. A través de esta investigación encontraron que 18 de las 32 universidades analizadas resultaron ser totalmente eficientes, es decir, no presentan algún tipo de ineficiencia. Por otro lado, los resultados obtenidos reiteran la necesidad que tienen aquellas instituciones que no alcanzaron la eficiencia en aumentar la cantidad de estudiantes matriculados

en los programas tanto de pregrado como de posgrado; aumentar la cantidad de revistas indexadas y los artículos publicados en las revistas indexadas, así como el número de docentes con movilidad.

Por otro lado, Aristizábal, Castro, Echeverri & Valencia (2017) aplican la metodología DEA a través de la cual buscan medir la eficiencia de la producción científica del país a través de su investigación titulada “Assessing the efficiency of science, technology and innovation using Data Envelopment Analysis (DEA): The case of Colombia”, en donde toman como unidades de medida los departamentos que conforman el territorio nacional junto con el distrito capital Bogotá DC. En su estudio aplican el modelo CCR orientado a las salidas y adicionalmente el modelo *Assurance Region (AR)*, el cual toma como referencia al primer modelo. Como resultado de esta investigación los autores encontraron que el departamento de Boyacá se destaca como el líder global ya que obtiene la calificación de *eficiente* en los dos modelos empleados y es el que más veces es tenido como referencia por los departamentos clasificados como no eficientes. Adicionalmente los resultados de los dos modelos coinciden en que las unidades no eficientes deben realizar mejoras en todas las variables de salida (productos) para poder alcanzar la eficiencia.

5.2 Universidad Industrial de Santander

Una de las aplicaciones del Análisis Envolvente de Datos realizadas en la Universidad Industrial de Santander, es el trabajo de grado elaborado por Barbosa (2010) titulado “Evaluación de la eficiencia de las escuelas de la Universidad Industrial de Santander (UIS) aplicando Análisis Envolvente de Datos (DEA)”. Este se dio debido al interés expresado por la universidad de realizar una inspección interna que permitiera vislumbrar el estado actual de la gestión de los recursos que se lleva a cabo en sus distintos departamentos, a partir de la adaptación del Modelo de Indicadores de Gestión que fue diseñado por el Sistema de Unidades Estatales (SUE) para evaluar instituciones

de educación superior (Ministerio de Educación Nacional, 2008). Para el desarrollo del estudio la autora decide excluir el recurso tecnológico y utilizar como entradas los recursos humanos, financieros y físicos; como asume resultados en formación, investigación y extensión. Excluyó los resultados de bienestar de las salidas ya que la asignación del recurso tecnológico y la oferta de lo que se considera en bienestar, se da a nivel institucional y no es posible categorizarlo por departamentos. Ahora bien, Barbosa decide medir la eficiencia de las escuelas de tres formas: considerando únicamente las variables de formación académica; utilizando sólo las variables de investigación y extensión; y haciendo uso simultáneo de las variables de formación académica, investigación y extensión. La orientación del modelo se enfoca a las salidas, y se aplica tanto el modelo BCC como el modelo CCR, a modo de realizar una comparación entre estos. Finalmente, y según la información obtenida, la autora señala, para cada caso, el número de escuelas calificadas como eficientes, e indica el porcentaje en que deben incrementar las variables de salida, aquellas unidades que no alcanzaron la frontera eficiente.

Por su parte López y Suárez (2011) en su proyecto “evaluación de la eficiencia de las instituciones educativas oficiales de Bucaramanga mediante el análisis envolvente de datos DEA”, plantean la necesidad de conocer si las instituciones educativas oficiales de Bucaramanga están siendo eficientes. Para llevar a cabo esta investigación, las autoras optaron por utilizar DEA, aplicando el modelo CCR y el modelo BCC orientado a las salidas, por medio del software *Efficiency Measurement System*⁵. Para este estudio, tomaron un total de 78 unidades de medida, las cuales corresponden a 39 instituciones oficiales de Bucaramanga evaluadas en los años académicos de 2009 y 2010. En cuanto a la selección de las variables a considerar dentro del

⁵ Efficiency Measurement System (EMS) es un software para Windows 9x / NT que calcula las medidas de eficiencia del Análisis Envolvente de Datos (DEA).

estudio, realizaron una revisión bibliográfica, un análisis de la información estadística y documentación del sector educativo en Colombia, consulta de expertos y finalmente un análisis estadístico de las anteriores etapas, lo anterior con el fin de seleccionar correctamente dichas variables. Con base en el análisis realizado, seleccionaron un total de seis variables de entrada (total docentes, total personal administrativo, total estudiantes matriculados, inventario, ingresos ejecutados y áreas aulas tipo A) y cinco variables de salida (estudiantes graduados, estudiantes aprobados, saber 5°, saber 9° y saber 11°).

Por medio de esta investigación, pudieron establecer no sólo cuáles planteles educativos de Bucaramanga no están siendo eficientes, sino que adicional a eso, identificaron las falencias que éstos tienen en cuanto al buen manejo de sus recursos y/o a la producción de sus productos, comparado con aquellas entidades que sí alcanzaron la eficiencia según los datos suministrados para el respectivo análisis.

También en la UIS, Yáñez, García & González (2015) llevaron a cabo una “evaluación de la eficiencia investigativa de las universidades públicas colombianas 2003-2012 utilizando el Análisis Envoltante de Datos”. Para ello clasificaron las 32 universidades evaluadas en seis grupos (Muy grande, Grande, Medio, Medio-Bajo, Pequeño y Muy pequeño) de acuerdo con tres indicadores de competencias (indicadores tradicionales adoptados por el MEN⁶; indicadores referentes a la actividad investigativa e indicadores de competencia con más consistencia en el tiempo). Aplicando el modelo CCR orientado a las salidas obtienen la medida de eficiencia de cada una de las instituciones a lo largo del periodo analizado (2003-2012), identifican el comportamiento de las mismas durante este tiempo y encuentran que las universidades Militar y Pedagógica son las únicas que permanecen en la clasificación *eficiente* durante todo el periodo

⁶ Ministerio de Educación Nacional (MEN).

estudiado. Por otro lado, encontraron que la mayoría de las instituciones pertenecientes al grupo *pequeño* resultaron ser ineficientes, en donde solo una de éstas, Sucre, obtuvo la calificación de *eficiente* en el año 2009.

Como se observa en los estudios anteriormente expuestos, la medición de la eficiencia en el ámbito educativo no es tema nuevo, por el contrario, con el paso de los años éste ha venido ganando popularidad; y aunque muchos métodos son utilizados para tal fin, en la literatura se puede evidenciar que el Análisis Envolvente de Datos es la metodología más comúnmente utilizada para este propósito. En cuanto a los trabajos que se enunciaron previamente cabe resaltar que éstos aportan significativamente a esta investigación puesto que fueron una guía al momento de definir el modelo a aplicar, las variables a tener en cuenta en el estudio y la posible aplicación de técnicas estadísticas adicionales.

6. Marco teórico

Como se mencionó en el análisis preliminar de la literatura, existen diversas metodologías para estimar el grado de eficiencia de alguna entidad, y están clasificadas en dos grandes corrientes: métodos paramétricos y no paramétricos. Los dos métodos se distinguen básicamente, porque a diferencia de los modelos no paramétricos, en los paramétricos se especifica una función para calcular la frontera de referencia –pues al tratarse de una medida relativa, la comparación entre las entidades se realiza observando la desviación existente entre cada una y una frontera construida a partir de las mismas.

Ahora bien, los modelos paramétricos pueden recurrir a técnicas estadísticas o de programación matemática, en el primer caso se encuentran con enfoque determinista los métodos de Máxima Verosimilitud (Afriat, 1972; Schmidt, 1976), y de Mínimos Cuadrados Corregidos (Richmond, 1974); en el segundo caso están los modelos de frontera de producción estocástica (Aigner, Lovell & Schmidt, 1977; Meeusen & Van den Broeck, 1977). Aigner & Chu (1968) estiman la función de producción mediante programación matemática y Tintner (1960) trata un problema de programación estocástica.

Entre los modelos no paramétricos que recurren a programación matemática, se destaca el Análisis Envolvente de Datos, propuesto por Charnes, Cooper & Rhodes (1978), el cual se da bajo un enfoque determinista que asume que todos los parámetros son numéricamente conocidos; más tarde Sengupta (1990) hará aplicable este modelo a enfoques donde las variables son aleatorias.

6.1 Análisis Envolvente de Datos

Hablar de la eficiencia como una medida relativa quiere decir que ésta sólo podría estimarse a través de una comparación con entidades similares entre sí; tal idea se refleja explícitamente en la metodología Análisis Envolvente de Datos (DEA, Data Envelopment Analysis) que utiliza la programación lineal para identificar las combinaciones de insumo-producto factibles, según la información disponible de éstas, a fin de obtener una frontera respecto a la cual se va a observar la desviación existente con las demás entidades para así obtener el grado de ineficiencia que presentan.

Para abordar el desglose de esta metodología, cuya ventaja principal es la construcción de una frontera sin la necesidad de asumir una forma funcional para ella, es importante, en primera instancia, mencionar la terminología utilizada en el DEA. En la primera publicación que dio a

conocer el modelo, Charnes, Cooper & Rhodes (1978) señalan que “el término *programa* hace referencia a una colección de unidades de decisión (DMU, Decision Making Units, por sus siglas en inglés) con entradas y salidas comunes”.

El DEA recibe tal denominación por la esencia que constituye esta metodología, “por la forma como envuelve las observaciones a fin de identificar una frontera que será utilizada para evaluar las observaciones que representan el desempeño de todas las entidades que están siendo evaluadas” (Cooper, Seirford & Tone, 2006, p.xix).

Como puede observarse en la Figura 3, la línea OF representa la frontera de producción del modelo CCR⁷, y dado que el punto C se encuentran sobre ella, éste se califica como eficiente.

De este modo, podría decirse que el concepto de envolvente surge de la definición de la función de producción: representación de la acción conjunta de los distintos factores con los que se obtienen máximos niveles productivos, ésta se determina a partir de las mejores prácticas observadas en la relación existente entre las variables de entrada y salida que sean consideradas en el modelo. Podría decirse que es el reflejo de cómo las distintas entidades procesan los recursos de que disponen para obtener los productos deseados en comparación con las demás entidades, a fin de identificar aquellas observaciones que pertenecen a la frontera de producción.

A grandes rasgos, el Análisis Envolvente de Datos consta de tres pasos para realizar la correspondiente estimación de la eficiencia. En primer lugar, y como es evidente, lo que requiere el investigador es definir las características del proceso productivo que llevan a cabo las entidades. Para ello existen supuestos cualitativos como lo son la libre disponibilidad⁸ de entradas y salidas

⁷ Ver apartado 6.3 Modelo CCR.

⁸ Para ejemplificar esta característica se va a utilizar la Figura 3. Observe el punto P, el cual indica un nivel de producción que puede obtenerse a partir de determinada cantidad de insumos, sin embargo, aún sin variar estos últimos

y los rendimientos a escala, que hacen referencia a la manera en que responde una entidad en cuanto a su producción al momento de presentarse una variación proporcional en sus insumos, es decir, al aumentar los insumos, qué comportamiento asume la producción. En segundo lugar, se debe precisar la disposición de las entidades a incrementar su beneficio, es decir, según el tipo de unidades productivas a evaluar el investigador tendrá que identificar el criterio que más las favorezca, bien sea una reducción en los recursos que mantenga el mismo nivel productivo, la maximización de las salidas del sistema con las entradas que disponen, o incluso un incremento de los productos y una reducción de los insumos simultáneamente; es en este momento donde las técnicas de programación lineal juegan su papel. Finalmente, ajustar y correr el programa computacional para obtener el índice de eficiencia especificado.

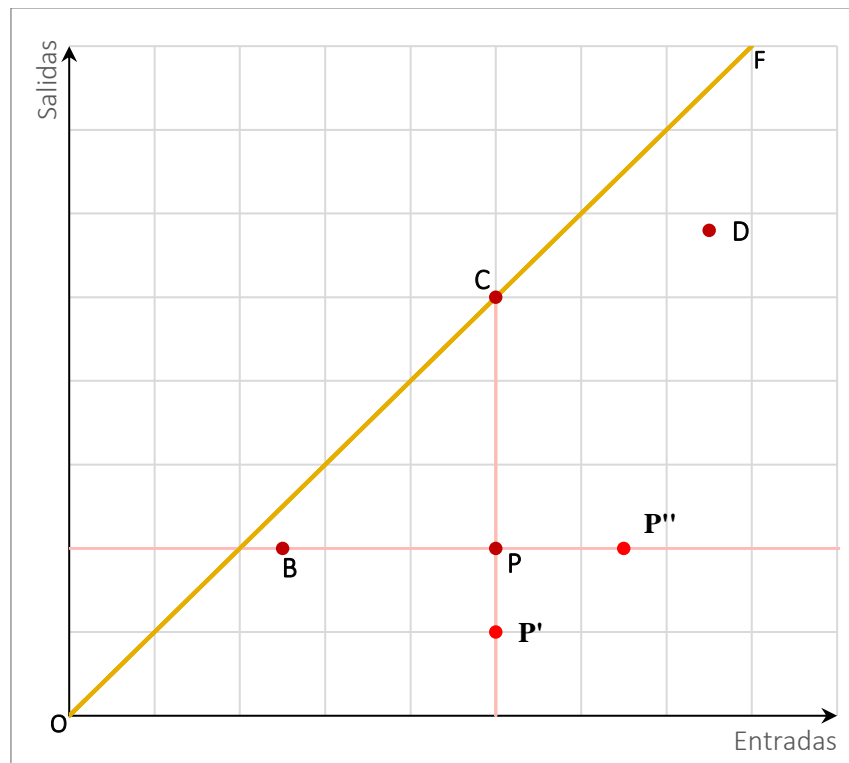


Figura 3. Frontera envolvente modelo CCR

es posible obtener una cantidad menor de productos (el punto P'); en caso contrario, es posible obtener la misma producción con un mayor consumo de recursos (el punto P'').

6.2 Escala de producción

La escala de producción se deriva de lo que en la teoría económica se conoce como economía de escala –definida en términos de costos–. Cuando se habla de la escala de un proceso productivo se está haciendo referencia a cómo se ven afectados los niveles de producción cuando se modifican en determinada proporción las entradas al proceso. Suponga que para obtener una unidad del producto P se requieren de dos unidades del recurso A y una del recurso B, entonces si se desea obtener el doble o el triple del producto, se requiere el doble o el triple de los insumos. Lo anterior es lo que se conoce como retornos constantes a escala, que indican que las salidas del proceso productivo aumentan o disminuyen en la misma proporción en que lo hacen las entradas.

Ahora bien, los retornos variables a escala indican que al aumentar (o disminuir) en determinada proporción los recursos, los productos podrán aumentar (o disminuir) en una proporción menor o mayor.

Matemáticamente los rendimientos a escala podrían expresarse de la siguiente manera:

- Rendimientos constantes a escala

$$F(tx_{ij}, ty_{rj}) = tF(x_{ij}, y_{rj}), \forall t > 0$$

- Rendimientos variables a escala

$$F(tx_{ij}, ty_{rj}) > tF(x_{ij}, y_{rj}) \text{ Crecientes, } \forall t > 1$$

$$F(tx_{ij}, ty_{rj}) < tF(x_{ij}, y_{rj}) \text{ Decrecientes, } \forall t < 1$$

Donde:

x_{ij} cantidad de recurso i utilizado por la entidad j

y_{rj} cantidad de producto r obtenido por la entidad j

$F(x_{ij}, y_{rj})$ representa la función de producción

t cualquier escalar positivo

6.3 Modelo CCR

El modelo CCR es el primer modelo DEA propuesto, considera rendimientos constantes a escala; es conocido como CCR por la inicial de los autores que lo desarrollaron en el año 1978, Charnes, Cooper & Rhodes. Los autores reiteran la utilización de este modelo en entidades sin ánimo de lucro en una posterior publicación realizada por ellos mismos en 1981, pues a pesar de que sus supuestos se basan en los propuestos por la teoría económica, tienden a desviarse de ella o a no considerarlos en su totalidad (como la *maximización de ganancias*), razón por la cual ha sido mayormente aplicado en la medición de eficiencia de entidades educativas como Colegios y Universidades.

6.3.1 Conjunto de posibilidades de producción. Se define el conjunto de posibilidades de producción como “el conjunto de procesos productivos tecnológicamente factibles” (Álvarez, 2013 p.14). Esto quiere decir que estará conformado por todas aquellas combinaciones de entradas y salidas que son viables; como puede observarse en la Figura 3 tal conjunto estaría conformado por todos los puntos sobre la línea OF y todos aquellos que se encuentren por debajo de ésta.

Considere un proceso productivo que cuenta con $i = 1, \dots, m$ entradas al sistema (recursos o insumos) y $r = 1, \dots, s$ salidas del proceso (productos), y se tienen $j = 1, \dots, n$ entidades (DMU's) para comparar entre sí, de modo que cada una de ellas hará uso de cada entrada en la cantidad x_{ij} para producir distintas salidas en la cantidad y_{rj} .

Ahora bien, los rendimientos constantes que asume el modelo CCR indican que la tecnología de producción de las entidades a evaluar es rígida, lo que en otras palabras sería que el producto sólo puede obtenerse de una sola combinación de recursos; este supuesto “permite que toda la información relevante sea presentada en un diagrama simple isocuante” (Farrel, 1957, p.254).

De este modo, el conjunto de producción P para el modelo CCR se define de la siguiente manera:

$$P = \{(\mathbf{x}, \mathbf{y}) \mid \mathbf{x} \geq X\lambda, \mathbf{y} \leq Y\lambda, \lambda \geq 0\} \quad (1,0)$$

$$X = \begin{bmatrix} x_{11} & \dots & x_{1n} \\ \vdots & \ddots & \vdots \\ x_{m1} & \dots & x_{mn} \end{bmatrix} \quad Y = \begin{bmatrix} y_{11} & \dots & y_{1n} \\ \vdots & \ddots & \vdots \\ y_{s1} & \dots & y_{sn} \end{bmatrix} \quad \lambda = [\lambda_1 \quad \lambda_2 \quad \dots \quad \lambda_n]^T = \begin{bmatrix} \lambda_1 \\ \lambda_2 \\ \vdots \\ \lambda_n \end{bmatrix}$$

Donde, (\mathbf{x}, \mathbf{y}) ⁹ representa una actividad.

Así P , conocido como Conjunto de Posibilidades de Producción, se puede definir como el conjunto de actividades factibles, dicha actividad conformada por $\mathbf{x} \in R^m$ (vector semipositivo de entradas), y $\mathbf{y} \in R^s$ (vector semipositivo de salidas).

Las propiedades de este conjunto específicamente para un modelo DEA, y que se deducen en (1,0), se expresan en Cooper *et al.* (2006) de la siguiente manera:

- Las actividades observadas $(\mathbf{x}_j, \mathbf{y}_j)$ ($j = 1, 2, \dots, n$) pertenecen a P .
- Si una actividad (\mathbf{x}, \mathbf{y}) pertenece a P , entonces la actividad $(t\mathbf{x}, t\mathbf{y})$ también pertenece a P , para cualquier escalar positivo t .
- Para cualquier actividad (\mathbf{x}, \mathbf{y}) en el conjunto P , cualquier actividad semipositiva $(\bar{\mathbf{x}}, \bar{\mathbf{y}})$ con $\bar{\mathbf{x}} \geq \mathbf{x}$ y $\bar{\mathbf{y}} \leq \mathbf{y}$ también pertenece al conjunto P .
- Cualquier combinación lineal semipositiva de actividades en P , pertenece a P .

6.3.2 Modelo Fraccional. Charnes *et al.* (1978) en su publicación *Measuring the efficiency of decision making units*, expresan que el objetivo de la medición de la eficiencia es obtener la maximización de la razón entre las salidas ponderadas y las entradas ponderadas del proceso productivo, y formulan el modelo de la siguiente la manera:

⁹ Los caracteres que se encuentren en negrita hacen referencia a notación matricial.

$$\max_{u,v} h_o = \frac{\sum_{r=1}^s u_r y_{ro}}{\sum_{i=1}^m v_i x_{io}} \quad (1,1)$$

$$\text{sujeto a } \frac{\sum_{r=1}^s u_r y_{rj}}{\sum_{i=1}^m v_i x_{ij}} \leq 1, \quad j = 1, 2, \dots, n \quad (1,2)$$

$$u_1, \dots, u_s \geq 0 \quad (1,3)$$

$$v_1, \dots, v_m \geq 0 \quad (1,4)$$

Donde:

x_{ij} cantidad de entrada i consumida por la DMU j

x_{io} cantidad de entrada i consumida por la entidad que está siendo evaluada o (observada)

y_{rj} cantidad de salida r producida por la DMU j

y_{ro} cantidad de salida r producida por la entidad que está siendo evaluada o (observada)

u_r peso asignado a la salida r

v_i peso asignado a la entrada i

En la literatura, cuando se hace referencia a la suma ponderada de las salidas $\sum_{r=1}^s u_r y_{ro}$, y la suma ponderada de las entradas $\sum_{i=1}^m v_i x_{ij}$, éstas reciben el nombre de *salidas virtuales* y *entradas virtuales* respectivamente.

El modelo anterior busca obtener la ponderación (peso) óptima de cada uno de los factores (tanto de entrada como de salida) de la entidad que está siendo observada, de manera que estos pesos la favorezcan en su puntuación de eficiencia.

La utilización de ponderaciones provee adimensionalidad, lo que permite solventar la disparidad existente entre las unidades de medida de los distintos factores al momento de considerarlos en una única expresión. Así pues, las variables del modelo resultan ser los pesos asociados a cada factor (v_i, u_r), mientras que la actividad (x_{ij}, y_{rj}) son datos que se emplean para dar solución al modelo.

Los pesos óptimos de la entidad observada se obtienen cuando se cumple que, al evaluar la eficiencia de todas las entidades con estos mismos, el resultado es menor que la unidad, como lo

refleja la restricción (1,2). De aquí que la puntuación de la eficiencia tome valores entre cero y uno.

6.3.3 Modelo Lineal. Con el propósito de facilitar la resolución del problema, Charnes, Cooper y Rhodes, realizan la correspondiente transformación para hallar el equivalente lineal del programa fraccional mostrado anteriormente.

A continuación, se presenta la formulación del modelo en la forma lineal (1,5)–(1,9), también conocida como forma multiplicativa, en la cual se hace evidente un cambio en la notación de los pesos asignados a cada factor, esto no es más que para hacer énfasis de que se trata de un problema diferente al que se ha mostrado en la forma fraccional.

En este caso, el problema busca maximizar el numerador del modelo fraccional incluyendo la restricción de normalización (1,6), obteniéndose así un resultado de eficiencia igual en ambas formulaciones (fraccional y lineal) pero con diferentes valores en las variables, razón por la cual la notación cambia de uno a otro.

$$\max_{\mu, \nu} \sum_{r=1}^s \mu_r y_{ro} \quad (1,5)$$

$$\text{sujeto a} \quad \sum_{i=1}^m \nu_i x_{io} = 1 \quad (1,6)$$

$$\sum_{r=1}^s \mu_r y_{rj} \leq \sum_{i=1}^m \nu_i x_{ij} \quad j = 1, 2, \dots, n \quad (1,7)$$

$$\mu_1, \dots, \mu_s \geq 0 \quad (1,8)$$

$$\nu_1, \dots, \nu_m \geq 0 \quad (1,9)$$

Nótese que aquí la puntuación de eficiencia corresponde al valor de la salida virtual, de la cual puede obtenerse la contribución hecha por cada producto a ese valor ($\mu_r y_{ro}$). La importancia de cada factor productivo está dada por $\nu_i x_{io}$. Como lo manifiestan Boussofiane, Dyson y Thanassoulis (1991) “los valores de los Inputs y Outputs virtuales expresan información sobre la

importancia que una unidad atribuye a determinados Inputs y Outputs con la intención de obtener su máxima puntuación de eficiencia” (p.6).

6.3.4 Modelo Dual. Charnes *et al.* (1978) afirman que utilizar un modelo dual “provee una nueva forma de estimar relaciones extremas a partir de los datos observados” (p.429) y adicionalmente permite la identificación de los grupos eficientes que sirven de referencia para aquellos ineficientes, a partir de la variable λ_j .

El programa Dual asociado al modelo lineal está formulado de la siguiente manera:

$$\min_{\theta, \lambda} \theta \quad (2,0)$$

$$\text{sujeto a } \theta x_{io} \geq \sum_{j=1}^n \lambda_j x_{ij} \quad i = 1, 2, \dots, m \quad (2,1)$$

$$\sum_{j=1}^n \lambda_j y_{rj} \geq y_{ro} \quad r = 1, 2, \dots, s \quad (2,2)$$

$$\lambda_1, \dots, \lambda_n \geq 0 \quad (2,3)$$

Donde:

$$\theta \text{ libre}$$

θ es el grado de eficiencia asociado a cada entidad

λ es un vector no negativo de variables asociado a cada entidad.

En este caso, la puntuación de eficiencia está dada por θ y corresponde al mismo valor obtenido en el modelo primal –nombrado anteriormente como el programa lineal o multiplicativo–. Tal como lo indican Coll y Blasco (2006) “para todo programa lineal original (programa primal) existe otro programa lineal asociado, denominado programa dual, que puede ser utilizado para determinar la solución del problema primal”

Es importante señalar que ambos modelos –primario y dual– se resuelve n veces, es decir, existe un programa para cada una de las entidades que se van a evaluar. Por otro lado, el número de restricciones asociadas al modelo primal corresponde a $n + 1$, mientras que para el dual esta cantidad depende de los factores, es decir, serán $m + s$ restricciones –sin contar aquellas que expresan la positividad de las variables–.

Del mismo modo, hay que apuntar que del programa primal se va a obtener el valor de $m + s$ variables, que corresponden a los pesos (v_i, μ_r) asignados a cada factor; mientras que en el programa dual la cantidad de variables resulta ser $n + 1$ correspondientes a λ_j y θ .

Adviértase que el programa dual y el programa primal presentan funciones objetivo opuestas, en el sentido de que mientras el primal maximiza, el dual minimiza, siendo este último el que refleje la orientación del modelo especificado por el investigador. Es decir, si en el estudio se busca la máxima reducción de los insumos manteniendo el mismo nivel de producción –enfoque a las entradas– deberá formularse el modelo primal (1,5)– (1,9) o el modelo dual (2,0)– (2,3); éste último “busca una actividad P que garantice al menos un nivel de producción y_{ro} de la entidad observada en todos los componentes, mientras se reduce el vector x_{io} proporcionalmente al más pequeño posible” (Cooper *et al.*, 2006, p.44).

6.3.5 Modelo orientado a las salidas. En caso de que el objetivo de la investigación –dada la naturaleza de las entidades a evaluar– sea buscar una maximización de las salidas, sin alterar el nivel de entradas, Charnes *et al.* (1978) proponen el siguiente modelo para este enfoque:

$$\min_{\mu, v} \sum_{i=1}^m v_i x_{io} \quad (2,4)$$

$$\text{sujeto a } \sum_{r=1}^s \mu_r y_{ro} = 1 \quad (2,5)$$

$$\sum_{r=1}^s \mu_r y_{rj} \leq \sum_{i=1}^m v_i x_{ij} \quad j = 1, 2, \dots, n \quad (2,6)$$

$$\mu_1, \dots, \mu_s \geq 0 \quad (2,7)$$

$$v_1, \dots, v_m \geq 0 \quad (2,8)$$

El programa lineal (primal) expresado por las ecuaciones (2,4)– (2,8) busca la minimización de la suma ponderada de las entradas. Las restricciones (2,5) y (2,6) hacen que la solución óptima de

este modelo tome valores por encima de la unidad, sin embargo, la puntuación eficiencia corresponde al recíproco de este valor.

En cambio, el programa dual o envolvente asociado (2,9)– (3,2) que se presenta a continuación, tiene como objetivo encontrar la máxima producción que una entidad j puede alcanzar con el nivel de entradas que utiliza. Así que cuando la entidad está produciendo lo más que puede con lo que tiene, η toma un valor igual a uno; mientras que una entidad que esté en capacidad de producir más con sus insumos actuales, obtendrá valores de $\eta > 1$.

$$\max_{\eta, \lambda} \eta \quad (2,9)$$

$$\text{sujeto a } x_{io} \geq \sum_{j=1}^n \lambda_j x_{ij} \quad i = 1, 2, \dots, m \quad (3,0)$$

$$\sum_{j=1}^n \lambda_j y_{rj} \geq \eta y_{ro} \quad r = 1, 2, \dots, s \quad (3,1)$$

$$\lambda_1, \dots, \lambda_n \geq 0 \quad (3,2)$$

$$\eta \text{ libre}$$

En este enfoque también se cumple que la solución óptima del modelo primal es igual a la obtenida en el programa dual.

6.3.6 Solución óptima CCR. Resolver ya sea un programa primal o un modelo dual, lleva a la obtención de valores que representan el grado de eficiencia de las entidades evaluadas. Cuando el modelo está enfocado a las entradas (CCR-I), la puntuación de eficiencia corresponde al valor óptimo de θ , mientras que, en un enfoque a las salidas (CCR-O), ese valor está dado por el recíproco del valor óptimo de η .

Las soluciones óptimas serán expresadas por θ^* y λ^* en el modelo CCR-I, y por η^* y λ^* en el modelo CCR-O.

No obstante, como lo mencionan Coelli *et al.* (2005), dar solución únicamente en una etapa puede ocasionar la obiedad de valores de holgura, por lo que Ali & Seiford (1993) proponen

realizar una segunda etapa a fin de encontrar inicialmente lo que se conoce como Eficiencia de Farrell¹⁰ y a partir de ella hallar lo que se conoce como Eficiencia de Pareto-Koopmans¹¹.

En ese orden de ideas, puede hablarse de la metodología de dos fases: que resuelve en primera instancia el programa dual en busca de un valor óptimo η^* –en el caso de un modelo orientado a las salidas–, para posteriormente (en la segunda fase) utilizarlo para “ajustar” los valores de entrada: $\eta^* x_{io}$ y calcular la máxima suma total de valores de holgura¹².

La segunda fase para un modelo orientado a las salidas se expresa de la siguiente manera:

$$\max_{\lambda, s^-, s^+} \omega = \sum_{i=1}^m S_i^- + \sum_{r=1}^s S_r^+ \quad (3,3)$$

$$\text{sujeto a} \quad S_i^- = x_{io} - \sum_{j=1}^n \lambda_j x_{ij} \quad (3,4)$$

$$S_r^+ = \sum_{j=1}^n \lambda_j y_{rj} - \eta^* y_{ro} \quad (3,5)$$

$$\lambda_j \geq 0, \quad S_i^- \geq 0, \quad S_r^+ \geq 0 \quad (3,6)$$

Donde:

S_i^- variable de holgura que representa un exceso en la entrada i .

S_r^+ variable de holgura que representa un déficit en el producto r .

Entonces, una entidad será calificada como CCR-O eficiente si y sólo si obtiene en la primera fase un valor $\eta^* = 1$ (Eficiencia de Farrell) y obtiene en la segunda fase todas las holguras (Eficiencia de Pareto-Koopmans) $S_i^- = S_r^+ = 0$.

Ahora bien, dichos valores de holgura resultan ser de gran importancia en el análisis de la eficiencia, pues “estos vectores miden la distancia entre una entidad catalogada ineficiente (x_{io}, y_{ro}) y su punto proyectado $(\hat{x}_{io}, \hat{y}_{ro})$ sobre la frontera” (Ali & Seiford, 1993, p.126).

¹⁰ La Eficiencia de Farrell hace referencia a la obtención de $\eta^* = 1$ para la entidad observada.

¹¹ La Eficiencia de Pareto-Koopmans indica que para calificar como eficiente a alguna entidad, ésta debe obtener no sólo un valor de $\eta^* = 1$ sino que además no debe tener ninguna holgura.

¹² Las holguras representan ya sea una escasez en las salidas o un exceso de las entradas. Este valor indica la cantidad en que se deben aumentar o disminuir los productos y/o recursos respectivamente, de manera que se siga obteniendo el mismo valor de eficiencia para el caso de las entidades con $\eta^* = 1$ y que presenten holguras, y para aquellas con $\eta^* \neq 1$ las cantidades que necesitan para ser catalogadas como eficientes.

$$\hat{x}_o = X\lambda = \theta^* x_o - s^- \quad (3,7)$$

$$\hat{y}_o = Y\lambda = y_o + s^+ \quad (3,8)$$

Por supuesto, para determinar las mejoras potenciales que una entidad calificada como ineficiente debe adoptar para alcanzar la eficiencia, es necesario que se indique cuáles son sus *pares*. Así pues, la técnica DEA genera grupos de entidades que han sido calificadas como eficientes y que presentan un proceso productivo similar al de la entidad calificada como ineficiente, estos grupos reciben el nombre de conjunto de referencia y se formula como se muestra en (3,9).

$$E_o = \{j | \lambda_j^* > 0\} \quad (j \in \{1, \dots, n\}) \quad (3,9)$$

Es decir, “es la existencia de estas unidades eficientes son las que fuerzan a la unidad observada a ser ineficiente” (Cooper, Seiford, Tone, 2006).

En conclusión, cuando una entidad obtenga una puntuación de eficiencia $\theta^* = 1$ se dice que es *técnicamente eficiente* o que presenta una eficiencia *débil* debido a que la presencia de alguna holgura indica que la entidad podría ser aún mejor en términos de su proceso productivo. En caso de que el valor de la eficiencia sea $\theta^* < 1$, la entidad será calificada como ineficiente y deberá reducir radialmente la utilización de sus recursos en $(1 - \theta^*)$ en caso de estar orientado a las entradas, o aumentar radialmente el nivel de producción en $(1 - \eta^*)$ en caso de que el modelo se oriente a las salidas; además de esto, la existencia de holguras indicará que la entidad debe realizar alguna reducción adicional de sus entradas, o algún incremento adicional de sus salidas, y en consecuencia se verán alteradas las proporciones (la mezcla) que caracterizan los rendimientos constantes. Es por ello por lo que “las ineficiencias asociadas con cualquier holgura no nula identificada en el procedimiento de dos fases se denominan *ineficiencias de mezcla*” (Cooper et

al., 2006, p.45). Por lo tanto, si una entidad además de obtener una puntuación de eficiencia $\theta^* = 1$ no presenta valores de holgura, se dirá que es *fuertemente* eficiente.

6.4 Modelo BCC

Seis años después de haberse propuesto el modelo CCR de rendimientos constantes, surge en la literatura lo que se conocería como el modelo BCC por la sigla de los autores R. D. Banker, A. Charnes, y W. W. Cooper, quienes en su publicación de 1984 proponen una técnica para estimar las ineficiencias técnicas y de escala para una tecnología¹³ de rendimientos variables.

La diferencia entre un modelo y otro puede observarse de manera simple en relación con la productividad, mientras que en el modelo CCR los rendimientos constantes conceden que un incremento en la producción –mayor eficiencia– no afecte la productividad de la entidad, en el modelo BCC no sucede, pues en caso de rendimientos decrecientes, dicho incremento en los niveles productivos está asociado a una menor productividad. Otro aspecto diferenciador entre los modelos resulta ser que, a diferencia del programa BCC, el modelo CCR es más restrictivo, por lo cual va a generar una menor cantidad de entidades catalogadas como eficientes. Así mismo, la puntuación de eficiencia asignada a las entidades mediante el modelo BCC será mayor que la establecida por CCR.

Lo anterior es fácilmente observable a través de la Figura 4, donde al identificar la frontera eficiente CCR por la línea OF es posible notar que tan sólo la entidad C ha sido catalogada como eficiente, mientras que sobre la línea azul que representa la frontera para el modelo BCC se hallan tres entidades (B, C y D), exhibiendo así mayor flexibilidad. Ahora bien, si se observa el punto A

¹³ “La tecnología de producción puede representarse como esa «fórmula» que determina la manera en que los inputs pueden transformarse en distintos vectores de output” (Gonzales, 2013, p.142).

que se encuentra por debajo de las fronteras, su puntuación de eficiencia vendría dada por el cociente entre la distancia desde el eje x hasta el punto proyectado sobre la frontera, y la distancia desde el eje x al punto A.

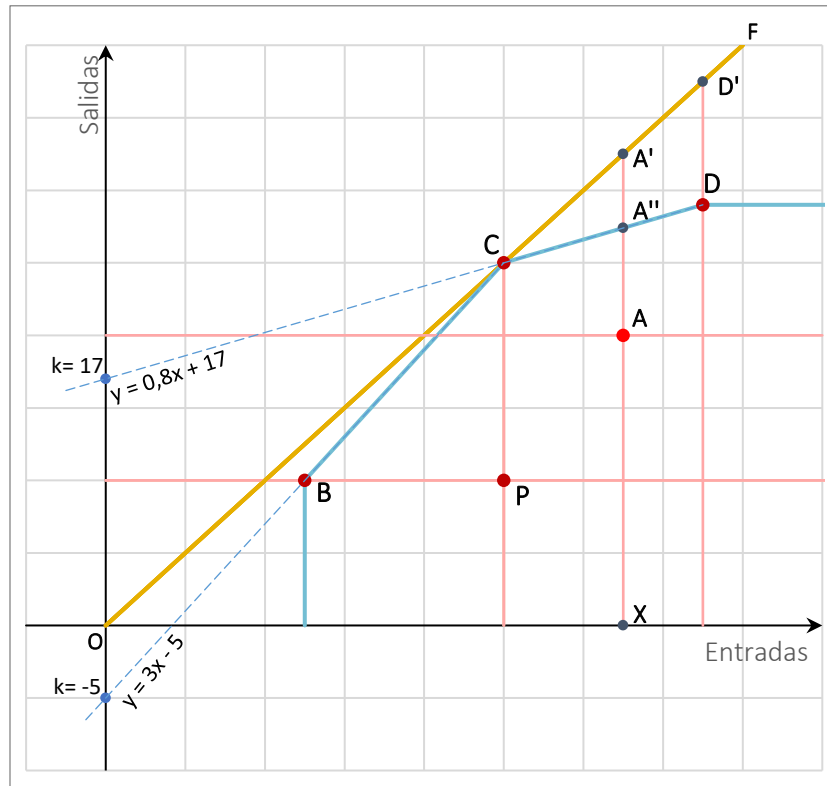


Figura 4. Frontera envolvente CCR y BCC.

Para el modelo CCR dicho cociente estaría representado por $\frac{XA}{XA'}$ en caso de una orientación a las salidas, mientras que para el modelo BCC sería $\frac{XA}{XA''}$. Aquí se hace evidente que el valor de eficiencia asociado a este último será siempre mayor que el valor obtenido en el modelo CCR, pues la distancia XA' es mayor que la distancia XA'' .

6.4.1 Conjunto de posibilidades de producción. En este modelo, al igual que se expuso para el programa CCR, el conjunto de producción está conformado por todas aquellas combinaciones de insumo–producto que son viables, pero con una restricción adicional, que busca relajar la

condición de rendimientos constantes, expresada por $\sum_{j=1}^n \lambda_j = 1$ y que es conocida como *restricción de convexidad*.

$$P_B = \{(\mathbf{x}, \mathbf{y}) \mid \mathbf{x} \geq \mathbf{X}\boldsymbol{\lambda}, \mathbf{y} \leq \mathbf{Y}\boldsymbol{\lambda}, \mathbf{e}\boldsymbol{\lambda} = 1, \boldsymbol{\lambda} \geq 0\} \quad (4,0)$$

Donde:

\mathbf{e} es un vector fila con todos los valores igual a 1.

$\boldsymbol{\lambda}$ es un vector columna cuyos elementos son no-negativos pertenecientes a R^n

$$\mathbf{X} = (\mathbf{x}_j) \in R^{m \times n}$$

$$\mathbf{Y} = (\mathbf{y}_j) \in R^{s \times n}$$

La restricción de convexidad, que puede ser expresada en términos vectoriales como $\mathbf{e}\boldsymbol{\lambda} = 1$, hace que las entidades calificadas como ineficientes sean comparadas únicamente con aquellas que manejen la misma escala de producción.

6.4.2 Modelo lineal y envolvente. En la Tabla 2 se muestran los modelos BCC dual y primal tanto para una orientación a las entradas como para las salidas. Si se compara con el modelo CCR primal la diferencia en BCC radica en la aparición de la variable k_o que representa el intercepto de los segmentos de la frontera BCC (observe las líneas punteadas en la Figura 4); el signo correspondiente al escalar k_o indicará el tipo de rendimientos que presenta la entidad que está siendo observada. Así pues, si éste toma valores negativos quiere decir que predominan rendimientos crecientes en una orientación a las entradas y decrecientes en una orientación a las salidas; por otro lado, si k_o toma valores positivos quiere decir que en la tecnología prevalecen rendimientos decrecientes en caso de orientar el modelo a las entradas y crecientes en caso de estar orientado a las salidas. Finalmente, si el valor que toma k_o es cero, esto indica que los rendimientos que predominan son constantes.

Ahora, al comparar el modelo dual del CCR con BCC se evidencia la aparición de la restricción de convexidad para limitar la forma en que se pueden combinar las actividades para cada una de las entidades.

Tabla 2
Modelo BCC orientado a las entradas y orientado a las salidas.

Modelo BCC		
	Primal	Dual
Orientado a las entradas	$\max_{v,u,k_o} z = \mathbf{u}y_o - k_o$ $\mathbf{v}x_o = 1$ $\mathbf{u}Y - \mathbf{v}X - \mathbf{e}k_o \leq 0$ $\mathbf{u} \geq 0, \quad \mathbf{v} \geq 0, \quad k_o \text{ libre en signo}$	$\min_{\theta_B, \lambda} \theta_B$ $X\lambda \leq \theta_B x_o$ $y_o \leq Y\lambda$ $\mathbf{e}\lambda = 1$ $\lambda \geq 0$
	Primal	Dual
Orientado a las salidas	$\min_{v,u,k_o} z = \mathbf{v}x_o - k_o$ $\mathbf{u}y_o = 1$ $-\mathbf{u}Y + \mathbf{v}X - \mathbf{e}k_o \geq 0$ $\mathbf{u} \geq 0, \quad \mathbf{v} \geq 0, \quad k_o \text{ libre en signo}$	$\max_{\eta_B, \lambda} \eta_B$ $X\lambda \leq x_o$ $\eta_B y_o \leq Y\lambda$ $\mathbf{e}\lambda = 1$ $\lambda \geq 0$

Nota: Adaptado de Cooper, W., Seiford, L., & Tone, K. (2006). *Introduction to Data Envelopment Analysis and its uses*. Springer Science+Business Media, Inc.

6.4.3 Solución óptima BCC. Al igual que para el modelo CCR, el programa BCC es resuelto en dos etapas a fin de garantizar una adecuada identificación de las entidades eficientes.

Como se presentó en el apartado del modelo CCR, la primera etapa corresponde a la solución del modelo dual o primal, según se prefiera, pues como ya se ha mencionado, tanto el modelo primal como el dual arrojan la misma solución óptima, que en el caso de una orientación a las salidas corresponde a la tasa de aplicación de los productos η_B y en el caso de una orientación a las entradas corresponde al nivel de eficiencia de la entidad θ_B . Posteriormente, la solución óptima de la primera fase será utilizada en la segunda etapa para buscar la maximización de las holguras en insumos y productos.

En la Tabla 3 se muestra en un solo modelo las dos etapas para cada orientación.

Tabla 3
Modelo BCC de dos etapas

Orientación a las entradas	Orientación a las salidas
$\min_{\theta_B, \lambda, s^-, s^+} \theta_B - es^- - es^+$ $s^- = \theta_B x_o - X\lambda$ $s^+ = Y\lambda - y_o$ $e\lambda = 1$ $\lambda \geq 0 \quad s^+ \geq 0 \quad s^- \geq 0$	$\max_{\eta_B, \lambda, s^-, s^+} \eta_B + es^- + es^+$ $s^- = x_o - X\lambda$ $s^+ = Y\lambda - \eta_B y_o$ $e\lambda = 1$ $\lambda \geq 0 \quad s^+ \geq 0 \quad s^- \geq 0$

Nota: Adaptado de Cooper, W., Seiford, L., & Tone, K. (2006). *Introduction to Data Envelopment Analysis and its uses*. Springer Science+Business Media, Inc.

Es importante señalar que –sin importar la orientación del modelo– la solución dada por el programa CCR se trata de una eficiencia técnica global (ETG) ya que en éste se tiene en cuenta tanto la eficiencia técnica (ETP) como la eficiencia de escala (EE), mientras que el resultado arrojado por el programa BCC está libre de cualquier efecto de escala por lo cual se dice que es una medida de eficiencia técnica pura (ETP). En el caso de la Figura 4, las eficiencias mencionadas anteriormente se pueden obtener de la siguiente manera para el punto A:

$$ETG = ETP * EE = \frac{XA}{XA''} * \frac{XA''}{XA'} = \frac{XA}{XA'}$$

Luego de resolver las dos etapas, una entidad será calificada como BCC–eficiente si y sólo si cumple con las condiciones: $\theta_B^* = 1$ y $s^- = s^+ = 0$ ó $\eta_B^* = 1$ y $s^- = s^+ = 0$, para un modelo orientado a las entradas o a las salidas respectivamente, esto según la definición de eficiencia de Pareto–Koopmans.

Ahora bien, si una entidad es calificada como CCR–eficiente y BCC–eficiente, es decir, con $\theta^* = \theta_B^* = 1$ y $s^- = s^+ = 0$, quiere decir que dicha entidad no posee ningún tipo de ineficiencia y por tanto genera la máxima productividad. En caso contrario, se dice que posee ineficiencia de escala. Para ejemplificar lo expuesto anteriormente observe el punto C en la Figura 4, el cual se

encuentra sobre las dos fronteras cumpliendo la primera característica formulada, por su parte el punto D, aunque es BCC–eficiente, presenta ineficiencia de escala equivalente a la distancia entre una frontera y otra, es decir, la distancia $\overline{D'D}$.

Para una entidad calificada como BCC–ineficiente existe un conjunto de referencia el cual estará conformado por todas aquellas entidades calificadas como eficientes y que produzcan bajo los mismos rendimientos. Lo anterior es posible gracias a la restricción de convexidad ya que ésta permite que la entidad evaluada sea comparada únicamente con otras que presenten escalas de producción similares. Dicho conjunto de referencia se expresa matemáticamente así:

$$E_o = \{j | \lambda_j^* > 0\} \quad (j \in \{1, \dots, n\}) \quad (4,1)$$

En definitiva, hay que apuntar que en la literatura es común encontrar que el modelo BCC es una extensión del CCR, así mismo, al permitir identificar la tipología de los rendimientos (entre ellos una escala constante) algunos autores afirman que podría considerarse al modelo CCR un caso particular del modelo BCC (Coll y Blasco, 2006). Adicionalmente, cabe señalar que a diferencia de un modelo que supone rendimientos constantes (CCR), en donde se cumple que el cálculo de la eficiencia bajo una orientación a las salidas es el recíproco de la misma orientada a las entradas, en el modelo BCC esto no se cumple.

6.5 Otros modelos DEA

Como se ha expuesto en apartados anteriores, la metodología DEA permite clasificar las entidades observadas como eficientes o ineficientes según la puntuación arrojada por el programa matemático utilizado –CCR y/o BCC– y la orientación del mismo. Sin embargo, la clasificación generada a partir de dicha metodología no ordena las DMU de acuerdo con el nivel de eficiencia

obtenido por cada una de éstas, ya que todas aquellas puntuadas como eficientes reciben un valor igual a la unidad aun cuando el nivel de productividad de éstas no es necesariamente equivalente. Por lo anterior, surgen modelos como extensión del DEA, a través de los cuales se pueden llevar a cabo una ordenación total de las entidades. Tal es el caso de métodos de Súper Eficiencia, Eficiencia Cruzada y líder global.

6.5.1 Súper Eficiencia. En 1993 Andersen & Petersen proponen un modelo que permitiera identificar cuáles entidades, dentro de las que han sido clasificadas como eficientes en los modelos principales del DEA, son aún mejores de manera que se pueda dar un orden fácilmente observable desde las más destacadas.

El modelo consiste, básicamente, en comparar la entidad observada con las demás DMU, sin tenerse en cuenta a ella misma, esto se observa en las restricciones del programa dual en las cuales se excluye la j observada de la sumatoria, por lo cual los puntajes de eficiencia obtenidos podrán ser mayores a la unidad en caso de un modelo orientado a las entradas; mientras que en un modelo orientado a las salidas (que se muestra en la Tabla 4), el valor de η obtenido luego de correr el modelo de Súper Eficiencia, se encontrará por debajo de la unidad, y al realizar el cálculo de la eficiencia técnica pura ($1/\eta$) usando ese valor, se obtendrán valores superiores al 100% –de ahí el nombre de *Súper Eficiencia*.

6.5.2 Eficiencia cruzada. En 1986 Sexton, Silkman & Hogan desarrollan este modelo con el fin de ordenar las entidades según su puntuación de eficiencia, y para el año 1994 Doyle & Green realizan una mejora al mismo. Dicho modelo consiste en puntuar cada entidad con base en los pesos óptimos arrojados por el programa primal del modelo DEA–CCR, es decir, se obtienen n valores de eficiencia para cada DMU con base en los multiplicadores óptimos u y v de la DMU

observada; posteriormente se calcula un promedio de tal valor para cada entidad y este recibirá el nombre de *Eficiencia Cruzada* (CE).

Los resultados de Eficiencia Cruzada, además de permitir listar en orden de eficiencia a las entidades, pondrán de manifiesto si las actividades productivas entre unas y otras son similares, o bien si no lo son.

Tabla 4
Programa matemático de Eficiencia Cruzada y Súper Eficiencia

Eficiencia Cruzada	Súper eficiencia
$E_{oj} = \frac{\sum_{r=1}^s u_{ro} y_{rj}}{\sum_{i=1}^m v_{io} x_{ij}}$	$\max_{\eta, \lambda} \eta$
$CE_j = \frac{\sum_{o \neq j}^n E_{oj}}{n - 1}$	$\sum_{j \neq o}^n \lambda_j x_{ij} \leq x_{io} \quad \forall i$
$j = 1, 2, \dots, n$	$\sum_{j \neq o}^n \lambda_j y_{rj} \geq \eta y_{ro} \quad \forall r$
$o \text{ indica la DMU observada}$	$\lambda_1, \dots, \lambda_n \geq 0$
	$\eta \text{ libre}$

Nota: Adaptado de Cooper, W., Seiford, L., & Tone, K. (2006). *Introduction to Data Envelopment Analysis and its uses*. Springer Science+Business Media, Inc.

6.5.3 Líder Global. Este concepto, introducido por Oral & Yolalan (1990), hace referencia a la aparición reiterada de una entidad eficiente en los grupos de referencia de aquellas entidades calificadas como ineficientes. Esto estaría indicando que la entidad eficiente presenta un mejor desempeño que las demás, lo que se traduce en una eficiencia más sólida.

En el caso de que la DMU sea calificada como eficiente y no se encuentre en ningún grupo de referencia, apunta a que ésta posee una actividad poco común (Coll & Blasco, 2006).

7. Metodología

La aplicación del Análisis Envolvente de Datos se realiza en tres fases, 1) la especificación, donde se define y justifica la elección de las entidades a evaluar, la orientación del modelo y el conjunto de variables que lo van a conformar; 2) la recolección de la información, en la que será posible identificar los insumos y productos que se van a utilizar en la corrida del modelo, según la información disponible acerca de éstos, para luego realizar el ajuste del modelo y establecer cuáles unidades productivas son las más eficientes y cuáles requieren mejorar sus prácticas para alcanzar el máximo rendimiento; por último 3) el análisis de resultados¹⁴, donde se mencionan los grupos eficientes y los que no lo son, y se indican las mejoras potenciales que deberán adoptar los grupos de investigación de las facultades de ingeniería de la Universidad Industrial de Santander, para alcanzar la eficiencia.

7.1 Especificación del modelo

7.1.1 Selección de unidades de decisión (UD). Asignar una calificación de eficiencia requiere de una comparación entre las entidades objeto de estudio, como se ha expuesto en apartados anteriores, debido a que este tipo de característica solo es observable respecto a un punto de referencia. Entonces, dada la naturaleza de dicho contraste, resulta irrefutable la necesidad de trabajar con entidades que presenten elementos similares. Como se dice comúnmente: «no es posible comparar peras con manzanas», lo que en este caso traduce que no es posible comparar

¹⁴ El análisis de resultados se encuentra en el Apartado 9.

grupos de investigación del área de ciencias exactas con grupos de investigación del área de ciencias humanas.

Es por lo anterior que uno de los requerimientos iniciales para aplicar la metodología DEA es que las DMU cumplan con la característica de homogeneidad, en el sentido de que usen el mismo tipo de recursos para obtener el mismo tipo de resultados, aunque en cantidades variables. En cuanto a esto, Golany & Roll (1989) establecen que:

Un grupo de DMU es homogéneo si:

- Las entidades desempeñan las mismas tareas con objetivos similares.
- Todas las unidades se desempeñan bajo las mismas ‘condiciones de mercado’ (lo cual es de especial importancia cuando se analizan organizaciones sin ánimo de lucro como escuelas, hospitales del estado, etc.)
- Los factores (tanto de entrada como de salida) que caracterizan el desempeño de todo el grupo de DMU, son idénticos, excepto por diferencias en intensidad o magnitud (p.239).

En este caso la homogeneidad está garantizada por el hecho de que los grupos, aunque pertenecen a escuelas distintas, conservan el distintivo de que su objeto y actividad fundamentalmente es la ingeniería. Además, sus insumos y sus productos son exactamente los mismos porque corresponden con los de una institución universitaria.

Ahora, observando en detalle cómo es la distribución de los Grupos de Investigación de la Universidad Industrial de Santander (UIS), se encontró que la institución cuenta con un total de 92 grupos de investigación lo cuales se encuentran clasificados como se refleja en la Tabla 5. Allí se evidencia que las ingenierías es el área con mayor cantidad de grupos de investigación en la UIS, de modo que para el presente trabajo se cuenta con una cantidad significativa de entidades a

evaluar. Así mismo, según Colciencias, el área de las ingenierías es el segundo con mayor cobertura a nivel departamental, Santander (Ver Tabla 6).

Tabla 5
Grupos de investigación en la UIS por áreas del conocimiento

Áreas del conocimiento	Nº de grupos de investigación	Porcentaje	Porcentaje acumulado
Ingenierías Fisicomecánicas	23	25,0%	25,0%
Ingenierías Fisicoquímicas	11	12,0%	37,0%
Ciencias	22	23,9%	60,9%
Ciencias Humanas	13	14,1%	75,0%
Salud	18	19,6%	94,6%
Vicerrectoría de investigación y extensión	3	3,3%	97,8%
IPRED	2	2,2%	100,0%
Total	92		

Nota: Consolidado general de categorías, grupos UIS. Adaptado de Universidad Industrial de Santander (2017). Recuperado de <https://www.uis.edu.co/webUIS/es/investigacionExtension/gruposInvestigacion/index.html>

Tabla 6
Grupos de investigación en Santander por área de conocimiento

Gran Área	Total
Ciencias Agrícolas	3
Ciencias Médicas y de la Salud	50
Ciencias Naturales	47
Ciencias Sociales	61
Humanidades	9
Ingeniería y Tecnología	60

Nota: Estado de la ciencia en Colombia. Convocatoria 737 de 2015. Adaptado de Colciencias (2017). Recuperado de <https://sites.google.com/a/colciencias.gov.co/estado-de-la-ciencia-2015/mapa>

7.1.2 Orientación del modelo. Según las entidades a evaluar, se requiere definir el aspecto decisorio que las involucra, es decir, que el analista debe observar el contexto en el que se encuentran las entidades e identificar con relación a qué factores se encuentran las mejoras potenciales más convenientes. Por ejemplo, en el sector industrial, para una empresa dedicada a la fabricación de algún producto –al contar con un determinado número de producción dependiente de los pedidos realizados por los clientes– resultaría más beneficioso identificar la mínima cantidad de insumos requeridos en la fabricación de determinada cantidad de unidades del producto; pues

buscar un incremento en la producción sin modificar las entradas al proceso no vendría siendo la decisión más adecuada si no se cuenta con la certeza de que tales unidades de producto adicionales van a generar un beneficio económico –considerando que al no tener un cliente fijo para la producción, ésta podría hacer que la organización incurra en gastos adicionales como lo son el almacenamiento. Por otra parte, en el sector educativo –y tratándose en este caso de una universidad pública– el mayor beneficio no vendría dado por una reducción en el nivel de recursos, ni siquiera el más mínimo beneficio podría resultar de ello; mientras que la búsqueda de un incremento de la producción que genera este tipo de entidades resultaría provechosa en términos de prestigio y reconocimiento, lo cual se traduce en un incremento de la demanda estudiantil, en un aumento del número de inscritos en los programas académicos que ofrece el plantel educativo y, por consiguiente, en un aumento en los recursos que el Estado pone a su disposición.

Todo lo anterior hace referencia a la orientación del modelo, esto es, que el modelo deberá enfocarse ya sea a maximizar las salidas del proceso productivo manteniendo constantes las entradas, o bien a buscar la máxima reducción de las entradas para producir la misma cantidad de salidas que actualmente se obtienen.

Dado que en el presente trabajo las entidades a evaluar son los grupos de investigación nacidos dentro de la Universidad Industrial de Santander, no cabe duda de que el enfoque más apropiado es una orientación a las salidas, pues es un despropósito pensar en reducir los insumos de los grupos en procura de una mejor eficiencia. Las reducciones en los insumos, en el caso de darlo, responden a políticas del grupo que no pueden nunca asumirse como la búsqueda de *mayores ganancias* como es propio de entidades con ánimo de lucro.

7.1.3 Selección de variables de entrada y salida. Como lo afirman Torrinco *et al.* (2007) “la selección de las variables es una etapa trascendental en la metodología Análisis Envolvente de

Datos” no sólo porque éstas caracterizan en gran parte el proceso productivo de la entidad a evaluar, sino también por la importancia de hacerlas medibles, es decir, que se puedan traducir en datos que posteriormente se utilicen para ajustar el modelo. Por consiguiente, la recolección de la información resulta ser un aspecto fundamental teniendo en cuenta que los resultados a obtener van a depender de la calidad de la misma (Alberto *et al.*, 2010).

Dentro del análisis bibliométrico se han encontrado trabajos que también pretendieron medir la eficiencia de grupos de investigación, como lo son los estudios realizados por Arenas, Soto y Rivera (2004), Restrepo y Villegas (2007), Villaveces y otros (2007), Giraldo (2007), Ruiz y Otros (2010), Pino, y otros (2009), Rojas (2010), y Visbal, Martínez y Guijarro (2017), a partir de los cuales se han esquematizado las entradas y salidas que utilizan en su medición (Apéndice A) con el fin de identificar aquellas variables que constituyen un componente frecuente en esas investigaciones, y que por consiguiente resultan determinantes para estimar la eficiencia de grupos de investigación según su experiencia.

De este modo, luego de examinar los factores productivos que se toman en consideración en los distintos estudios, ha resultado evidente que el tamaño de los grupos de investigación – entendiéndose como el número de integrantes que conforman cada unidad productiva, sin reparar en el grado de escolaridad de cada uno de ellos– es un elemento reiterativo; y no era de esperarse menos, como se dice popularmente «dos cabezas piensan mejor que una».

Así pues, en cuanto a la elección de las entradas a considerar en el modelo se tiene que: el tamaño de los grupos de investigación es un insumo principal para la obtención de resultados, y dado que es de interés en este trabajo conocer la relación que existe con la productividad, se justifica la intención de tener en cuenta esta característica como entrada. Además, si se observa la lógica del aprendizaje, se podría afirmar que, entre más tiempo transcurra el investigador

desarrollando sus habilidades en la práctica, el conocimiento empírico y el desempeño –del investigador y el grupo al cual pertenece– mejorará, por lo tanto, la experiencia podría considerarse como una variable de entrada razonable. Del mismo modo, se esperaría que el nivel de estudios de los integrantes del grupo de investigación influya positivamente en la productividad de éstos, así que también será un insumo del modelo. Por último, y teniendo en cuenta que este aspecto influye significativamente en el desarrollo de las actividades de los grupos, se hace indispensable tener en cuenta el presupuesto que las entidades reciben como un factor de producción importante.

Respecto a los productos que consideran como variables de salida, podría decirse que todos los estudios utilizan las variables asociadas a la producción de conocimiento, como lo son el número de artículos publicados en revistas, y las tesis y trabajos de grado enfocados a investigación dirigidos por profesores miembros del grupo.

Ahora bien, además de tener en cuenta el criterio propio y los de otros investigadores al estimar la eficiencia de grupos de investigación utilizando DEA, es indispensable considerar el ente que rige y evalúa los grupos de investigación en Colombia con la intención de considerar los aspectos que este órgano considera indispensables en la medición de la eficiencia de grupos de investigación. Así pues, Colciencias –como entidad encargada de regular todos los grupos de investigación formalmente constituidos– clasifica la producción de los grupos de investigación en cuatro grandes áreas: productos de generación de nuevo conocimiento; productos de resultados de actividades de investigación, desarrollo e innovación; productos de apropiación social y circulación del conocimiento y productos de formación de recursos humanos. Cada una de esas categorías comprende distintos subtipos de producción que las caracterizan, y a partir de los cuales Colciencias nutre sus indicadores que resumen la actividad productiva de los grupos de

investigación, a fin de conceder un puntaje de eficiencia a cada uno de ellos y categorizarlos a partir de éste.

En el Apéndice B puede encontrarse una tabla con la clasificación detallada dada por Colciencias a la producción de los grupos de investigación, a partir de la cual se ha decidido considerar adicionalmente, a los Artículos y las Tesis de grado, los productos correspondientes a libros, capítulos de libro y productos de divulgación.

Para el ajuste del modelo se ha utilizado el uso del software *Cplex* desarrollado por IBM en su versión gratuita.

7.2 Recolección de la información

Una vez que planteadas las posibles variables de entrada y salida era necesario saber si la información asociada a ellas estaba disponible en las bases de datos de la Universidad, para lo cual se estableció un periodo de tiempo a analizar de tres años, 2014–2016.

En las averiguaciones previas en la Vicerrectoría de Investigación y Extensión (VIE), se estableció la no disponibilidad de toda la información requerida en ese periodo de tiempo. Según funcionarios de este departamento, la captura de información correspondiente a la variable de salida *Trabajos de grado* se realizó únicamente a partir del año 2016, haciendo imposible la utilización de esta variable en el periodo de tiempo establecido. De igual forma, existieron inconsistencias con la información correspondiente a otras variables, es el caso de aquella correspondiente al dinero asignado a los grupos, que no consideraba en todos los años los mismos rubros.

La falta de información homogénea y consistente impidió tomar el periodo de tiempo correspondiente a 2014–2016, decidiéndose finalmente analizar únicamente el año 2016 pues en éste sí se contaba con toda la información necesaria para realizar el estudio.

Fue así que se solicitó a la VIE la información correspondiente a las variables de entrada o insumos que se refieren al tamaño de los grupos, experiencia o años de antigüedad, estudiantes vinculados al grupo y presupuesto recibido en el año 2016. Dentro de la información suministrada se encontró que los integrantes de los grupos de investigación estaban clasificados según la vinculación con la universidad al año 2016, es decir, para cada miembro se especificaba si era profesor planta o cátedra, estudiante de pregrado, especialización, maestría o doctorado. En el caso de los profesores, se listan los distintos estudios realizados y la fecha en la que éstos fueron finalizados, de modo tal que se procedió a filtrar la información hasta conocer el máximo título obtenido hasta el primer trimestre del año 2016. Sin embargo, no fue posible contar con información adecuada acerca de las horas dedicadas a la investigación.

Igualmente, las posibles variables de salida que se consideraron fueron los artículos clasificados según el reconocimiento de la revista en la cual fueron publicados, los trabajos de grado realizados por los estudiantes vinculados al grupo de investigación, los libros, capítulos de libro y productos de divulgación. La información fue solicitada a la VIE y a la división de Recursos Humanos de la Universidad, y se especificó que ésta correspondiera al año 2016 dado que es el periodo de tiempo que finalmente se decidió analizar en el presente trabajo.

Se recibió la información sobre la producción realizada por los grupos de investigación, concretamente los artículos publicados y las tesis de grado en todos los niveles: pregrado, maestría y doctorado. Los libros, los capítulos de libro y los productos de divulgación no hicieron parte de la base de datos proporcionada por la VIE.

Así pues, en la Tabla 7 se encuentran definidas las variables de entrada y salida a utilizar en la formulación del modelo, con base en la información que se dispone para su conformación.

Tabla 7
Variables para la formulación del modelo DEA

VARIABLES DE ENTRADA	VARIABLES DE SALIDA
Docentes	Artículos indexados
Estudiantes	Trabajos de grado
Años de formación del grupo	
Presupuesto	

Ahora bien, según la literatura correspondiente a la metodología DEA algunos autores proponen la utilización de reglas que determinan la cantidad adecuada de variables a partir del número de entidades que se pretende evaluar, tal como lo afirman Cooper *et al.* (2006) “si el número de entidades es menor que la combinación de entradas y salidas, una gran proporción de entidades serán calificadas como eficientes y el poder de discriminación entre éstas será cuestionable”.

Para solventar este problema, los siguientes autores proponen:

- Golany & Roll (1989):

$$n \geq 2(m + s)$$

Donde n es el número de entidades a evaluar, m el número de variables de entrada, y s el número de variables de salida.

En este caso:

$$34 \geq 2(4 + 2) \rightarrow 34 \geq 12$$

- Charnes, Cooper & Rhodes (1994):

$$n \geq 3(m + s)$$

En este caso:

$$34 \geq 3(4 + 2) \rightarrow 34 \geq 18$$

- Murias (2004):

$$n \geq m \times s$$

En este caso:

$$34 \geq 4 \times 2 \rightarrow 34 \geq 8$$

Resulta evidente que en cualquiera de los casos se cumple la condición de que determinada combinación de variables no supere el número de entidades a evaluar, incluso se estarían cumpliendo estas tres reglas si como mínimo se estudiaran 18 grupos de investigación.

8. Resultados y su discusión

Esta fase conlleva, en primera instancia, el *análisis descriptivo de la información*, el cual permite conocer y resumir los datos de manera que se puedan traducir a información más comprensible. Seguido de ello, se presenta *el tratamiento de la información* en donde se hace posible reflejar la significancia de cada entrada en el proceso de obtención de los productos, y el impacto de estos últimos. Finalmente, se corre el modelo de programación lineal y se identifican aquellos grupos de investigación que resulten eficientes.

8.1 Análisis descriptivo de la información

En primera instancia se realiza este análisis desde el punto de vista de las Escuelas a fin de observar su estado respecto a las actividades de investigación, y luego, más detalladamente, desde el punto de vista de los grupos de investigación, que son las entidades de interés.

El análisis de la información suministrada por la VIE se ha realizado aislando cada variable de entrada y salida y detallando sus categorías. En el caso de los integrantes de cada grupo de investigación, se han apartado las variables Docentes y Estudiantes, y dentro de cada una se ha distinguido el tipo de vinculación con la Universidad para definir sus categorías. En las variables de salida, por ejemplo, la información acerca de la producción de artículos será observada según

el tipo de revista especializada en la que han sido indexados. De igual forma con las demás variables que contemplan categorías decisorias para el análisis.

8.1.1 Análisis por Escuelas. Como se puede apreciar en la Tabla 8, la Escuela de Ingeniería de Sistemas e Informática (EISI) y la Escuela de Ingeniería Eléctrica, Electrónica y de Comunicaciones (E3T) son las que cuentan con la mayor cantidad de grupos de investigación, 6 grupos (18%) cada una. Mientras que las Escuelas de Geología (ESGEO) y de Ingeniería Mecánica (EIM) cuentan con el menor número de grupos, 2 cada una (6%). Las otras escuelas correspondientes a Diseño Industrial, Estudios Industriales y Empresariales, Ingeniería Civil, Ingeniería de Petróleos, Ingeniería Metalúrgica y Ciencia de los Materiales, e Ingeniería Química cuentan con 3 grupos de investigación cada una (9%).

Tabla 8

Cantidad de grupos de investigación por Escuela

Escuela	Cantidad de GI
EISI - Ingeniería de Sistemas e Informática	6
E3T - Ingeniería Eléctrica, Electrónica y de Comunicaciones	6
EDI - Diseño Industrial	3
EEIE - Escuela de Estudios Industriales y Empresariales	3
EIC - Ingeniería Civil	3
EIP - Ingeniería de Petróleos	3
EIMT - Ingeniería Metalúrgica y Ciencia de los Materiales	3
EIQ - Ingeniería Química	3
ESGEO – Geología	2
EIM - Ingeniería Mecánica	2
Total	34

Nota: Consolidado general de categorías, grupos UIS. Adaptado de Universidad Industrial de Santander (2017). Recuperado de <https://www.uis.edu.co/webUIS/es/investigacionExtension/gruposInvestigacion/index.html>

En las Tablas 9 y 10 se observa la clasificación de los grupos de investigación de las facultades de ingeniería según la Escuela a la cual pertenecen. Así mismo, se incluyen las siglas de los

nombres de los Grupos de Investigación (GI)¹⁵ y de las Escuelas, con el fin de facilitar la identificación de éstos a lo largo del documento.

Tabla 9
Grupos de Investigación por Escuela - Facultad Ingenierías Fisicomecánicas

Escuela	Grupo de Investigación
Diseño Industrial (EDI)	GEPS - Ergonomía, producto y significado
	INTERFAZ
	GIROD - Grupo de investigación en robótica de servicio y diseño industrial
Escuela de Estudios Industriales y Empresariales (EEIE)	INNOTEC - Grupo de investigación en gestión de la innovación tecnológica y del conocimiento
	F&M - Finance & Management
	OPALO - Grupo en optimización y organización de sistemas productivos, administrativos y logísticos
Ingeniería Eléctrica, Electrónica y de Comunicaciones (E3T)	CIDLIS - Centro de innovación y desarrollo para la investigación en ingeniería del software
	RadioGIS - Grupo de investigación en radiocomunicaciones y sistemas de información geográfico
	CPS - Grupo de investigación en conectividad y procesamiento de señales
	CEMOS - Grupo de investigación en control, electrónica, modelado y simulación
	GISEL - Grupo de investigación en sistemas de energía eléctrica
	CIDIC - Grupo de diseño de circuitos integrados
Ingeniería Mecánica (EIM)	GIEMA - Grupo de investigación en energía y medio ambiente
	DICBoT - Grupo de investigación en sistemas dinámicos multifísicos, control y robótica
Ingeniería Civil (EIC)	GEOMÁTICA - Gestión y optimización de sistemas
	INME - Grupo de investigación en materiales y estructuras de construcción
	GPH - Grupo de investigación en recursos hídricos y saneamiento ambiental
Ingeniería de Sistemas e Informática (EISI)	GIIB - Grupo de investigación en ingeniería biomédica
	FILOEN - Filosofía y enseñanza de la filosofía
	STI - Grupo de investigación en sistemas y tecnología de la información
	SIMON - Grupo de investigación en modelamiento y simulación
	CAGE - Grupo de cómputo avanzado y a gran escala
HDSP - Grupo de investigación en diseño de algoritmos y procesamiento de datos multidimensionales	

Nota: Consolidado general de categorías, grupos UIS. Adaptado de Universidad Industrial de Santander (2017). Recuperado de <https://www.uis.edu.co/webUIS/es/investigacionExtension/gruposInvestigacion/index.html>

¹⁵ De aquí en adelante, cada vez que se haga referencia a grupos de investigación se utilizará la sigla GI

Tabla 10
 Grupos de Investigación por Escuela - Facultad Ingenierías Físicoquímicas

Escuela	Grupo de Investigación
Geología (ESGEO)	GIGBA - Grupo de investigación en geología básica y aplicada GIGHC - Grupo de investigación en geología de hidrocarburos y carbones
Ingeniería Metalúrgica y Ciencia de los Materiales (EIMT)	GIMAT – Grupo de investigación en desarrollo y tecnología de nuevos materiales GIC – Grupo de investigación en corrosión GIMBA – Grupo de investigación en materiales, biohidrometalúrgica y ambiente
Ingeniería de Petróleos (EIP)	GMPH - Modelamiento de procesos de hidrocarburos GIEP - Grupo de investigación en estabilidad de pozo GRM - Grupo de investigación recobro mejorado
Ingeniería Química (EIQ)	CIDES - Centro de investigación para el desarrollo sostenible en industria y energía INTERFASE - Grupo de investigación en tecnologías de valorización de residuos y fuentes agrícolas e industriales para la sustentabilidad energética GIP - Grupo de investigación en polímeros

Nota: Consolidado general de categorías, grupos UIS. Adaptado de Universidad Industrial de Santander (2017). Recuperado de <https://www.uis.edu.co/webUIS/es/investigacionExtension/gruposInvestigacion/index.html>

En el año 2016 la Universidad Industrial de Santander contaba con 454 docentes vinculados a las facultades de ingeniería en la sede principal –171 planta y 283 cátedra–, de los cuales 223¹⁶ hacían parte de grupos de investigación (166 planta y 57 cátedra). Las Escuelas que mostraron en ese año una mayor participación docente en actividades de investigación fueron E3T y EISI, con un mayor porcentaje de profesores adscritos a sus GI, 20,2% (53 docentes en total, de los cuales 38 son de tiempo completo) y 14,5% (correspondiente a 33 docentes planta y 5 cátedra) respectivamente. Además, como se observa en la Figura 5, es importante señalar que, en todas las Escuelas, el total de profesores planta vinculados a GI sobrepasa al total de docentes cátedra.

¹⁶ Este valor contempla el hecho de que hay profesores que pertenecen a más de un GI, por tal razón no se obtiene de sumar el total de profesores vinculados a cada grupo de investigación por escuela.

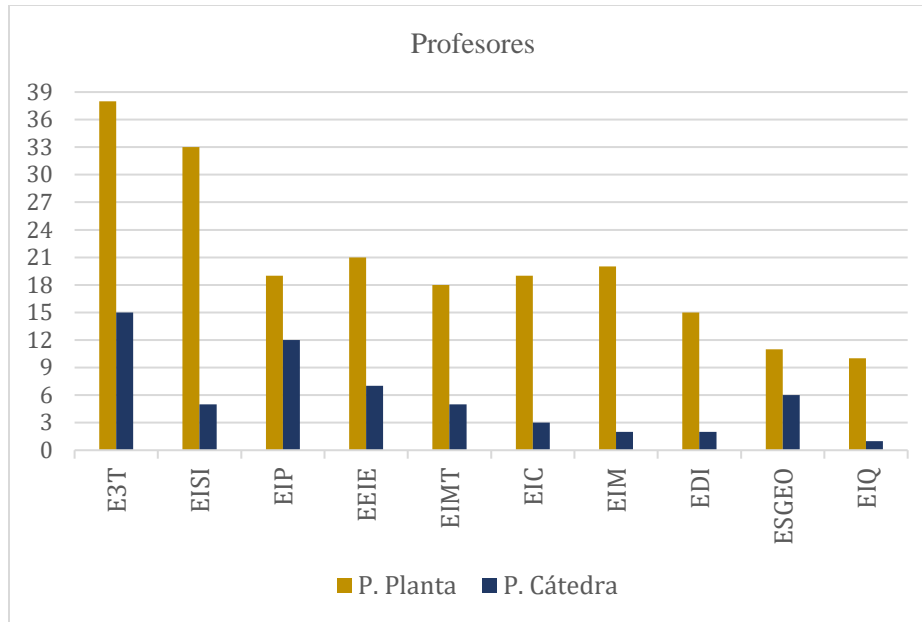


Figura 5. Profesores vinculados a grupos de investigación por Escuela. Adaptado de la información suministrada por la VIE (2017).

Si se trata de los estudiantes, hay que señalar que para el año 2016 la universidad contaba con un total de 14.343 estudiantes de pregrado en su campus principal, de los cuales, según datos suministrados por la Vicerrectoría de Investigación y Extensión (VIE), 55 de ellos se encontraban adscritos a algún GI de ingeniería. Como se observa en la Figura 6, las Escuelas ESGEO y EIQ contaron con la mayor participación por parte de éstos, pues abarcan el 25,5% (14) y el 20% (11) de los estudiantes de pregrado vinculados a GI en ese año, respectivamente.

Ahora bien, de los 1.150 estudiantes de maestría en el año 2016, 122 estaban vinculados a grupos de investigación de ingeniería y, como se observa en la Figura 6, de estos últimos el 31,9% (39) pertenecieron a GI en la EISI, seguido por un 19,7% (24) que se encontraron adscritos a grupos de la E3T. Las Escuelas con menor número de estudiantes de maestría vinculados a grupos de investigación son EIC y ESGEO, con tan sólo el 1,6% de los 122 estudiantes (2 estudiantes cada una).

Además, en el año 2016 la universidad contaba con un total de 131 estudiantes de Doctorado de los cuales 41 hacen parte de algún GI de ingeniería. De estos últimos, el 31,7% (13) estaban vinculados a la Escuela de Sistemas, EISI, como se observa en la Figura 6. Cabe resaltar que las Escuelas EIC, EDI, EIP, ESGEO, EEIE y EIM contaban con dos estudiantes a lo más, de hecho, entre todas ellas solo tenían 5 estudiantes de doctorado, es decir, el 12,2% del total.

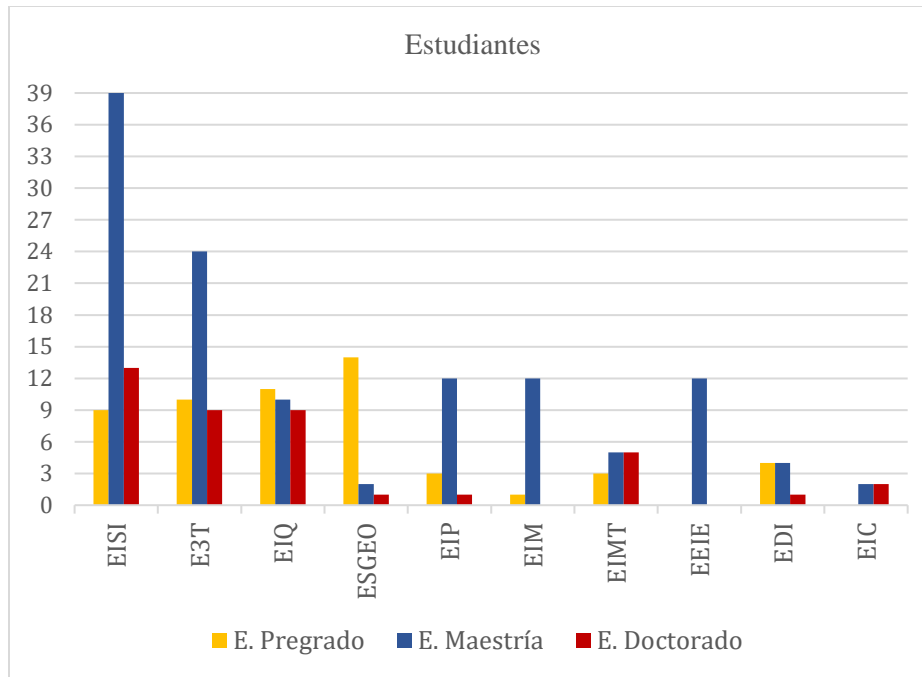


Figura 6. Estudiantes vinculados a grupos de investigación por Escuela. Adaptado de la información suministrada por la VIE (2017).

Entonces, si se reúnen todos los estudiantes en una única categoría, se tendría que para el año 2016 en el campus central de la universidad, 218 estudiantes comprendidos en las modalidades de pregrado, maestría y doctorado, y un estudiante de especialización adscrito a un GI de la E3T, participaron en actividades de investigación; y que 3 de las 10 escuelas de las facultades de ingeniería, abarcan el 61,2% de estos 219 estudiantes, EISI (61), E3T (44) y EIQ (30).

Tratándose ahora de las salidas del modelo, de acuerdo con los datos suministrados por la VIE, en el año 2016 se realizaron 186 publicaciones en revistas especializadas (A1, A2, B y C), excluyendo la categoría *Otros*, por parte de los 34 GI de las ingenierías (véase Tabla 11).

Tabla 11
Publicación según el tipo de revistas

Tipo de revista	Total	Porcentaje
A1	54	26,2%
A2	51	24,8%
B	54	26,2%
C	27	14%
Otros	19	9%
Total	205	

Como se puede observar en la Figura 7, EISI es la escuela que más publicaciones totales realizó (42 artículos), seguida de la E3T, EIMT y EIQ con 27 artículos cada una. EISI además fue la Escuela que indexó una mayor cantidad de artículos en revistas tipo A1 (16), seguida por la Escuela de Ingeniería Química con 15 artículos. La Escuela de Ingeniería Mecánica, por su parte, presentó una producción nula en esa categoría.

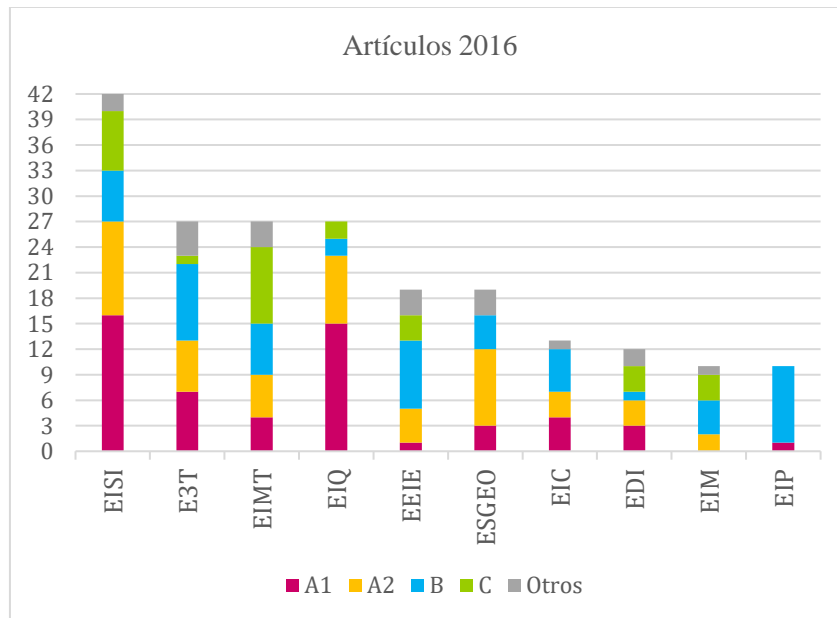


Figura 7. Producción de artículos por Escuela en 2016. Adaptado de la información suministrada por la VIE (2017).

Considerando todos los grupos, se produjeron 51 artículos en revistas tipo A2 de los cuales el 21,6% (11) fueron publicados por la EISI, siendo ésta la mayor producción por escuela, seguida por ESGEO con el 17,6% (9 artículos). La Escuela de Petróleos no reflejó producción alguna en revistas A2.

Del total de publicaciones en revistas especializadas tipo B, el 33,3% se dio en grupos de investigación de la E3T y de la EIP (9 artículos en cada Escuela), seguido por el 14,8% correspondiente a 8 artículos publicados por GI de la EEIE. La escuela que menos artículos indexó en estas revistas fue Diseño Industrial (1 artículo), sin embargo, a diferencia de EIP, EDI sí realizó publicaciones en revistas tipo A2 y C.

Así mismo, se puede observar que la mayor producción de artículos tipo C se dio en la EIMT, con 9 artículos correspondientes al 33,3%, seguido por la EISI con el 26% de un total de 27 publicaciones tipo C (7 artículos). Las Escuelas EIC, EIP y ESGEO no tuvieron producción en este sentido.

Los artículos publicados en revistas no indexadas u homologadas por Colciencias (Véase *otros* en la Figura 7) es la categoría con menor producción en los GI analizados para el periodo de tiempo evaluado con un total de 19 artículos publicados, de los cuales el 21,1% (4) corresponde a grupos de investigación de la E3T. Las demás Escuelas presentan una menor cantidad de artículos en esta categoría, EIP y EIQ no tuvieron producción de este tipo de artículos.

En general, es evidente que las Escuelas de la Facultad de Ingenierías Fisicomecánicas son las que cuentan con los mayores niveles en los insumos *Docentes* y *Estudiantes*, así como en los dos productos considerados (*Tesis* y *Artículos*). Lo anterior era de esperarse si se tiene en cuenta que dicha Facultad cuenta con dos escuelas más en comparación con la Facultad de Ingenierías Fisicoquímicas, así como una cantidad superior de GI.

8.1.2 Análisis por grupos de investigación. Una vez realizado el análisis por escuelas, resulta pertinente desglosar la información por grupos de investigación teniendo en cuenta que estos últimos son el eje central de esta investigación. A diferencia del análisis realizado en el apartado anterior, en éste se tendrán en cuenta la totalidad de variables listadas en la Tabla 7, es decir, aquí serán examinados también los *Años de conformación* y el *Dinero asignado* de cada grupo.

La Universidad Industrial de Santander cuenta con 34 grupos de investigación en las facultades de ingeniería, de los cuales 17 (50%) fueron clasificados por Colciencias, en la convocatoria 737 de 2015, en las categorías B y C (8 y 9 respectivamente), 7 GI en la categoría A1, y sólo los grupos CIDIC y GIROD, pertenecientes a la Facultad de Fisicomecánicas, fueron clasificados como No Reconocidos.

Por cada grupo de investigación la VIE suministró la información correspondiente al año 2016 de los integrantes de cada uno de ellos, clasificados por el tipo de vinculación que a ese año tenían con la Universidad; así mismo, entregaron la fecha exacta en la que fueron conformados, el dinero asignado por estímulo y movilidad, el dinero asignado a proyectos que se encontraban vigentes al año en cuestión, el número de artículos publicados según la revista en la que fueron indexados y el número de trabajos de grado dirigidos en cada GI según la modalidad estudiantil.

Así pues, en lo relativo a los miembros de cada grupo, fue posible observar que 8 de los 34 GI no cuentan con profesores cátedra vinculados, lo que corresponde al 23,5%; y que 19 de los 34 GI analizados están conformados al menos en un 50% por profesores planta, dejando nuevamente en evidencia una fuerte participación por parte de estos docentes en cada GI, salvo en CPS (de la E3T) donde los docentes cátedra duplican en número a los de tiempo completo. De igual manera, se puede observar que GIEMA, GISEL y CEMOS –éstos dos últimos pertenecientes a la E3T– son los grupos con mayor cantidad de docentes planta vinculados (16, 15 y 11 docentes

respectivamente); y que CPS, GMPH y GIGBA son los grupos con mayor participación de catedráticos con 8, 7 y 5 respectivamente (Ver Figura 8).

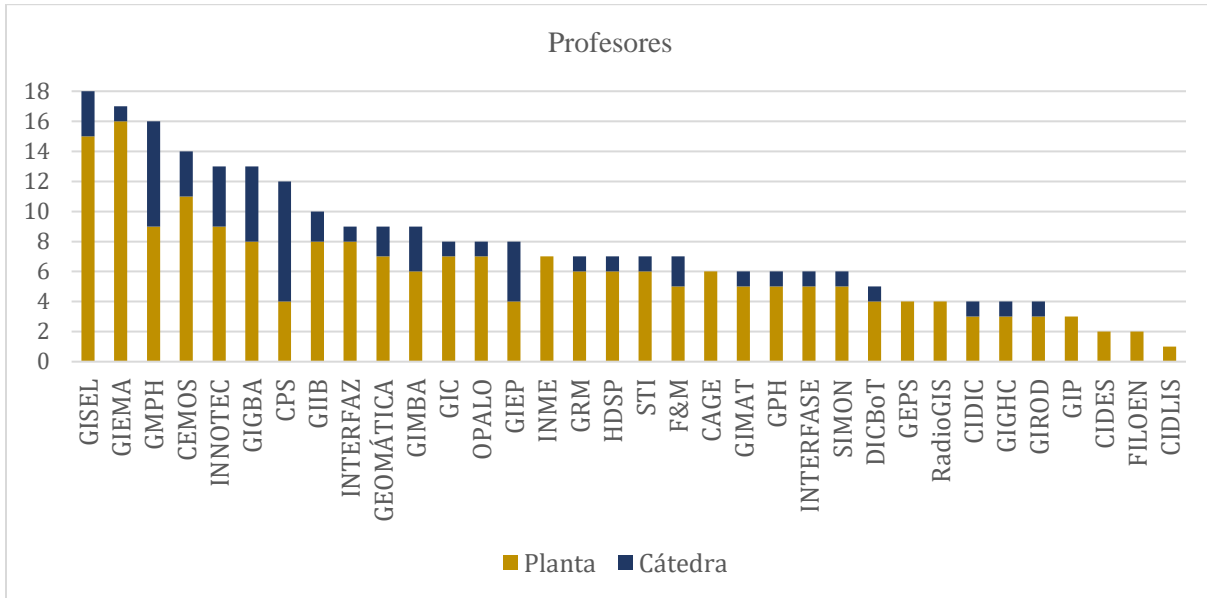


Figura 8. Docentes vinculados a cada grupo de investigación. Adaptado de la información suministrada por la VIE (2017).

Otro aspecto significativo acerca de los integrantes de los GI, resulta ser la existencia de grupos conformados únicamente por profesores, siendo estos: GPH, GIROD y FILOEN; o bien GI constituidos tan sólo en un 17,6% por docentes, como es el caso de GIP que de un total de 17 integrantes únicamente 3 de ellos son profesores.

Siguiendo con los asociados a GI, respecto a los educandos, se observa en la Figura 9 que los estudiantes de maestría tienen una fuerte participación, siendo EISI la Escuela que tiene los grupos de investigación con mayor número de éstos: HDSP y GIIB con 12 (9,8%) y 11 (9%) miembros respectivamente; además de CAGE que es el grupo conformado en su mayoría por estos estudiantes (46,7%), pues de un total de 17 miembros entre docentes y estudiantes en las diversas categorías, ésta es la que prevalece con un total de 7 estudiantes de maestría; seguido por HDSP en el que 12 (44,4%) de sus 27 integrantes son estudiantes de maestría y 7 (25,9%) son estudiantes de doctorado, lo que lo convierte en el grupo de investigación de las ingenierías con la mayor

cantidad de estudiantes de posgrado vinculados (19 en total). Sin embargo, el GI conformado en su mayoría por estudiantes de doctorado resulta ser CIDES, con 4 (44,4%) de sus 9 miembros en este rubro, seguido por INTERFASE con 5 (33,3%) de 15 miembros totales. Cabe aclarar que sólo 17 de los 34 grupos de investigación cuentan con estudiantes de doctorado vinculados y que, de los estudiantes de especialización, RadioGIS es el único grupo que cuenta con un integrante vinculado.

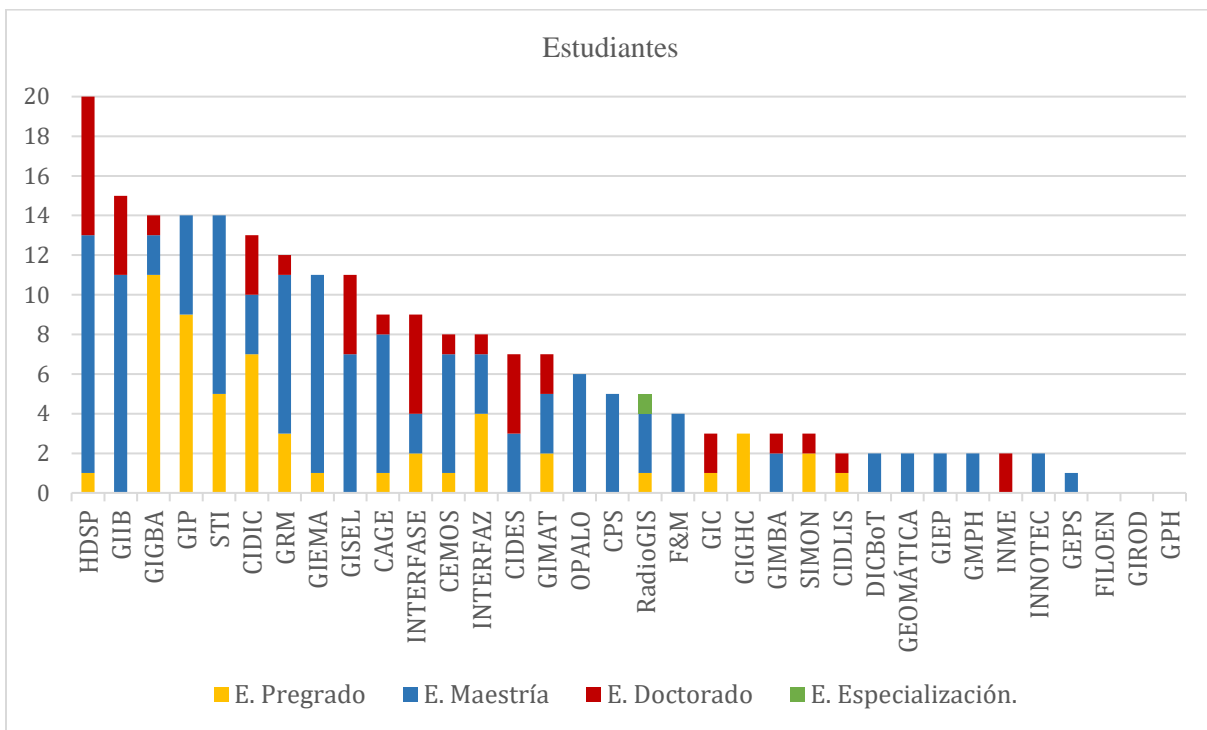


Figura 9. Estudiantes vinculados a cada grupo de investigación. Adaptado de la información suministrada por la VIE (2017).

Otro aspecto importante que analizar en los grupos de investigación es el presupuesto total que se les asigna, y que corresponde a la suma de los presupuestos de cada proyecto aprobado, los estímulos que se les otorga y el total financiado en términos de movilidad de investigadores.

En la Figura 10 se observa cómo el rubro *proyectos* es el que predomina en 15 de los 34 grupos de investigación (44,11%); así mismo, en el 26,65% (9 GI en total) más del 90% del dinero total recibido corresponde a este mismo ítem, *proyectos*. En cuanto a *movilidad* cabe resaltar que en los grupos GMPH y CIDIC el total de dinero recibido es por este rubro, es decir, no registran dinero

por *estímulos* o *proyectos*. Así mismo, los grupos GIEP, STI, y CIDLIS no recibieron por concepto de *movilidad* ni por *proyecto*, eso quiere decir que el 100% recibido fue por *estímulos* a pesar de tener una trayectoria de 13,33 años, 15,92 y 26 respectivamente.

Los grupos CEMOS, HDSP, GIMAT, INTERFASE y CIDIC recibieron, entre los cinco, más de la mitad (53,97%) del total destinado a las facultades de ingeniería en el año 2016. En contraste, el 50% de los GI (17 grupos) recibieron, cada uno, menos del 1% del dinero total en ese mismo año. Por último, es necesario resaltar el hecho de que tres grupos, GEOMÁTICA, GIROD y RadioGIS no recibieron durante el año 2016 dinero por ninguno de los tres rubros considerados.

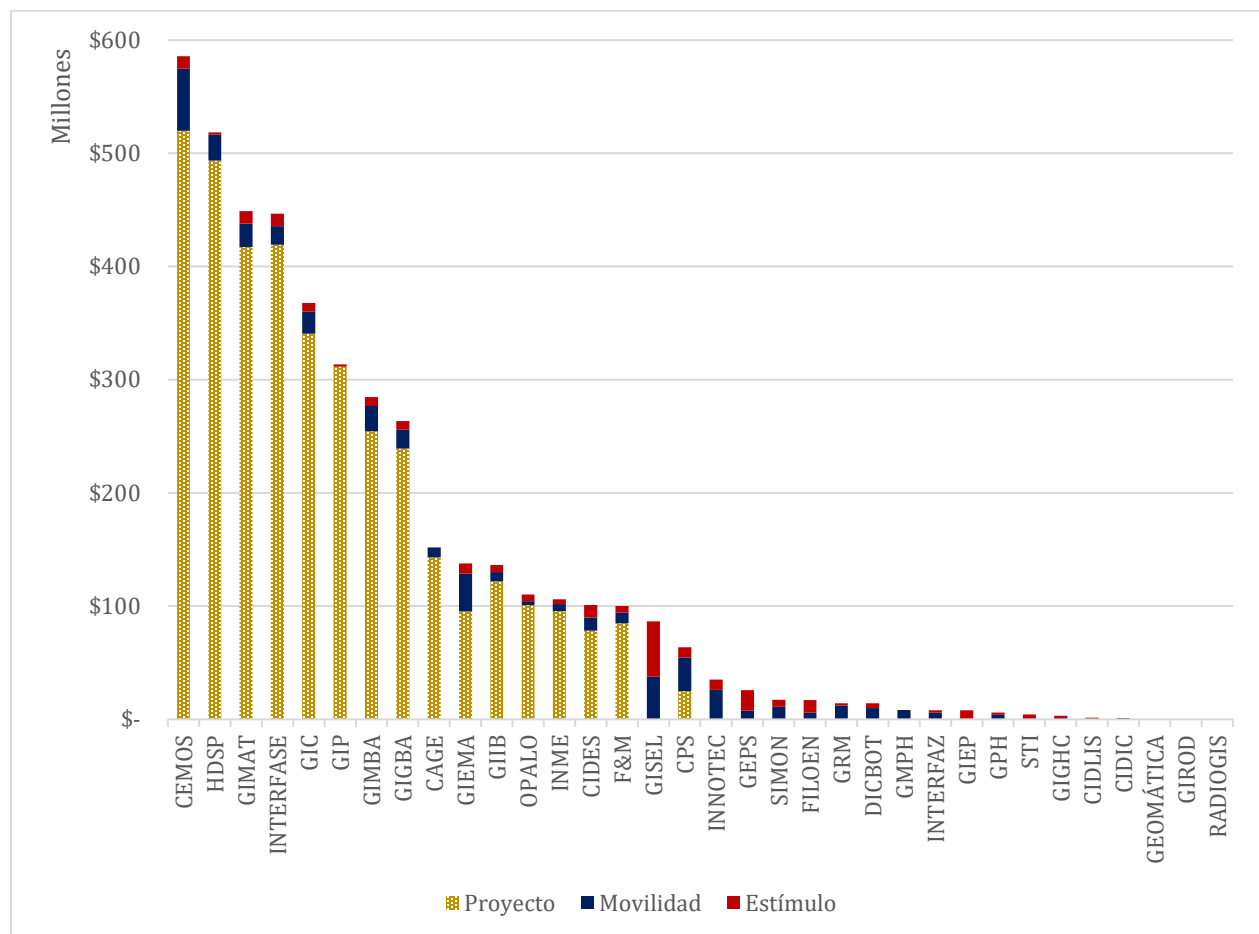


Figura 10. Dinero asignado a los grupos de investigación. Adaptado de la información suministrada por la VIE (2017).

En el caso de la experiencia o años de antigüedad de los grupos (ver Figura 11), se encontró que GIP, de la Escuela IQ, es el que mayor trayectoria tiene, cuenta con más de 40 años de

experiencia. Posterior a éste están los grupos GIC y GIMBA de la Escuela de Metalúrgica con 31 y 30 años respectivamente. Es de resaltar que el 41,17% (14) de los 34 GI analizados, tienen una trayectoria de al menos 20 años; y que los grupos GIROD, HDSP y CAGE son aquellos con menor tiempo de experiencia, contando con menos de 10 años desde su formación.

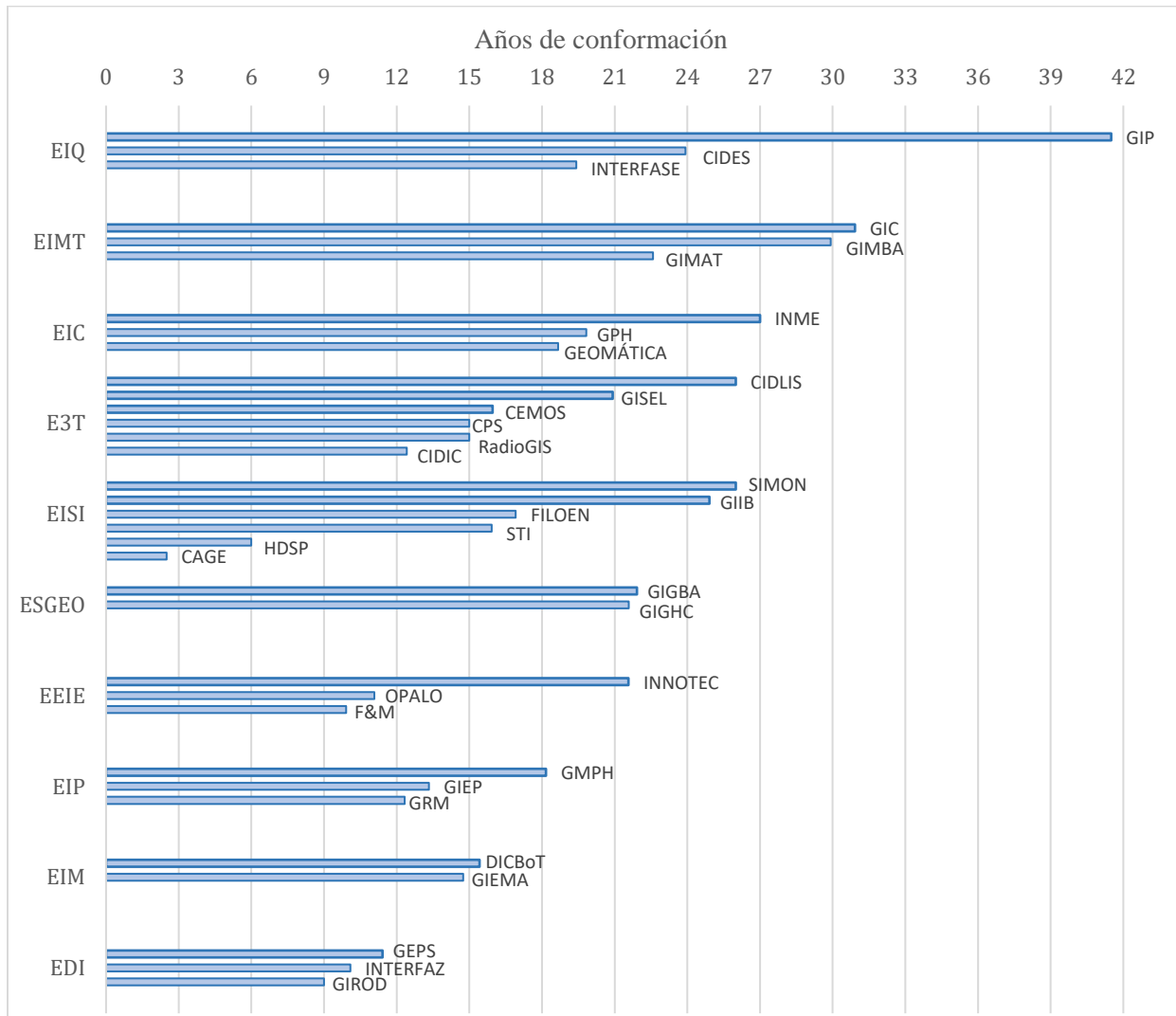


Figura 11. Años de conformación de cada grupo de investigación. Adaptado de la información suministrada por la VIE (2017).

La Figura 11 muestra los años transcurridos desde la fecha de conformación de cada grupo hasta diciembre de 2016, lo que permite afirmar que EIQ es la Escuela cuyo compromiso con la investigación resultó ser pionero al crear el Grupo de Investigación en Polímeros (GIP),

manteniéndose por 10 años como el único grupo de las facultades de ingeniería, hasta que en la EIMT se crearon GIC y GIMBA.

Por otra parte, la Figura 12 muestra el número de artículos publicados en el año 2016 según las revistas en las que fueron indexados. Como se puede observar, el grupo que más producción tuvo en ese año fue GIGBA, grupo de investigación de la Escuela de Geología, seguido por HDSP, de la Escuela de Ingeniería de Sistemas, con un total de 15 y 14 publicaciones respectivamente; sin embargo, no representan un porcentaje significativo frente a la producción total realizada por las facultades de ingeniería, pues tan sólo corresponden al 7,3% y al 6,8% respectivamente. Por su parte, los grupos de investigación GIROD, CIDIC y FILOEN, pertenecientes a la EDI, E3T y EISI respectivamente, no realizaron ninguna publicación en ese año –aun cuando presentan una trayectoria superior a diez años– sin embargo, como se mencionó anteriormente, estos dos grupos no recibieron dinero por proyectos, pero sí dinero por movilidad y/o estímulos.

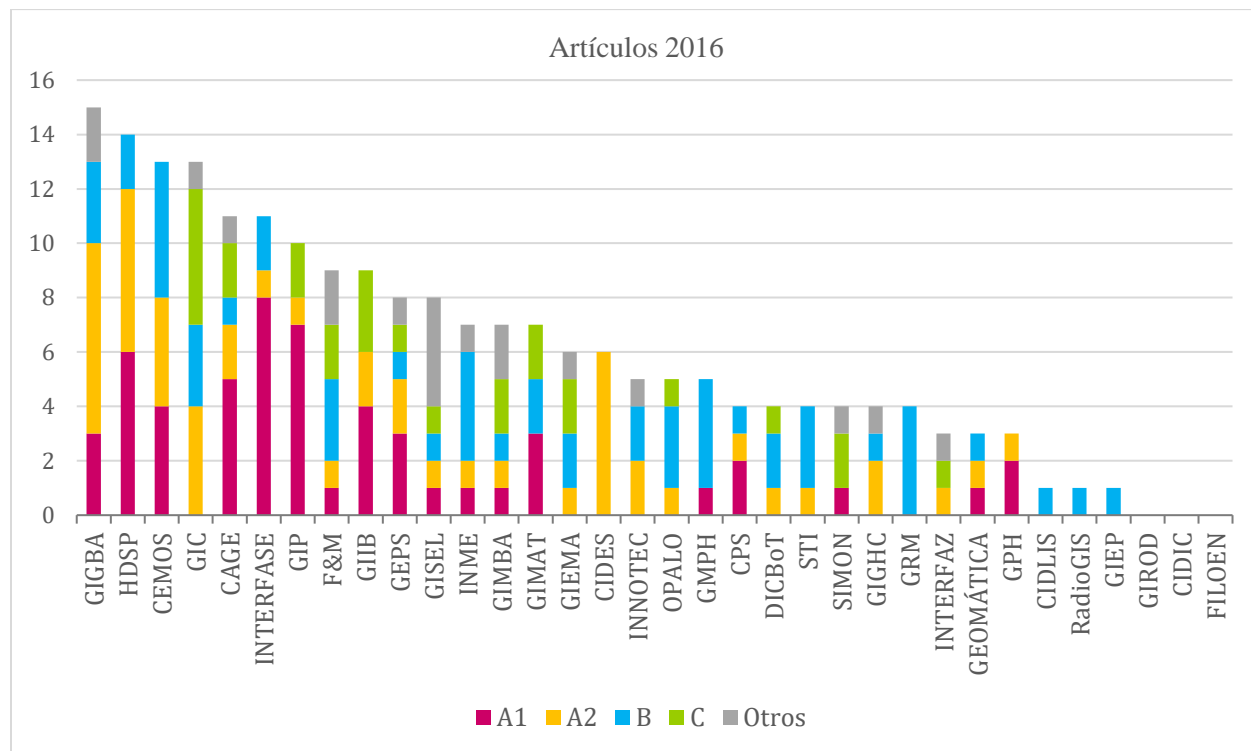


Figura 12. Producción de artículos por grupo de investigación en 2016. Adaptado de la información suministrada por la VIE (2017).

Además, del total de 206 publicaciones realizadas en el año 2016 por los GI de las facultades de ingeniería, tan sólo el 9,2% (19) correspondió a publicación de artículos en revistas no indexadas u homologadas por Colciencias (Ver *otros* en la Figura 12); mientras que el 26,2% (54) correspondió a publicaciones en revistas tipo A1; 24,8% (51) a publicaciones en revistas A2; 26,2% (54) en revistas tipo B, y el 13,6% (28) a publicaciones en revistas tipo C.

En cuanto a la calidad de las publicaciones, hay que apuntar que INTERFASE fue el grupo de investigación que del total de su producción presenta un mayor porcentaje de artículos publicados en revistas especializadas A1 (72,7% correspondiente a 8 artículos de 11 en total), seguido por GIP con un 70% de su producción indexada en este tipo de revistas (7 de 10 artículos), CIDES publicó en revistas especializadas tipo A2 el total de su producción, todos ellos pertenecientes a la Escuela de Ingeniería Química.

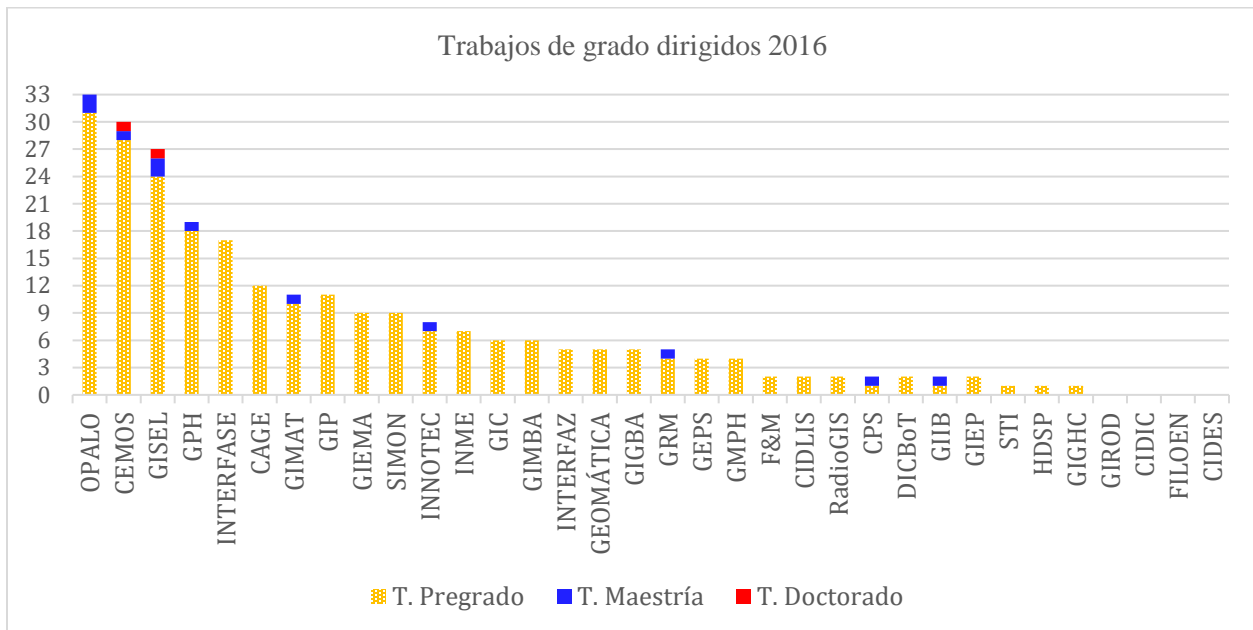


Figura 13. Trabajos de grado dirigidos a 2016. Adaptado de la información suministrada por la VIE (2017).

Finalmente, la información correspondiente al número de trabajos de grado dirigidos por los grupos de investigación de las facultades de ingeniería se muestra en la Figura 13. El total de tesis dirigidas en el año 2016 fue de 255 de las cuales el 94,9% (242) corresponde a tesis de pregrado,

el 4,3% (11) a tesis de maestría y tan sólo el 0,8% (2) a tesis doctorales; siendo estas últimas dirigidas por los grupos de investigación CEMOS y GISEL, pertenecientes a la E3T.

En definitiva, los grupos que se destacan por tener la mayor cantidad tanto de insumos como de producto son: GISEL que encabeza el ranking en la variable *Docentes*; si se trata de *Estudiantes* el primer lugar lo ocupa HDSP; en cuanto a *Dinero asignado* CEMOS se gana la primera posición, y para finalizar con los insumos GIP es el grupo con más tiempo de trayectoria, *Años de conformación*. En los productos, HSDP es el GI que más producción tuvo en términos de *Artículos*, sin embargo, INTERFASE y GIGBA fueron los grupos con mayor cantidad de artículos A1 y A2 respectivamente. En el caso de *Tesis dirigidas* OPALO es el que mayor aporte hizo en este sentido, siendo el grupo con mayor cantidad de tesis tanto en pregrado como en maestría, CEMOS y GISEL fueron los únicos grupos con dirección de tesis doctorales.

Por otro lado, es importante considerar el diagnóstico estadístico de cada variable analizada previamente. En la Tabla 12 se presentan los estadísticos: Media, Mediana, Moda, Desviación, y los Percentiles de mayor interés, correspondientes a las variables analizadas y sus categorías, según la información de los 34 grupos de investigación (entidades) para el año 2016.

Los estadísticos fueron obtenidos a través del software SPSS de IBM, y en ellos resulta interesante observar que, en el caso de la producción de Artículos, el 75% de los grupos de investigación presenta cuando mucho 3 artículos A1, y 2 artículos A2; mientras que en la dirección de tesis de pregrado es evidente que en el 75% de las entidades la cantidad de trabajos de grado, en la modalidad de pregrado, no supera los 9.

Según las medidas de tendencia central, resulta evidente que los Años de conformación de los grupos presenta un comportamiento simétrico, pues los valores de la Media y la Mediana son cercanos, así mismo, indican que al menos la mitad de los grupos tiene más de 17 años de

trayectoria. Adicionalmente, es posible identificar que al menos la mitad de los grupos cuenta con docentes que tienen un nivel de formación superior a la especialización.

Tabla 12

Estadísticos descriptivos de las variables

		Estadísticos						
		Media	Mediana	Moda	Desviación	Percentiles		
						25	50	75
Años		18,291	17,545	15	7,864	12,4	17,55	22,92
Docentes Planta	Pregrado	0,118	0	0	0,327	0	0	0
	Especialista	0,206	0	0	0,479	0	0	0
	Magíster	1,941	1	1	1,858	0	1	3,25
	Doctor	3,735	3	1	2,767	1	3	5,25
Docentes Cátedra	Pregrado	0,529	0	0	0,992	0	0	1
	Especialista	0,147	0	0	0,436	0	0	0
	Magíster	0,882	1	0	1,066	0	1	1
	Doctor	0,147	0	0	0,359	0	0	0
Estudiantes	Pregrado	1,618	0,5	0	2,697	0	0,5	2
	Maestría	3,588	2,5	0	3,421	0,8	2,5	6
	Doctorado	1,206	0,5	0	1,754	0	0,5	2
Presupuesto	Estímulo	6,768	6	0	8,67	2	6	9
	Movilidad	12,222	8,505	0	12,815	1,3	9	19,66
	Proyecto	110,039	0	0	161,895	0	0	167,12
Artículos	A1	1,588	1	0	2,217	0	1	3
	A2	1,500	1	1	1,83	0	1	2
	B	1,588	1	1	1,417	0	1	3
	C	0,824	0	0	1,167	0	0	2
Tesis	Pregrado	7,118	5	1	8,297	1	5	9,25
	Maestría	0,324	0	0	0,589	0	0	1
	Doctorado	0,059	0	0	0,239	0	0	0

Si se observan las tres categorías de Estudiantes (Pregrado, Maestría y Doctorado) se evidencia que los de maestría son los que cuentan con la desviación más alta, es decir, que existe una mayor variación de los datos, indicando que en promedio el número de estudiantes de maestría por grupo de investigación se aleja de la media aproximadamente en 3. De forma similar, para las Tesis dirigidas, la categoría que presenta una mayor desviación es la correspondiente a las Tesis de pregrado, indicando que en promedio el número de estos trabajos de grado por grupo se aleja de

la media aproximadamente en 8; además se puede observar que las desviaciones de las categorías de Tesis doctorales y de maestría son valores cercanos a cero, es decir, que la cantidad de tesis dirigidas tanto en maestría como en doctorado por cada uno de los grupos es muy similar, y que al menos la mitad de los grupos no exhiben tesis doctorales y de maestría dirigidas por ellos.

Importante señalar que las medidas de tendencia central correspondientes a la variable *Presupuesto* hacen evidente que al menos la mitad de los grupos de investigación no recibe dinero por el rubro de Proyectos en el año 2016, y que en promedio el dinero por Proyecto recibido por grupo se desvía de la media aproximadamente en 162 millones. En cuanto al dinero asignado por Estímulos se observa gran similitud entre los valores de la media y la mediana, es decir, los datos correspondientes a este rubro son simétricos.

8.2 Tratamiento de la información

Para dar lugar al modelo DEA orientado a las salidas es importante considerar que las variables sólo soportan una medida, es decir, si las variables presentan categorías diferenciadoras es necesario unificarlas a fin de obtener una única medida por cada una de ellas. Por ejemplo, en el caso del insumo *Docentes vinculados*, la información los diferencia según el tipo de vinculación con la universidad y el máximo nivel de formación alcanzado, haciendo necesario determinar el peso (importancia) de dichas categorías para poder agrupar los docentes en una única variable.

Al aplicar este razonamiento a las salidas del modelo, el producto *Artículos de investigación* será observado según el tipo de revista en la que fue publicado y conforme a lo que establece el decreto 1279 del año 2002 acerca del reconocimiento asignado a cada docente investigador que realice publicaciones en revistas especializadas; de este modo, cada publicación efectuada recibirá un puntaje de acuerdo a la Tabla 13. Para facilitar el análisis de los resultados, se va a realizar el

cociente entre el puntaje total obtenido por cada grupo y la mayor ponderación asignada (15), de modo tal que la unidad de medida de esta variable son los artículos A1.

Tabla 13

Ponderación según el tipo de revista para cada publicación.

Tipo de revista	Puntaje
A1	15
A2	12
B	8
C	3

Nota: Adaptado del Decreto 1279 de junio 19 de 2002, por el cual se establece el régimen salarial y prestacional de los docentes de las Universidades Estatales. Artículo 10, inciso a. Consultado en 2018.

En el caso de los trabajos de grado asociados a cada grupo de investigación, se tendrá en cuenta la modalidad a la que pertenecen, pues no sería preciso equiparar tesis de pregrado con tesis de doctorado, de modo que para destacar las diferencias entre una y otra tesis se va a considerar el tiempo en que éstas deben realizarse, es decir, para el caso de una tesis de pregrado se han establecido dos periodos académicos en los cuales el estudiantes cursa Trabajo de grado I y trabajo de grado II, lo que corresponde a un año en términos de tiempo, y que en este caso se tomará como un (1) punto por cada uno de los realizados por el grupo de investigación; con la misma argumentación para una tesis de maestría se asignan dos (2) puntos y para una doctoral cuatro (4) puntos. Adicionalmente, se tomará como unidad de medida el número de tesis doctorales, de tal manera que el valor de esta variable para cada grupo es el cociente entre el total obtenido por el grupo dividido por 4.

Al considerar la variable de entrada docentes, es conveniente en primera instancia diferenciar los profesores planta y cátedra, para lo cual se ha tenido en cuenta el número de horas de dedicación semanales de cada uno de éstos, y se ha procedido a establecer la relación existente. De los datos obtenidos a través de UIS en cifras 2016 se conoce que el total de horas cátedra semanales contratadas en las facultades de ingeniería es de 2.868, y teniendo en cuenta que el número total

de profesores cátedra vinculados a éstas es 283, se obtiene un tiempo promedio de 10,13 horas semanales por profesor; mientras que los profesores de tiempo completo ejercen 40 horas semanales, lo cual indica la existencia de una relación 1:4 entre estos dos tipos de vinculación. Es por lo anterior que se ha decidido puntuar a los docentes planta con un valor de 4 y a los docentes cátedra con 1, sin embargo, para que la unidad de medida esté en función de los docentes de tiempo completo cada ponderación será dividida en 4. Esta variable tendrá una ponderación adicional correspondiente al nivel académico del docente, la cual tendrá el mismo argumento de ponderación que la variable *Tesis de grado*.

Por otro lado, el insumo *Estudiantes* también cuenta con subcategorías (pregrado, especialización, maestría y doctorado), por lo cual se hace necesaria la unificación de éstas para obtener una única variable. Para llevar a cabo tal unificación la ponderación es la misma usada para el insumo *Tesis de grado* y que se puede observar con mayor claridad en la Tabla 14.

Tabla 14
Puntaje asignado según el máximo título obtenido.

Título	Puntaje
Pregrado	1
Especialización	1.5
Maestría	2
Doctorado	4

En el Apéndice C se encuentran los valores originales de cada entrada y salida, según la información suministrada por la VIE, y su correspondiente valor ponderado.

8.3 Correlación bivariada.

Otro aspecto importante a conocer es cómo se asocia cada variable de salida con cada variable de entrada, por grupo de investigación. Para ello se ha realizado un diagrama de dispersión

enfrentando Artículos y Tesis (variables de salida) con Presupuesto, Años, Docentes y Estudiantes (variables de entrada).

La Figura 14 muestra el producto *Artículos* versus el insumo *Presupuesto*, allí se puede apreciar que estas dos variables se correlacionan altamente de forma positiva, obteniendo un coeficiente de correlación de Pearson (R) de 81,5%.

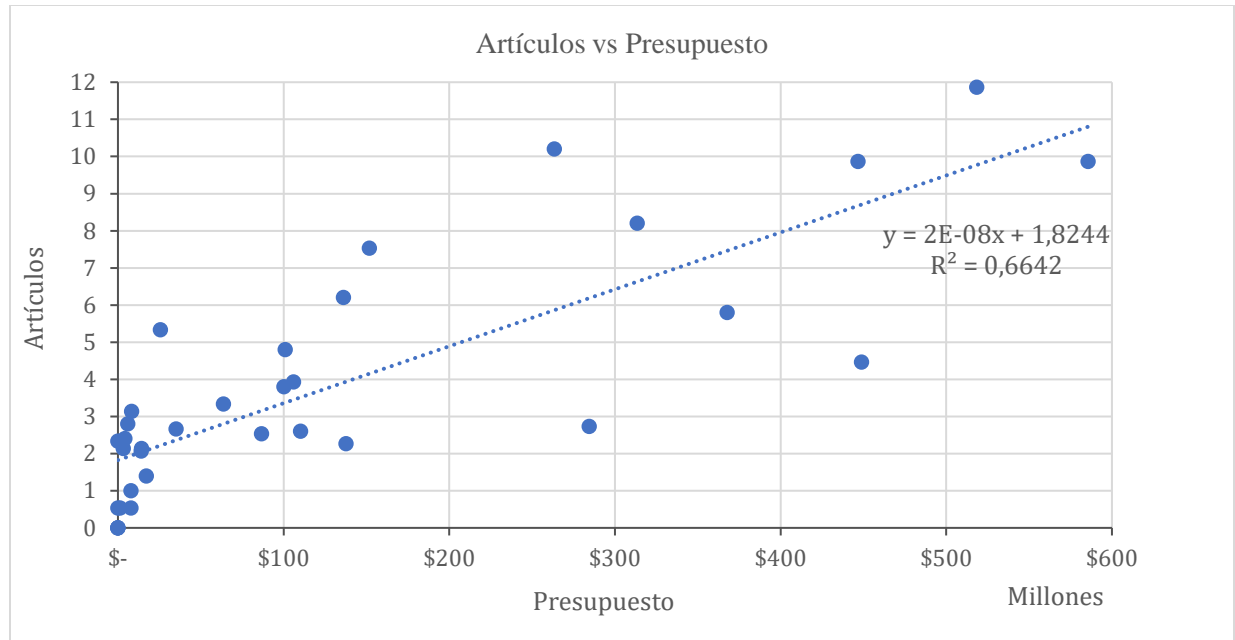


Figura 14. Dispersión cantidad de Artículos vs Presupuesto.

Por el contrario, al enfrentar el insumo *Presupuesto* con el producto *Artículos*, se observa que el insumo apenas explica el 12,87% de la variabilidad en la producción de *Tesis*, de modo que presentan una baja correlación positiva: $R = 35,9\%$. Sin embargo, esto se puede dar debido a que el coeficiente de correlación de Pearson (R) es sensible a valores extremos, en este caso aquellos grupos que presentan gran cantidad de dirección de *Tesis* y bajo *Presupuesto*, o en sentido contrario, de modo que al excluir esas observaciones extremas (Figura 15) el coeficiente de correlación R aumenta considerablemente, tomando un valor de 57,4% e indicando una relación moderada entre las variables.

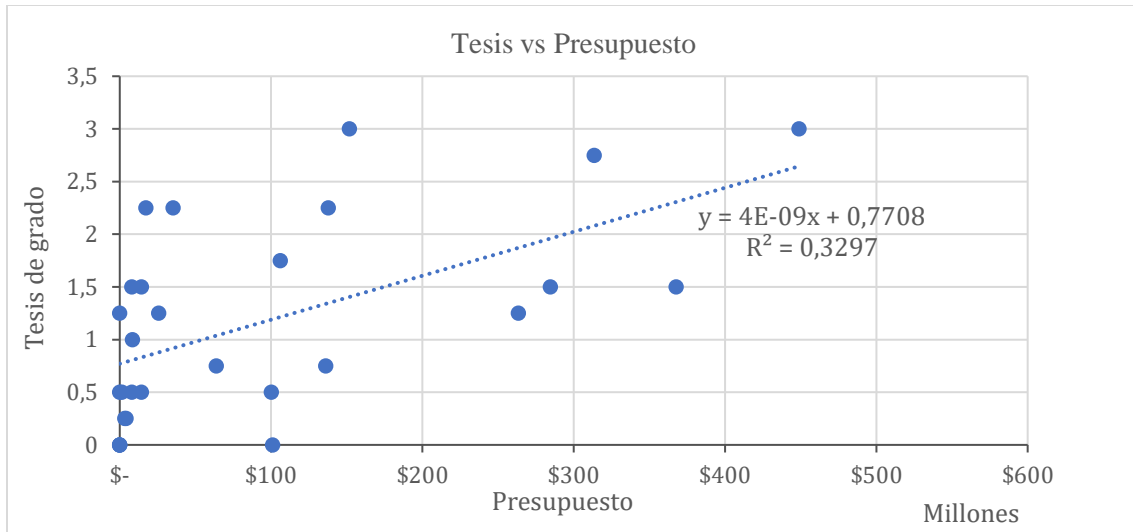


Figura 15. Dispersión Tesis vs Presupuesto.

Las Figuras 16 y 17 muestran la relación entre los productos *Artículos* y *Tesis* y la variable de entrada *Años de conformación*, luego de haber excluido los valores extremos correspondientes a grupos con una trayectoria de pocos años pero que exhiben la mayor cantidad de *Artículos* publicados, y viceversa. En ambos casos resulta evidente que los productos se relacionan moderadamente con el tiempo que llevan los grupos desde su conformación, con un coeficiente de correlación de Pearson $R = 40,3\%$ para el caso de los *Artículos* vs *Años*, y un $R = 49,5\%$ en la relación *Tesis* vs *Años*.

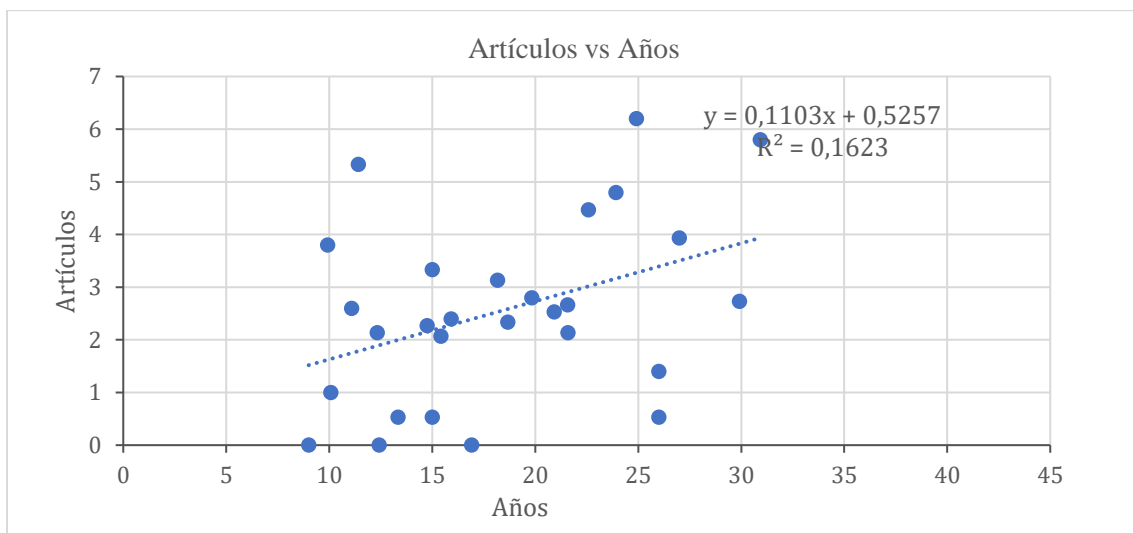


Figura 16. Dispersión Artículos vs Años de conformación.

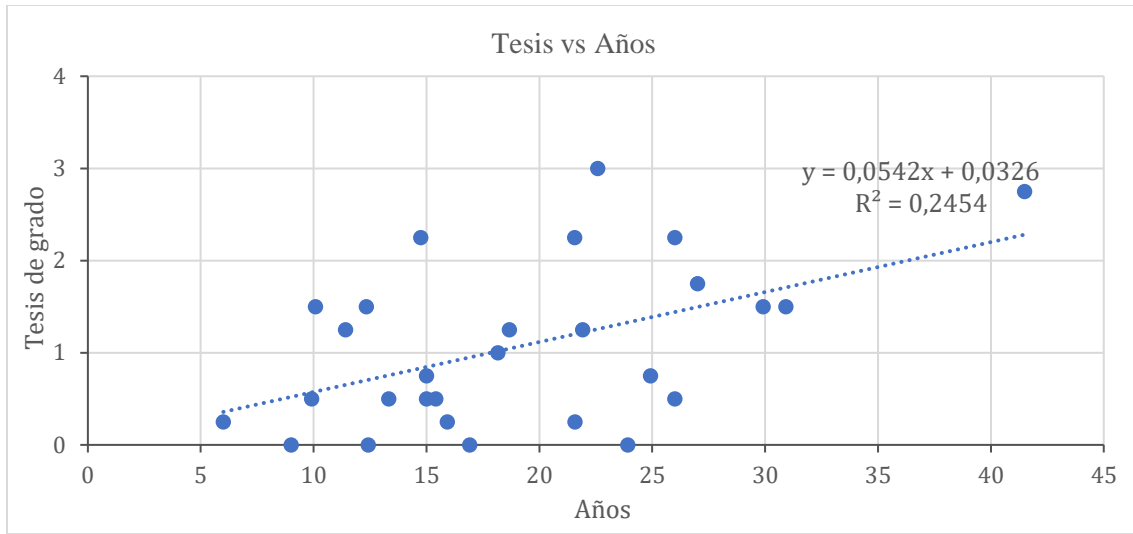


Figura 17. Dispersión Tesis vs Años de conformación.

En cuanto a la relación entre los *Estudiantes* adscritos a grupos de investigación y la producción de *Artículos*, se observa un nivel de correlación medio con un valor $R = 57,2\%$ como puede apreciarse en la Figura 18. Mientras que la relación entre las variables *Tesis* y *Estudiantes* resulta ser baja (Figura 19) aun cuando se omiten los valores extremos, obteniendo un coeficiente de correlación de Pearson $R = 11,5\%$ que se puede explicar basados en que la publicación de los trabajos de grado se efectúa cuando el autor ha dejado de ser estudiante.

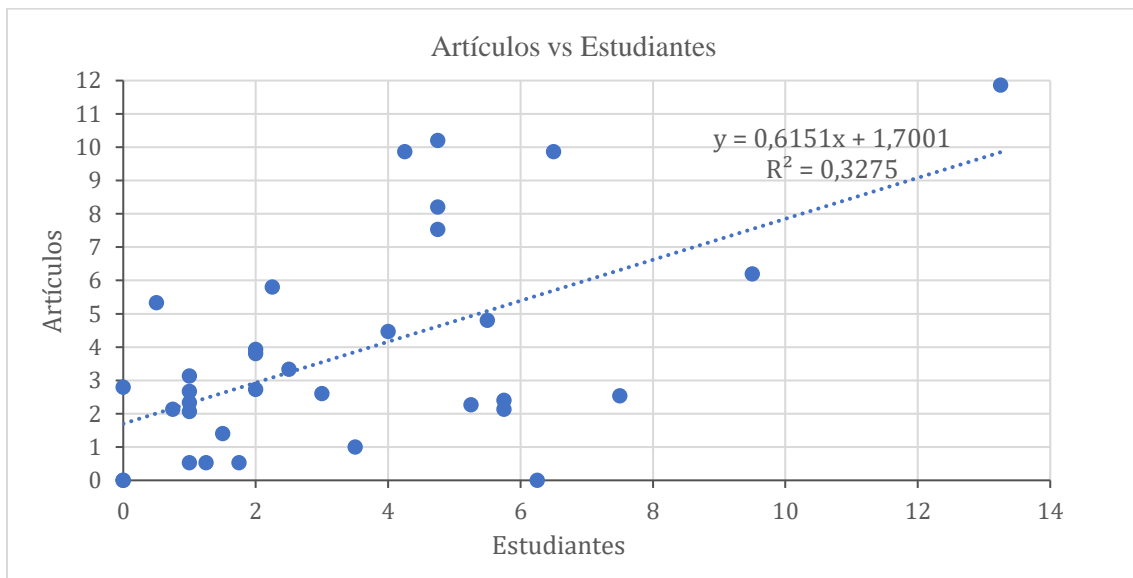


Figura 18. Dispersión Artículos vs Estudiantes.

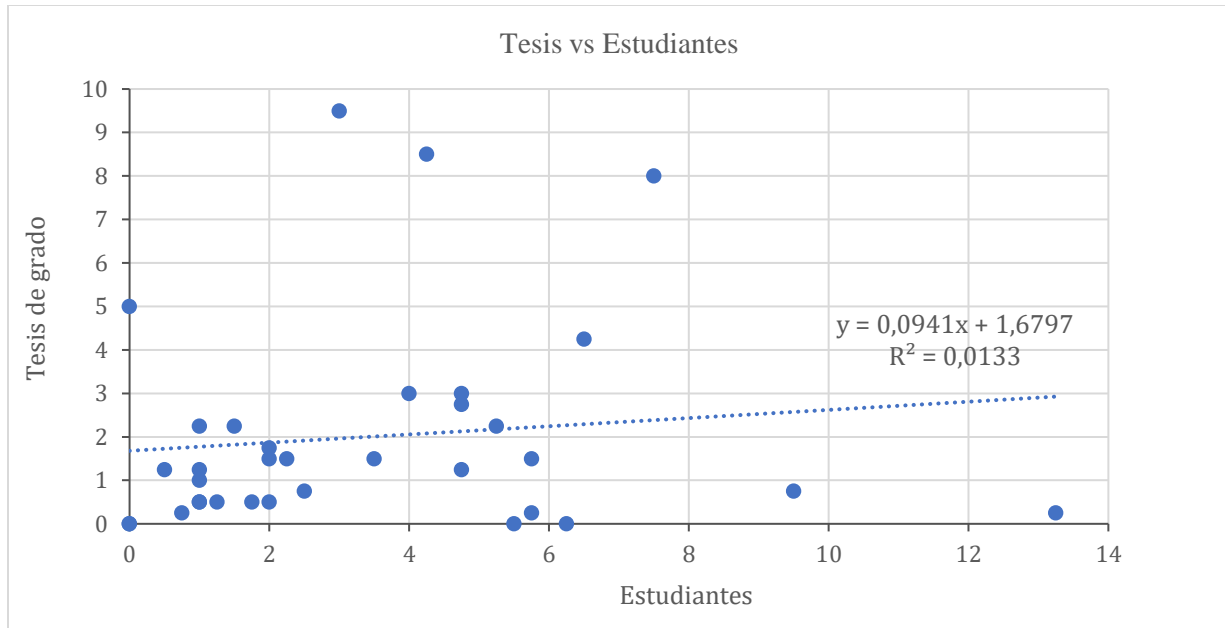


Figura 19. Dispersión Tesis vs Estudiantes.

La relación entre las variables *Artículos* y *Docentes* se muestra en la Figura 20, luego de haber excluido las observaciones extremas correspondientes a grupos con los más altos valores en este producto y pocos docentes adscritos, y viceversa. Allí se puede observar una correlación moderada entre estas variables, con un coeficiente de correlación de 62,7%.

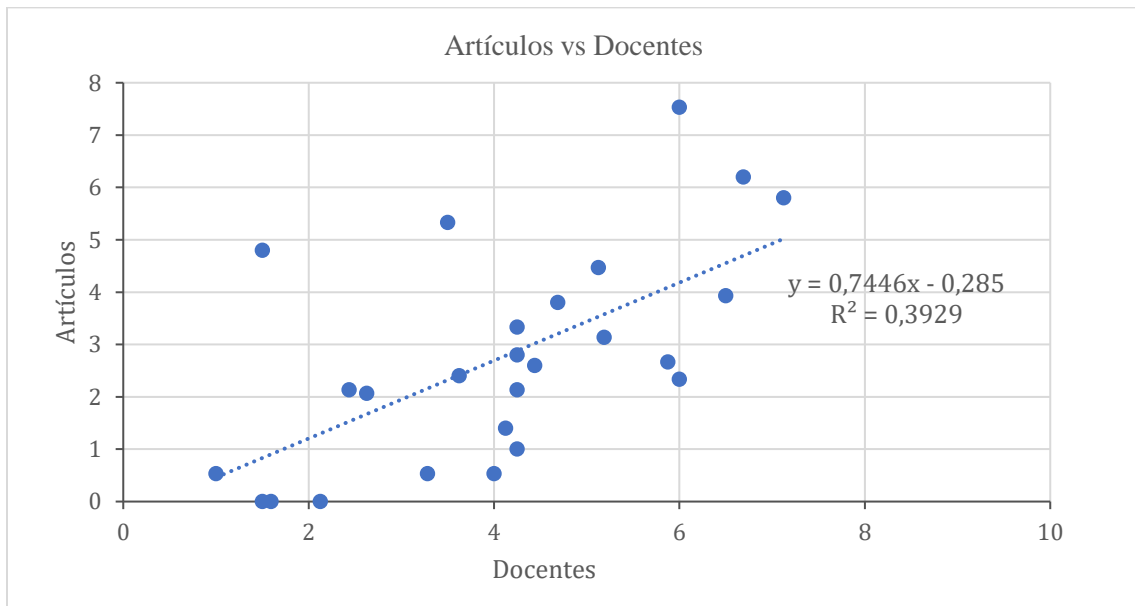


Figura 20. Dispersión Artículos vs Docentes.

Para el caso de las *Tesis* y los *Docentes*, se obtiene una relación media entre las variables en base a un coeficiente de correlación de Pearson $R = 49,7\%$ (véase Figura 21). Aun si se omitieran los valores extremos el coeficiente de correlación no cambiaría significativamente.

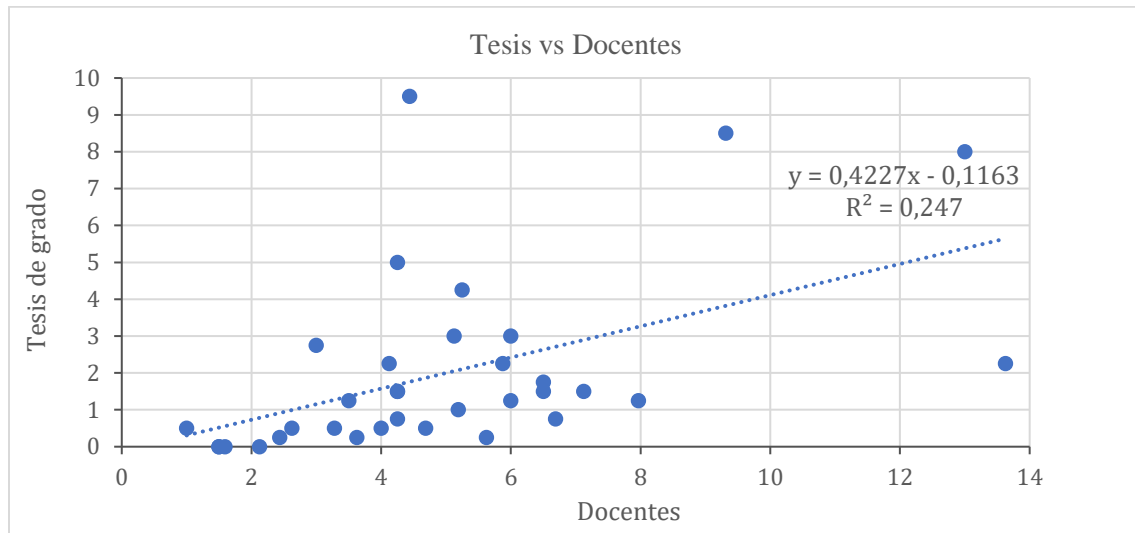


Figura 21. Dispersión Tesis vs Docentes.

8.4 Análisis Clúster de las variables

Con el propósito de conocer la semejanza de los diferentes grupos de investigación respecto a la información que se tiene en cada una de sus variables se ha decidido utilizar el Análisis Clúster, una herramienta estadística multivariante cuyo objetivo es generar conglomerados a partir de la agrupación de elementos homogéneos, de forma tal que aquellos elementos que sean similares se sitúen en un mismo grupo o clúster, mientras que aquellos que sean diferentes se ubiquen en grupos diferentes.

Dentro de los métodos de agrupación más comunes en esta técnica están el Clúster Jerárquico y el No Jerárquico. El primero genera el número de agrupaciones, las cuales tienen una estructura absorbente, es decir, se crean varios niveles jerárquicos donde los niveles inferiores son “absorbidos” por los de nivel superior. En contraste, en el No Jerárquico el investigador define, a

priori, el número de conglomerados que desea. En este caso, se optó por la aplicación del Clúster Jerárquico.

En principio, por medio de una matriz de proximidad, esta técnica identifica la distancia mínima que hay entre cada una de las entidades evaluadas, es decir, busca aquellas que están más cerca, siendo éstas las que más se parecen entre sí.

Para la ejecución de esta técnica se hizo uso del paquete estadístico *SPSS* de IBM, al cual se ingresó la información correspondiente a las variables de entrada y salida ponderadas según las categorías expuestas en el apartado anterior (Véase Apéndice C. Valores observados ponderados).

Se llevaron a cabo dos análisis separados: uno para las variables insumo (Presupuesto, Años, Docentes y Estudiantes) y otro para las variables producto (Artículos y Tesis). El Dendrograma resultante del primer análisis se muestra en la Figura 22, se trata de un gráfico en el cual aparecen los enlaces que se van formando entre los grupos según la similitud que estos presentan en los valores correspondientes a las variables insumo; es decir, en el Dendrograma se hace visible la formación de los clústeres (agrupaciones) y las distancias (eje horizontal) existentes entre unos y otros.

Así bien, en la Figura 22, se puede identificar el grupo que presenta las observaciones menos similares a las de los demás, que es HDSP (grupo No. 23) pues presenta la mayor distancia al clúster final. Claramente los últimos grupos en incorporarse al clúster final son HDSP, GISEL y GIP, lo que permite afirmar que son los que menos se parecen al resto de los grupos en términos de las variables de insumo. Por el contrario, los grupos que más se parecen son aquellos que presentan observaciones cercanas, siendo el caso de GEOMÁTICA, GMPH, INNOTECH, GPH, y los demás que le siguen en la Figura 22.

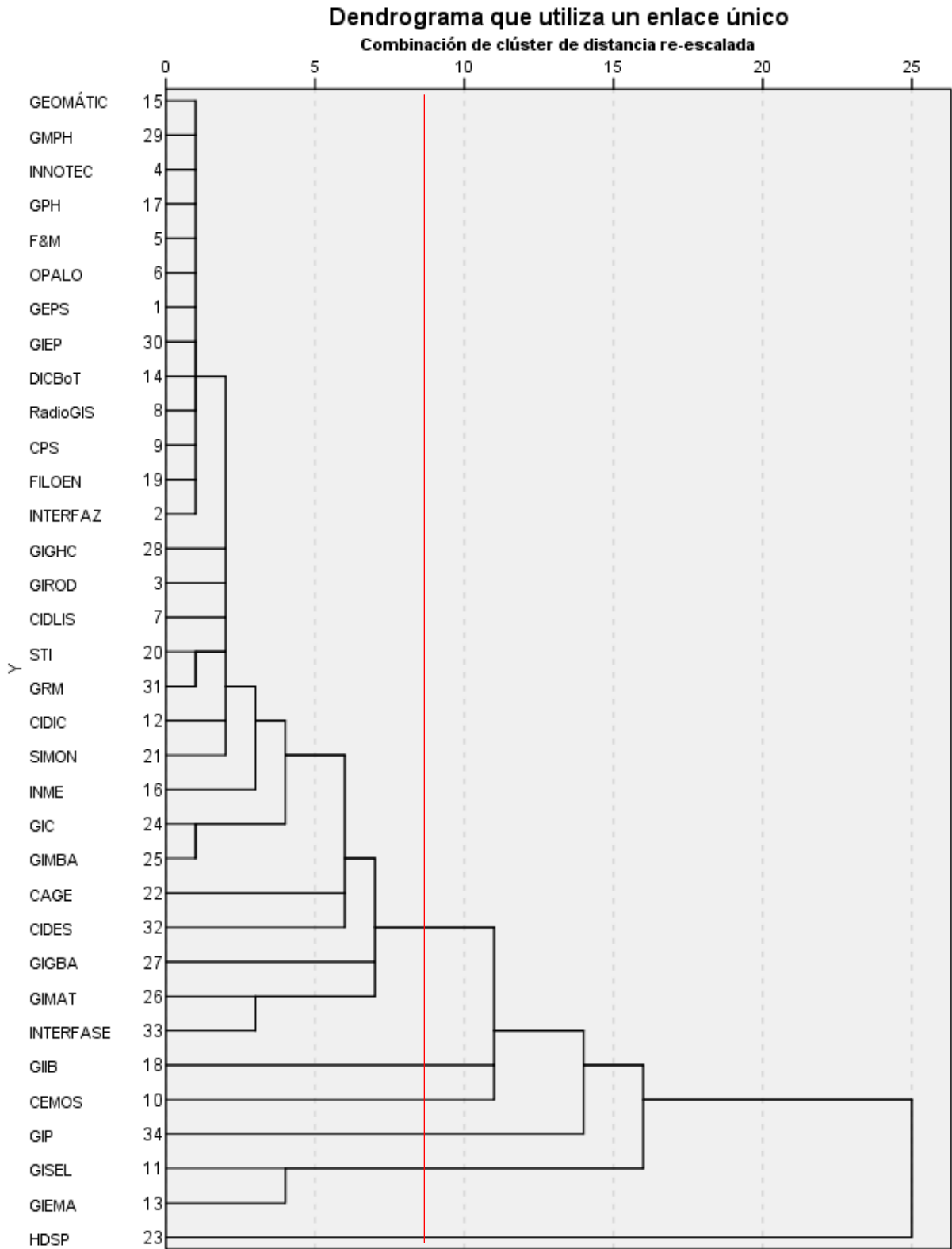


Figura 22 Dendrograma - Clúster Insumos.

Si la mayoría de los grupos presentan distancias cercanas entre sí podría decirse que los datos son homogéneos, como en este caso, ya que la mayoría de los grupos muestran una distancia inferior a 2 que el resto.

Para realizar una agrupación en seis clústeres, es necesario trazar una línea vertical que corte seis uniones en el diagrama (línea roja), de tal modo se obtiene que 28 grupos conforman el primer conglomerado, un solo grupo en los siguientes tres clústeres, dos grupos en la agrupación número 5, y uno solo en el último. En el primer clúster, donde se encuentra la mayoría de los grupos, se observa que se asemejan en las variables *Docentes* y *Estudiantes*.

En cambio, en el siguiente clúster, HDSP que presenta el máximo valor en Estudiantes, GIP con el máximo en Años y CEMOS en Presupuesto, conforman cada uno un clúster. GIIB, aunque no presenta el máximo en algún insumo, como los tres anteriores, conforma por si solo otro clúster. GISEL y GIEMA se agrupan en un clúster debido a que el comportamiento en los insumos Docentes y Estudiantes es parecido.

El Dendrograma correspondiente al análisis de las variables producto se muestra en la Figura 23. Allí se puede apreciar que el grupo cuya distancia al clúster final es la más grande es CEMOS, seguido de lejos por OPALO, identificándose como los grupos más distintos a los demás en términos de sus productos. Aun así, la mayoría de los grupos exhiben una distancia por debajo de 2, de modo que en términos de los productos también existe homogeneidad en los datos.

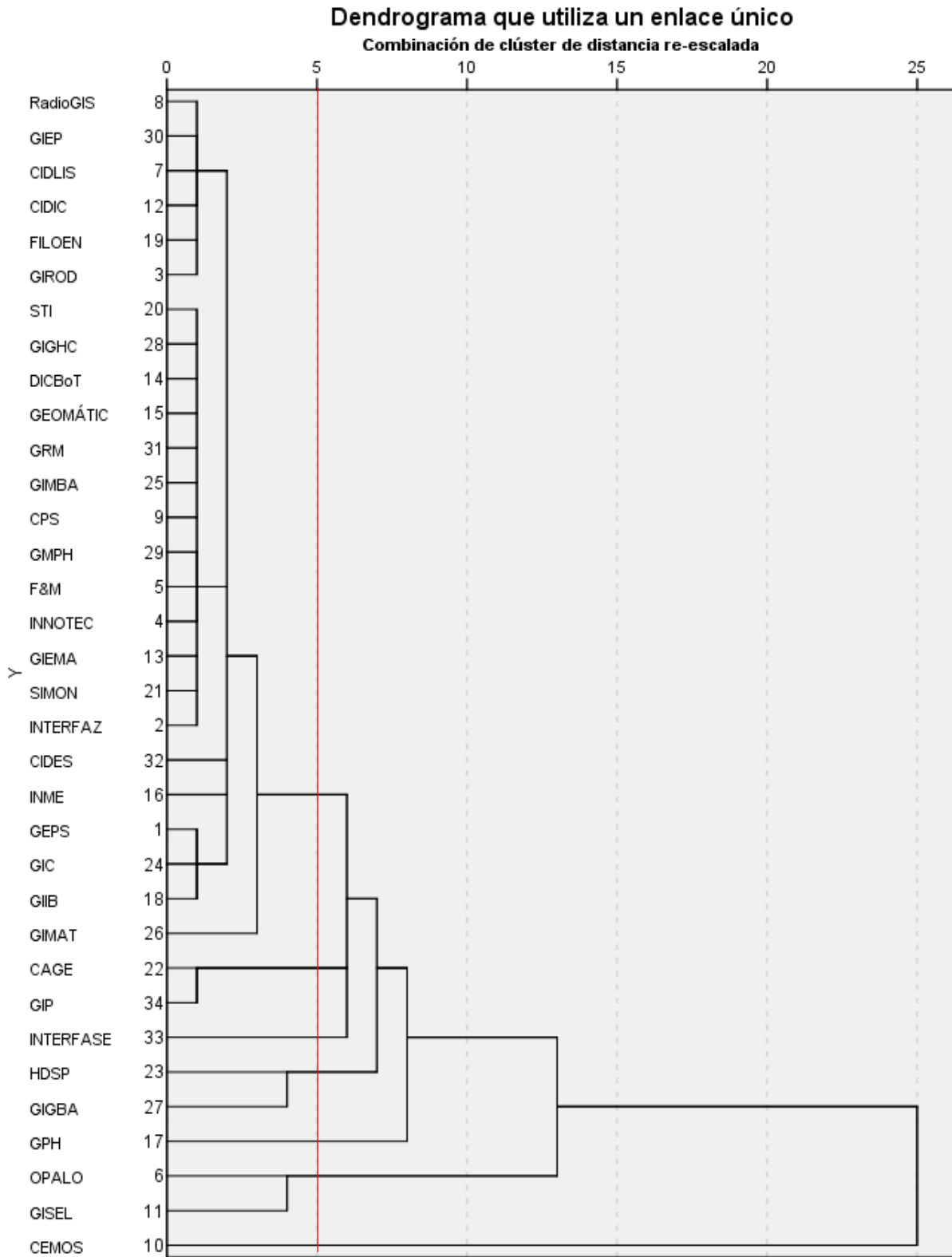


Figura 23. Dendrograma - Clúster Productos.

Se procede igual que en los insumos para establecer los grupos que van a integrar los 7 clústeres que se han decidido conformar a partir de la Figura 23. El primer conglomerado, en este caso, también está conformado por la mayoría de los grupos, 25 en total, que se parecen principalmente en el producto *Tesis de grado* y que no se diferencian significativamente en la variable *Artículos*. En el siguiente clúster se encuentran GIP y CAGE cuya similitud se aprecia en las dos variables de salida. En otra agrupación se encuentran los grupos HDSP y GIGBA, pues exhiben los dos valores más grandes en la variable *Artículos*. GISEL y OPALO se unen en una misma agrupación, pues se parecen principalmente en la producción de *Artículos* y presentan valores cercanos en *Tesis de grado*, los cuales resultan ser de los más altos. Por su parte, CEMOS conforma un clúster debido a que, a diferencia de los demás grupos, presenta grandes valores tanto en *Artículos* como en *Tesis*. Igualmente, INTERFASE y GPH conforman cada uno un clúster.

8.5 Corrida del modelo

Como se estableció en el marco teórico del presente trabajo, los modelos principales de la metodología DEA se distinguen por el tipo de rendimientos a escala que se asumen al momento de correr el programa; si se trata de retornos constantes el modelo correspondiente sería aquel propuesto por Charnes, Cooper y Rhodes, y si se trata de una tecnología de producción que presenta las características de una economía de escala el programa adecuado es el que plantean los autores Banker, Charnes y Cooper. Para conocer el tipo de tecnología que exhiben los grupos de investigación, se ha remitido a la literatura expuesta en el apartado 5. Marco de antecedentes, dentro de la cual se ha podido destacar el uso del modelo CCR bajo el supuesto de Rendimientos Constantes a Escala en todos los trabajos encontrados que utilizan DEA para medir la eficiencia de grupos de investigación. Así pues, Rojas (2010) asegura que, por las características propias de

los GI, lo más adecuado es utilizar un modelo que considere retornos constantes a escala. Por su parte, Restrepo y Villegas (2007) justifican el uso de un modelo CCR al hecho de obtener problemas de infactibilidad al momento de clasificar las DMU con el método de Súper eficiencia usando los resultados obtenidos a partir de modelos con retornos variables a escala. Giraldo (2007) después de aplicar los modelos CCR, BCC, SBM, entre otros, llega a la conclusión de que el modelo más adecuado para evaluar la eficiencia de los GI de la Universidad Tecnológica de Pereira es el CCR. Finalmente, según Aristizábal *et al.* (2017) el 59% de los investigadores usa modelos CCR, 40% modelos BCC y el restante 1% modelos de Súper eficiencia, Eficiencia Cruzada o algún modelo propio, para medir eficiencia.

Además, otros autores como Knorr, Mittermeir, Aichholzer & Wallers (1979), señalan que en el funcionamiento de las unidades académicas de investigación, intervienen factores como la motivación de sus integrantes, el prestigio de los investigadores líderes, la dedicación y cooperación, la frecuencia con que se reúnen los miembros, la inclusión de jóvenes investigadores, el desempeño como grupo, entre otros, y establecen que estos factores no se relacionan con una economía de escala.

Con base en esta información y por los resultados obtenidos corriendo los dos modelos, se optó por seleccionar el modelo CCR orientado a las salidas (CCR-O) en su forma dual identificado por las (2,9)–(3,2) y la segunda fase del mismo, dada por las ecuaciones (3,3)–(3,6).

El problema de programación lineal se resolvió utilizando el software Cplex de IBM que utiliza un lenguaje de programación C++. Para la corrida del modelo no se tuvo en cuenta los grupos CIDIC, FILOEN y GIROD debido a que su producción fue nula durante el periodo analizado en la presente investigación. Los datos ingresados al programa se pueden consultar en el Apéndice C.

Una vez corrido el modelo se obtuvo que 12 de los 31 grupos de investigación analizados, fueron calificados como 100% eficientes y los demás GI (19) con un porcentaje de eficiencia que varía entre 20,4% y 92,5%.

En la Figura 24 se muestra la agrupación de los grupos según la puntuación de eficiencia que cada uno ha obtenido. Así, dado que ninguna entidad fue calificada con un puntaje inferior al 20%, se ha definido como *Bajo nivel de eficiencia* a las puntuaciones por encima del 20% y por debajo del 45%, como *nivel de eficiencia Medio* se ha precisado que son las puntuaciones que se encuentren entre el 45% e inferiores al 70%, y como *Alto nivel de eficiencia* a aquellos puntajes mayores o iguales al 70% y por debajo del 100%, sin embargo, ya que el siguiente puntaje menor al 100% no supera el 95%, se ha establecido que este sea el límite de la última agrupación.

De este modo, se tiene que el 61,3% (19) de las entidades evaluadas obtuvieron puntajes de eficiencia por debajo del 100%, y que el 57,9% de éstas se encuentran en un nivel de eficiencia Medio. Cuatro de las 19 entidades ineficientes (21,1%) tienen puntuaciones por debajo del 45% y otras cuatro tienen puntajes por encima del 70% pero sin superar el 95%.



Figura 24. Distribución de los resultados de eficiencia de los GI.

9. Análisis de resultados

Siguiendo la teoría del Análisis Envolvente de Datos, en este caso se dirá que un grupo de investigación es CCR-Eficiente, si además de obtener una solución óptima igual a uno en la primera fase ($\eta^* = 1$), obtiene holgura cero en la segunda fase ($s^- = s^+ = 0$).

Teniendo en cuenta que en un modelo orientado a las salidas se busca la maximización de la producción manteniendo constante el nivel de insumos, los valores de holgura asociados a las salidas (s^+) serán los únicos resultantes en este modelo, pues son los que determinarán las mejoras potenciales a considerar para alcanzar la eficiencia.

La Tabla 15 presenta los valores de η^* obtenidos en la primera fase, la puntuación de eficiencia que corresponde al recíproco de η^* –dado que el modelo está orientado a las salidas–, y los valores de holgura de la segunda fase. Estos resultados muestran que justamente las DMU con Eficiencia de Farrell ($\eta^* = 1$) tienen holgura cero, por lo tanto, presentan Eficiencia de Pareto-Koopmans y se dice que son CCR-Eficientes.

Los resultados muestran 12 grupos con una puntuación de eficiencia máxima (100%) y holgura cero; 6 de los 19 grupos con puntajes inferiores al 100% obtuvieron holguras en alguno de sus productos (Artículos o Tesis), y los 13 restantes a pesar de haber sido calificados como ineficientes sus holguras resultaron ser cero, evidenciando que su nivel de ineficiencia es más bajo y por consiguiente está más próximo a lograr un nivel fuerte de eficiencia en comparación con aquellos que sí tienen holguras.

Es importante recordar que la medida de Eficiencia CCR involucra dos tipos de eficiencia: Técnica (ETP) y de Escala (EE), así que cuando una entidad obtiene un valor de eficiencia del

100% en el modelo CCR, se dice que no posee ningún tipo de ineficiencia, y por tanto genera la máxima productividad.

Tabla 15
Resultados del modelo DEA CCR-O

Grupo de investigación	CCR		Déficit de salidas	
	η^*	Puntuación de eficiencia	S^+ _{artículos}	S^+ _{tesis}
CAGE	1	100%	0	0
CEMOS	1	100%	0	0
CIDES	1	100%	0	0
GEOMÁTICA	1	100%	0	0
GEPS	1	100%	0	0
GIGHC	1	100%	0	0
GIP	1	100%	0	0
GPH	1	100%	0	0
HDSP	1	100%	0	0
INTERFASE	1	100%	0	0
OPALO	1	100%	0	0
STI	1	100%	0	0
GMPH	1,0817	92,45%	0	0,1124
GIGBA	1,2745	78,46%	0	0,7950
GISEL	1,2889	77,59%	0,4659	0
CIDLIS	1,4129	70,77%	0	0
GRM	1,6056	62,28%	0	0
DICBoT	1,6598	60,25%	0	0
RadioGIS	1,6667	60,00%	0,6667	0
INTERFAZ	1,7447	57,32%	0	0
GIMAT	1,7506	57,12%	0	0
F&M	1,7592	56,84%	0	1,0384
GIIB	1,8172	55,03%	0	0,6532
GIC	2,0307	49,24%	0	0
CPS	2,0392	49,04%	0	0
SIMON	2,1719	46,04%	0	0
INNOTEC	2,2194	45,06%	0	0
INME	2,4156	41,40%	0	0
GIMBA	3,2690	30,59%	0	0
GIEMA	3,3760	29,62%	0	0
GIEP	4,9098	20,37%	0	0

Además de las Holguras, cada DMU ineficiente tendrá un conjunto de referencia conformado por aquellas entidades eficientes mejores que ésta en términos de su producción; razón por la cual

la entidad en cuestión ha resultado ineficiente: existe al menos una DMU –o una combinación lineal de entidades– que obtiene una mayor producción utilizando, como mucho, las cantidades de factores productivos de la DMU observada. Por ejemplo, podría decirse que para que el grupo de investigación GISEL sea eficiente, debe guiarse de los grupos OPALO y GPH; y que INNOTEC por su parte deberá tomar como referencia a GEPS, OPALO y GPH para alcanzar la eficiencia (véase Tabla 16).

Tabla 16
Conjunto de referencia para los GI no eficientes

Grupo de investigación	Conjunto de referencia				
RadioGIS	GEOMÁTICA				
CIDLIS	GIGHC	GPH			
GISEL	GPH	OPALO			
GRM	GEPS	GPH			
DICBoT	GEPS	GIGHC	GPH		
F&M	CAGE	GEPS	HDSP		
GIC	GEPS	GIP	INTERFASE		
GIEMA	CAGE	GEPS	OPALO		
GIGBA	CAGE	GEPS	HDSP		
GIIB	CIDES	GEPS	HDSP		
INNOTEC	GEPS	GPH	OPALO		
INTERFAZ	GEPS	GPH	OPALO		
SIMON	GEPS	GPH	OPALO		
CPS	CIDES	GEPS	HDSP	INTERFASE	
GIEP	GEOMÁTICA	GEPS	GPH	STI	
GIMAT	GEPS	GIP	INTERFASE	OPALO	
GIMBA	GEPS	GIP	GPH	OPALO	
GMPH	GEOMÁTICA	GEPS	GIGHC	STI	
INME	GEPS	GIP	GPH	OPALO	

Como es evidente, un grupo de investigación eficiente puede ser referente de varios grupos que no alcanzaron la eficiencia. Dicha presencia reiterada en los conjuntos de referencia indica que una entidad posee una mayor solidez en su eficiencia. En este caso GEPS, GPH y OPALO son aquellos que presentan la mayor frecuencia, véase Figura 25.

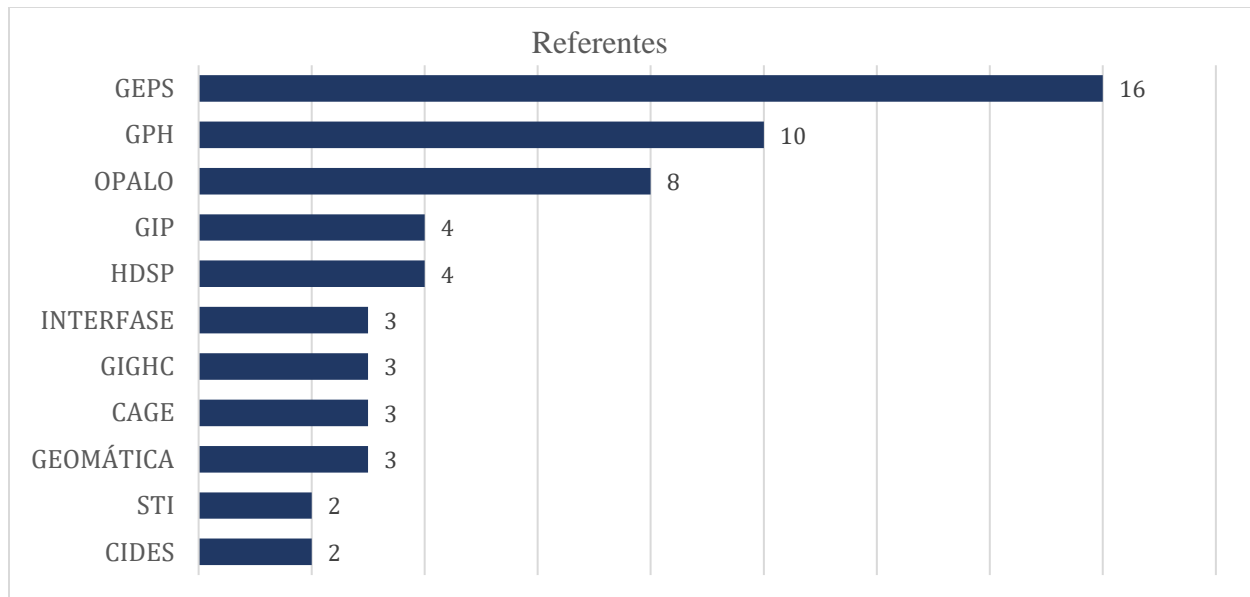


Figura 25. Frecuencia en conjuntos de referencia.

9.1 Valores objetivo y mejoras potenciales

Como se estableció en el apartado teórico de este trabajo, las entidades que resultan eficientes conforman lo que se denomina la *frontera de eficiencia* que envuelve a las demás DMU que no la alcanzaron. Sobre dicha frontera se localiza la versión mejorada $(\hat{x}_{ij}, \hat{y}_{ij})$ de cada entidad calificada como ineficiente –proyección que se determina a partir de la combinación lineal de todos sus referentes–, y cuyas coordenadas representan las cantidades de insumo-producto que la ineficiente debe lograr para ser eficiente (valores objetivo).

En los resultados de la segunda fase del modelo CCR-O se conocen las holguras y el conjunto de referencia de cada GI calificado como ineficiente. Los referentes se identifican cuando el valor obtenido de los pesos λ_j es mayor que cero en la solución óptima, y representan la contribución insumo-producto que los GI eficientes hacen a los valores objetivo de las entidades ineficientes, marcando de este modo, la trayectoria que estas últimas deben seguir para alcanzar la mejora más accesible.

En el caso del grupo GISEL, cuyos referentes OPALO y GPH tienen pesos de 0,75114 y 0,63504 respectivamente, los valores objetivo del producto *Artículos* se calculan de la siguiente

$$\text{manera: } \hat{y}_{rj} = \sum_{j=1}^n \lambda_j y_{rj} \quad \forall r$$

$$\hat{y}_{\text{artículos}, GISEL} = (\lambda_{OPALO} y_{\text{artículos}, OPALO}) + (\lambda_{GPH} y_{\text{artículos}, GPH})$$

$$\hat{y}_{\text{artículos}, GISEL} = (0,75114 * 2,6) + (0,63504 * 2,8)$$

$$\hat{y}_{\text{artículos}, GISEL} = 3,7311$$

De forma similar se obtienen las proyecciones para la salida *Trabajos de grado*, y para las demás entidades ineficientes. En el Apéndice D se pueden consultar los valores de λ_j para cada grupo de investigación.

Si bien los referentes y sus pesos por cada entidad ineficiente dictan las coordenadas de su proyección sobre la frontera, los valores de holgura sugieren las reducciones e incrementos (en entradas y salidas respectivamente) que las DMU ineficientes deberían efectuar para llegar a ser eficientes. Indican, en el caso de un modelo orientado a las salidas, que con los niveles de insumos que presenta actualmente, la entidad podría ser capaz de generar una cantidad de productos adicional s^+ .

De este modo, las holguras permiten obtener de manera más simple el valor de cada salida r , correspondiente a la proyección sobre la frontera de cada entidad ineficiente j , así:

$$\hat{y}_{rj} = \eta^* y_{rj} + S_r^+ \quad \forall r$$

Esto significa que la entidad ineficiente, utilizando sus cantidades actuales x_{ij} de insumos, deberá obtener al menos una proporción η^* de sus productos.

Siguiendo con el ejemplo del grupo GISEL, que presenta un valor observado de 2,5333 en el producto *Artículos*, una solución óptima en la primer fase equivalente a $\eta^* = 1,2889$ y un valor de holgura en la segunda fase correspondiente a $S_{\text{artículos}, GISEL}^+ = 0,46592$ que se puede consultar en la Tabla 15, se tiene que:

$$\hat{Y}_{\text{artículos, GISEL}} = (1,2889 * 2,5333) + 0,46592$$

$$\hat{Y}_{\text{artículos, GISEL}} = 3,7311$$

Resulta evidente que el valor de la proyección, obtenido a partir de las holguras, es el mismo que se obtiene a partir de los pesos λ_j asignados a los referentes.

Se procede de la misma forma a calcular los valores objetivo para cada grupo de investigación que haya obtenido una puntuación de eficiencia menor al 100% y se presentan en la Tabla 17. Es importante aclarar que debido a que los datos ingresados al modelo se ponderaron para la mayoría de las variables, los valores de holgura, y cualquier cálculo que se haga a partir de éstos, también estarán expresados en términos ponderados, según las unidades que se establecieron en el apartado 8.2 Tratamiento de la información, lo que da cierta flexibilidad a la hora de querer mejorar su eficiencia.

Tabla 17
Valores objetivo.

Grupo de investigación	Valores objetivo		Grupo de investigación	Valores objetivo	
	Artículos	Tesis		Artículos	Tesis
CIDLIS	0,7536	0,7065	GIMBA	8,9352	4,9035
CPS	6,7974	1,5294	GISEL	3,7311	10,3111
DICBoT	3,4303	0,8299	GMPH	3,3892	1,1941
F&M	6,6849	1,9180	GRM	3,4252	2,4083
GIC	11,7783	3,0461	INME	9,5012	4,2272
GIEMA	7,6523	7,5960	INNOTECH	5,9185	4,9937
GIEP	2,6186	2,4549	INTERFAZ	1,7447	2,6170
GIGBA	13,0004	2,3882	RadioGIS	1,5556	0,8333
GIIB	11,2669	2,0161	SIMON	3,0407	4,8868
GIMAT	7,8192	5,2517			

Ahora bien, según los valores objetivo es posible identificar el tipo de mejoras que un GI ineficiente tendría que experimentar para ubicarse en su proyección sobre la frontera. En este caso las mejoras sólo podrán estar enfocadas a los productos, así que para toda entidad ineficiente que presente valores de holgura $S_r^+ = 0$ se dirá que es necesario realizar mejoras radiales

correspondientes al incremento proporcional de sus productos: $(\eta^* - 1)y_{rj}$; pero si cuenta con holguras $S_r^+ > 0$ la entidad deberá realizar, además de las mejoras radiales, mejoras de holgura debidas a los incrementos de este tipo.

En la Tabla 18 se muestran las mejoras potenciales en que deberían incurrir los grupos ineficientes, para alcanzar la eficiencia. Nótese que las entidades con menores puntajes de eficiencia son aquellas que presentan el mayor porcentaje de mejora (GIEMA y GIEP), además, para aquellas entidades que sólo deben realizar mejoras radiales, ese porcentaje es equivalente al valor que le falta para obtener una puntuación del 100%.

Tabla 18
Mejoras potenciales

Grupo de investigación	Mejoras potenciales		Porcentaje de mejora	
	Artículos	Tesis	Artículos	Tesis
CIDLIS	0,22	0,21	29,2%	29,2%
CPS	3,46	0,78	51,0%	51,0%
DICBoT	1,36	0,33	39,8%	39,8%
F&M	2,88	1,42	43,2%	73,9%
GIC	5,98	1,55	50,8%	50,8%
GIEMA	5,39	5,35	70,4%	70,4%
GIEP	2,09	1,95	79,6%	79,6%
GIGBA	2,80	1,14	21,5%	47,7%
GIIB	5,07	1,27	45,0%	62,8%
GIMAT	3,35	2,25	42,9%	42,9%
GIMBA	6,20	3,40	69,4%	69,4%
GISEL	1,20	2,31	32,1%	22,4%
GMPH	0,26	0,19	7,5%	16,3%
GRM	1,29	0,91	37,7%	37,7%
INME	5,57	2,48	58,6%	58,6%
INNOTEC	3,25	2,74	54,9%	54,9%
INTERFAZ	0,74	1,12	42,7%	42,7%
RadioGIS	1,02	0,33	65,7%	40,0%
SIMON	1,64	2,64	54,0%	54,0%

Como se ha mencionado anteriormente, todos los valores en las Tablas 17 y 18 aluden a datos ponderados, es decir, si en la columna correspondiente a *Artículos* de la Tabla de *Mejoras potenciales* se halla un valor de 1,73 éste no quiere decir que se deba producir 1,73 artículos tipo

A1 adicionales, sino que el grupo debe obtener un puntaje adicional de 1,73 en ese producto, donde dos artículos A1 serían suficientes, o en su defecto, dos A2 y un C, etcétera.

En ese orden de ideas, se puede afirmar que el grupo de investigación CIDLIS habría alcanzado la frontera eficiente si hubiera tenido un puntaje adicional de 0,22 en el producto *Artículos*, y de 0,21 en *Trabajos de grado* (mejoras radiales). Sin embargo, enunciar las mejoras de esta manera no parece muy práctico, ya que no le indica a cada GI cómo debe incrementar su producción. Para solventar ese inconveniente, debe entenderse que los incrementos apuntan a un mínimo que el producto debe alcanzar.

Entonces, para el caso de los productos, lo que se quiere conocer es cuántos *Artículos* como mínimo debe producir adicionalmente el GI y en qué tipo de revista deben estar indexados cada uno de ellos (A1, A2, B o C) para incrementar su puntuación al menos en el valor que se muestra en la Tabla 18. De igual forma para los *Trabajos de grado*, lo que se busca es el mínimo de éstos que adicionalmente debe dirigir el grupo, según el nivel académico en que éstos se encuentren (Pregrado, Maestría o Doctorado), para alcanzar como mínimo el valor de mejora que se muestra en la Tabla 18. Para que la entidad pueda ser calificada como eficiente, deberá realizar el incremento correspondiente en cada una de las salidas de forma simultánea.

Así, el grupo de investigación CIDLIS, con el nivel de insumos que presentaba en el año 2016¹⁷, hubiera podido producir un Artículo tipo B adicional, y dirigir una Tesis de pregrado adicional. Entonces, CIDLIS sería calificado como un grupo de investigación eficiente si sus productos fueran al menos 2 Artículos indexados en revistas tipo B, y 3 Tesis de pregrado dirigidas.

Se procede de la misma manera para los demás grupos y se obtiene que GIEMA, manteniendo constante sus insumos, habría podido producir adicionalmente 6 Artículos (3 indexados en A1 y 3

¹⁷ Estos valores se pueden consultar en el Apéndice C

indexados en A2), una Tesis de grado de maestría y 5 Tesis doctorales –mejoras radiales en ambos productos. De este modo, habría alcanzado la frontera eficiente.

Un grupo al que le corresponde realizar tanto mejoras radiales como proporcionales es GIGBA, pues presenta valores de holgura en el producto Tesis de grado, de modo que debe dirigir adicionalmente al menos un Trabajo de grado de pregrado y una tesis doctoral. Y en Artículos una mejora radial equivalente a producir como mínimo tres A1.

GIEP, por su parte, no presenta valores de holguras en los productos, lo que indica que sus mejoras sólo serán radiales. Así que este grupo habría alcanzado la eficiencia si, manteniendo constante sus insumos, adicionalmente hubiera producido 3 Artículos –indexando dos de ellos en revistas A2, y uno en alguna revista tipo B– y dirigido 2 Tesis doctorales.

En cuanto a los incrementos en la producción, el grupo de investigación que tendría que producir una mayor cantidad de artículos es GIMBA, seguido de GIC e INME. Mientras que en la dirección de tesis GIEMA es el que más aumentos debe realizar.

Los incrementos en cada categoría de los productos, para los GI que requieren mejoras, se encuentran en las Tablas 19 y 20.

Cabe aclarar que las mejoras que se proponen en esas tablas señalan valores mínimos adicionales, y por consiguiente es posible obtener los puntajes de mejora, que se muestran en la Tabla 18, con otras combinaciones de incrementos. Los valores objetivo en términos de las categorías de cada producto se pueden consultar en el Apéndice E.

Tabla 19

Mejora potencial: Incremento en la producción de artículos

Grupos de Investigación	Producción adicional de Artículos			
	A1	A2	B	C
CIDLIS			1	
CPS	3		1	
DICBoT	1		1	
F&M	3			
GIC	6			
GIEMA	3	3		
GIEP		2	1	
GIGBA	3			
GIIB	3	2	1	
GIMAT	1	3		
GIMBA	5	1	1	
GISEL	1			1
GMPH			1	
GRM		1	1	
INME	4	2		
INNOTEC	2	1	1	
INTERFAZ		1		
RadioGIS			2	
SIMON	1	1		

Tabla 20

Mejora potencial: Incremento en la dirección de Trabajos de grado

Grupos de Investigación	Dirección adicional de Tesis		
	Pregrado	Maestría	Doctorado
CIDLIS	1		
CPS			1
DICBoT		1	
F&M		1	1
GIC			2
GIEMA		1	5
GIEP			2
GIGBA	1		1
GIIB		1	1
GIMAT		1	2
GIMBA		1	3
GISEL		1	2
GMPH	1		
GRM			1
INME		1	2
INNOTEC			3
INTERFAZ	1		1
RadioGIS		1	
SIMON			3

Tenga en cuenta para lo anterior que 1) las entidades o DMU son grupos de investigación y, por consiguiente, no es apropiado sugerir una reducción de sus insumos, razón por la cual el modelo escogido se ha orientado a la maximización de las salidas; 2) los insumos que hacen referencia a integrantes de los GI, y los productos, cuentan con categorías que fueron ponderadas¹⁸ para dar lugar a un único valor total en cada una de esas variables; y 3) el periodo analizado corresponde a un año (2016) debido a la no disponibilidad de información completa de años anteriores.

9.2 Clasificación de los Grupos de Investigación

A sabiendas de que en un modelo CCR orientado a las salidas, en el que el valor de eficiencia corresponde al recíproco de η^* obtenido en la solución óptima del modelo, la máxima puntuación que una entidad puede alcanzar es el 100%, se hace necesario remitirse a modelos alternativos que permitan el “desempate” de las DMU eficientes.

Dicho “desempate” tiene como objeto identificar cuáles son los mejores dentro del grupo de los mejores, a fin de asignarles una ubicación apropiada en el ranking de las entidades evaluadas.

9.2.1 Clasificación según Súper Eficiencia. El aspecto teórico de este modelo se detalló en el apartado 6.5.2, sin embargo, para facilitar la observación de éste, se presenta a continuación el programa matemático utilizado para efectuar su aplicación.

$$\max_{\eta, \lambda} \eta \quad (4,2)$$

$$\text{sujeto a } x_{io} \geq \sum_{j=1, j \neq o}^n \lambda_j x_{ij} \quad i = 1, 2, \dots, m \quad (4,3)$$

$$\sum_{j=1, j \neq o}^n \lambda_j y_{rj} \geq \eta y_{ro} \quad r = 1, 2, \dots, s \quad (4,4)$$

$$\lambda_1, \dots, \lambda_n \geq 0 \quad \eta \text{ libre} \quad (4,5)$$

¹⁸ Véase apartado 8.2 Tratamiento de la información.

El desempate en este caso sólo se da entre las entidades eficiente, dado que las ineficientes al haber obtenido puntuaciones de eficiencia por debajo del 100% difícilmente una entidad obtendrá un valor idéntico al de otra, así que éstas pueden ordenarse fácilmente de menor a mayor ineficiencia.

Como puede notarse, el modelo de Súper Eficiencia es casi exactamente el mismo modelo Dual CCR-O que se muestra en las ecuaciones (2,9)–(3,2) del apartado 6.3.5, salvo por un pequeño detalle: al evaluar la entidad observada no se tiene en cuenta a ella misma ($j \neq o$).

Debido a que la orientación corresponde a las salidas, la puntuación de eficiencia también se obtiene del recíproco del η^* que este modelo arroja en la solución óptima, de modo que un menor valor de η^* apunta a un mayor valor de eficiencia.

Las entidades señaladas como ineficientes por el modelo CCR-O, obtendrán la misma solución óptima en el modelo de Súper Eficiencia.

En la Tabla 21 se muestra la clasificación de los grupos de investigación según los resultados del modelo de Súper Eficiencia. Note que el grupo GPH fue el que obtuvo la solución óptima más pequeña ($\eta^* = 0,1$), lo que indica que es el GI más eficiente.

Los grupos que se encuentran en las primeras posiciones de este ranking, los que recibieron la mayor puntuación de eficiencia, son en orden: GPH, GEOMÁTICA, CAGE, OPALO y GEPS.

En el caso de Súper Eficiencia se observa que las dos primeras posiciones de la clasificación las ocupan los grupos GPH y GEOMÁTICA, esto podría explicarse con la relación que hay entre las variables de entrada y de salida. En GPH, a diferencia de los demás grupos, la variable *Estudiantes* es nula, además sus otros insumos presentan valores por debajo de la media (exceptuando la variable *Años*), mientras que en productos la variable *Tesis* se encuentra muy por

encima de esta. También se observa que 8 de los 10 grupos para los cuales GPH es referente no lo superan en producción aun cuando sus niveles de insumo son cercanos al de éste.

Tabla 21
Resultado según método de Súper Eficiencia

Clasificación	Grupo de investigación	Súper eficiencia	
		η^*	Puntuación de eficiencia
1	GPH	0,1	1000%
2	GEOMÁTICA	0,1306	765,62%
3	CAGE	0,3705	269,90%
4	OPALO	0,4124	242,46%
5	GEPS	0,4471	223,64%
6	HDSP	0,6537	152,97%
7	CIDES	0,7096	140,93%
8	GIGHC	0,7474	133,80%
9	GIP	0,7564	132,21%
10	INTERFASE	0,8931	111,96%
11	CEMOS	0,9041	110,61%
12	STI	0,9499	105,28%
13	GMPH	1,0817	92,45%
14	GIGBA	1,2745	78,46%
15	GISEL	1,2889	77,59%
16	CIDLIS	1,4129	70,77%
17	GRM	1,6056	62,28%
18	DICBoT	1,6598	60,25%
19	RadioGIS	1,6667	60,00%
20	INTERFAZ	1,7447	57,32%
21	GIMAT	1,7506	57,12%
22	F&M	1,7592	56,84%
23	GIIB	1,8172	55,03%
24	GIC	2,0307	49,24%
25	CPS	2,0392	49,04%
26	SIMON	2,1719	46,04%
27	INNOTEC	2,2194	45,06%
28	INME	2,4156	41,40%
29	GIMBA	3,2690	30,59%
30	GIEMA	3,3760	29,62%
31	GIEP	4,9098	20,37%

Por otro lado, de todos los grupos de investigación solo dos tienen un valor de cero en la variable *Presupuesto*, GEOMÁTICA y RadioGIS, sin embargo, este último resultó ser ineficiente puesto que su producción no fue muy significativa (un Artículo B y dos Tesis de pregrado), en comparación con GEOMÁTICA que produjo 3 artículos (A1, A2 y B) y cinco Tesis de pregrado dirigidas; razón por la cual GEMOMÁTICA es el único referente de RadioGIS. Así que, al existir una fuerte relación entre el *Presupuesto* y los *Artículos* (como se estableció en el Apartado 8.3) es posible justificar que GEOMÁTICA se catapulte a la segunda posición de este ranking. Además, es referente de otros dos grupos (GMPH y GIEP) cuya similitud se ve reflejada principalmente en la variable *Estudiantes*, y no se evidencia una diferencia muy marcada en los demás insumos, pero sí en sus productos.

9.2.2 Clasificación según Eficiencia Cruzada. Como se mencionó en el apartado 6.5.2 del marco teórico, esta metodología utiliza las ponderaciones óptimas asignadas por el DEA en el modelo Primal que se indica en las ecuaciones (2,4)–(2,8).

Para obtener los valores de Eficiencia Cruzada (*CE* por sus siglas en inglés) lo primero que se requiere es realizar el cálculo de la puntuación de eficiencia (*E*) de cada DMU utilizando las ponderaciones óptimas v^* y μ^* , obtenidas en la solución del modelo CCR-O Primal, de las demás entidades, así:

$$E_{oj} = \frac{\sum_{r=1}^s \mu_{ro}^* y_{rj}}{\sum_{i=1}^m v_{io}^* x_{ij}}$$

De este modo, la entidad observada o va a generar n puntuaciones de eficiencia para cada entidad j , siendo una de ellas la correspondiente a sus propios pesos óptimos, cuando $o = j$.

Como puede verse en la Tabla 22, al realizar este mismo procedimiento para cada entidad, se genera una matriz $n \times n$ en la cual las filas muestran la calificación de eficiencia al utilizar las ponderaciones de la DMU_o con las entradas x_{ij} y las salidas y_{rj} de las demás entidades, mientras

que en las columnas se encuentran las puntuaciones de cada DMU_j con los valores v^* y μ^* de las demás; y cuya diagonal refleja la puntuación de eficiencia obtenida directamente de correr el modelo CCR-O Primal o Dual.

Tabla 22
Matriz Eficiencia Cruzada

E_{oj}		DMU j									
		1	2	3	4	5	6	7	8	9	10
DMU observada	1	1	0,246	0,281	0,473	0,292	0,055	0,096	0,401	0,314	0,220
	2	1	0,573	0,451	0,355	1	0,104	0,182	0,369	0,360	0,753
	3	1	0,573	0,451	0,355	1	0,104	0,182	0,369	0,360	0,753
	4	1	0,142	0,280	0,568	0,362	0,081	0,076	0,440	0,779	0,140
	5	0,136	0,187	0,130	0,059	1	0,024	0,042	0,061	0,513	0,468
	6	1	0,323	0,371	0,257	0,284	0,708	0,216	0,329	0,154	0,213
	7	0,069	0,238	0,090	0,007	0,126	0,434	0,600	0,017	0,021	0,131
	8	1	0,263	0,377	0,505	0,962	0,223	0,117	0,490	0,777	0,330
	9	1	0,172	0,338	0,518	0,758	0,058	0,082	0,340	1	0,306
	10	0,293	0,526	0,311	0,058	1	0,081	0,143	0,100	0,210	0,776

Nota: La matriz completa se encuentra en el Apéndice F.

Dado que se están evaluando $n = 31$ entidades sólo se muestra, en la Tabla 22, una parte de la matriz resultante al efectuar el cálculo de E_{oj} . La matriz completa puede consultarse en el Apéndice F, donde además se especifica a cuál entidad hace referencia la numeración.

Entonces, si se observa la columna 1 puede notarse que la entidad 1 (DMU_1) es eficiente, y además obtiene la máxima puntuación de eficiencia al utilizar los pesos óptimos de otras 19 entidades¹⁹, lo que quiere decir que éstas exhiben actividades productivas similares.

Finalmente, para obtener los valores de Eficiencia Cruzada (CE) que servirán para clasificar las entidades, se calcula el promedio por columna de las calificaciones de eficiencia E_{oj} de la Tabla 22, sin tener en cuenta cuando $o = j$, razón por la cual el divisor corresponde a $n - 1$, así:

¹⁹ Ya que la Tabla 22 es un fragmento de la matriz original sólo puede apreciarse que la DMU_1 obtiene la máxima puntuación de eficiencia con las entidades 2, 3, 4, 6, 8, y 9. Sin embargo, si se consulta el Apéndice F será posible comprobar las demás entidades con las que esto sucede.

$$CE_j = \frac{\sum_{o=1, o \neq j}^n E_{oj}}{n-1}$$

La Tabla 23 resume los valores de Eficiencia Cruzada, ordenando las entidades de mayor a menor valor CE y asignando una posición de clasificación.

Tabla 23
Resultados según método de Eficiencia Cruzada

Eficiencia Cruzada		
Clasificación	Grupo de investigación	CE
1	GEPS	0,7798
2	GPH	0,7367
3	CAGE	0,6232
4	OPALO	0,6168
5	INTERFASE	0,4688
6	GIP	0,4562
7	CEMOS	0,4523
8	GIGHC	0,4254
9	GEOMÁTICA	0,4253
10	GMPH	0,4225
11	CIDES	0,4120
12	HDSP	0,3992
13	GIGBA	0,3946
14	STI	0,3721
15	DICBoT	0,3489
16	GRM	0,3449
17	INNOTEC	0,3070
18	F&M	0,3058
19	CPS	0,3025
20	GISEL	0,2983
21	GIIB	0,2780
22	INTERFAZ	0,2651
23	INME	0,2564
24	SIMON	0,2527
25	GIMAT	0,2503
26	GIC	0,2353
27	CIDLIS	0,2317
28	RadioGIS	0,1490
29	GIMBA	0,1440
30	GIEMA	0,1401
31	GIEP	0,1283

Como lo afirman Restrepo y Villegas (2007), esta metodología hace “que grupos de investigación que no son eficientes pero que están cerca de serlo obtengan buenas calificaciones”, que es lo que sucede con los grupos GMPH y GIGBA que se encuentran en las posiciones 10 y 13 de este ranking respectivamente, ubicándose por encima de grupos que sí son eficientes (CIDES, HDSP y STI para el caso de GMPH, y STI en el caso de GIGBA), con diferencias muy leves en sus valores *CE*.

Según esta clasificación, las entidades que obtienen las mejores calificaciones son GEPS, GPH, CAGE, OPALO e INTERFASE, en ese orden. Las primeras cuatro también obtuvieron las principales posiciones según el método de Súper Eficiencia, aunque en diferente orden; no obstante, utilizando el método de Eficiencia Cruzada, el grupo GEOMÁTICA no alcanza a ubicarse entre las primeras cinco posiciones.

GEPS y GPH que obtienen los dos primeros puestos en este ranking, además ocupan las dos primeras posiciones en el conteo del número de veces que aparecen como referentes de grupos ineficientes (véase Figura 25), dejando en claro una fuerte relación entre los resultados obtenidos por Eficiencia Cruzada y la ordenación obtenida según los conjuntos de referencia. Lo anterior podría justificarse en el hecho de que Eficiencia Cruzada utiliza los pesos obtenidos del modelo primal DEA y los conjuntos de referencia son un resultado de la segunda fase de este modelo.

9.2.3 Súper Eficiencia vs Eficiencia Cruzada.

Lo que se quiere conocer es qué tan similares son las puntuaciones de eficiencia asignadas por los dos métodos de clasificación. Para analizar la relación existente entre ambos resultados se ha recurrido al análisis de correlación.

La Figura 26 muestra las calificaciones asignadas a los 31 grupos de investigación enfrentadas según el método utilizado para su cálculo.

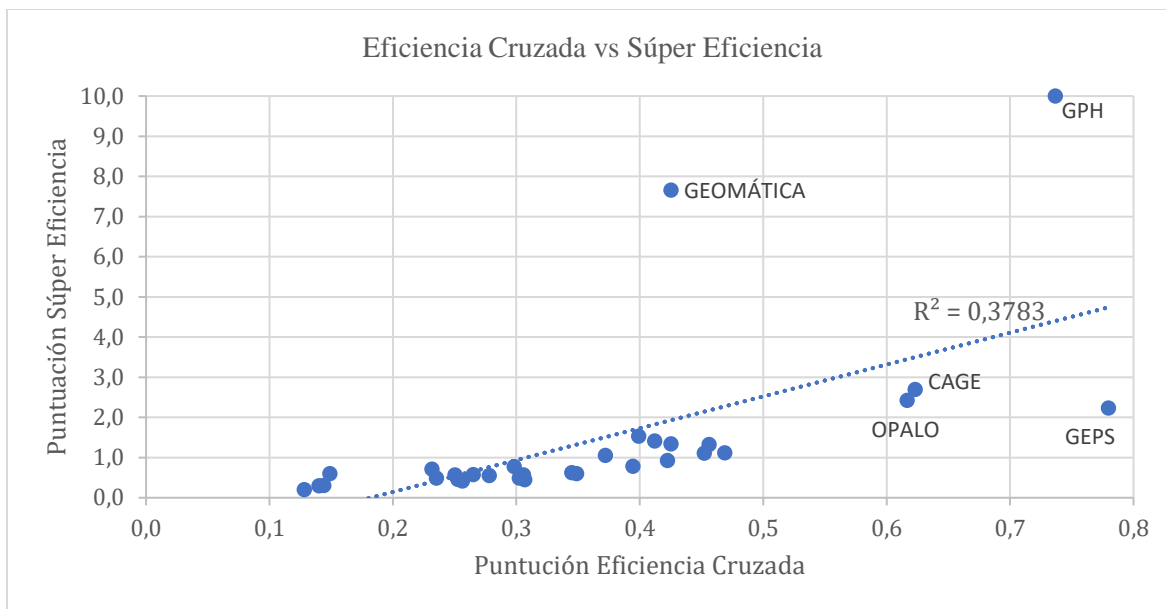


Figura 26. Dispersión Eficiencia Cruzada Vs. Súper Eficiencia.

Resulta evidente que GPH y GEOMÁTICA son observaciones atípicas que afectan el coeficiente de determinación (R^2), el cual indica que el 37,83% de la variación de la puntuación asignada por el método de Súper Eficiencia es explicada por la variación del puntaje dado por Eficiencia Cruzada. De igual forma, estas dos observaciones afectan el coeficiente de correlación de Pearson (R) que toma un valor de 61,51% reflejando así una correlación positiva media.

La aparición de estas observaciones extremas podría justificarse haciendo referencia a autores como Banker & Chang (2006) que afirman que la utilización del método de Súper Eficiencia no siempre arroja resultados satisfactorios para realizar clasificaciones, y que por el contrario es bastante útil para identificar valores atípicos.

Al remitirse a la información ingresada al modelo para obtener las puntuaciones de eficiencia con esta metodología, se encuentra que tanto GEOMÁTICA como GPH son los únicos grupos de investigación *eficientes* que poseen valores cero en alguna variable de entrada, *Presupuesto* y *Estudiantes* respectivamente, es decir, el grupo GEOMÁTICA en el año 2016 no recibió dinero por alguno de los rubros analizados (proyectos, estímulos y movilidad), y el grupo GPH no tenía

vinculado ningún estudiante al año 2016; lo que podría justificar la identificación de éstos como puntos extremos.

Por lo anterior se decidió eliminar las dos observaciones atípicas y realizar nuevamente el correspondiente gráfico de dispersión, obteniéndose así la Figura 27.

Claramente este cambio permite que la línea de tendencia se ajuste mejor a las observaciones, y por consiguiente se consiga un coeficiente de correlación $R = 90,1\%$ que manifiesta una fuerte correlación positiva.

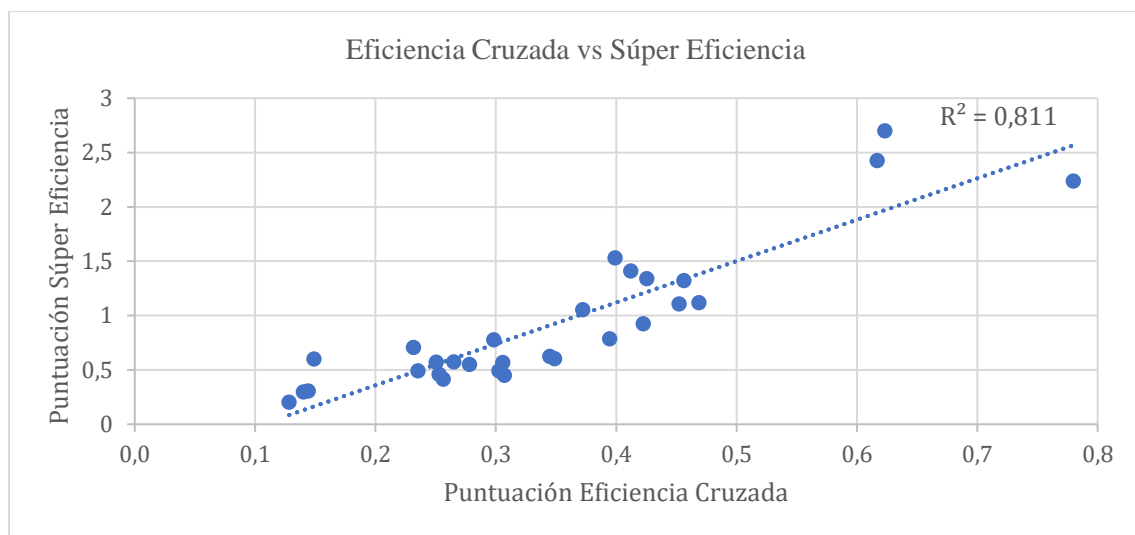


Figura 27. Dispersión Eficiencia Cruzada Vs. Súper Eficiencia sin puntos extremos.

Ahora bien, la relación entre las puntuaciones de eficiencia asignadas por el modelo CCR-O y las obtenidas por los dos métodos de clasificación también resulta de interés. En la Figura 28 se presenta el diagrama de dispersión correspondiente a los valores de eficiencia originales y los obtenidos con el método de Eficiencia Cruzada de los 31 grupos de investigación evaluados. En este caso el coeficiente de correlación de Pearson toma un valor de $R = 80,3\%$ dejando en evidencia una fuerte correlación positiva.

Es importante señalar que en el modelo CCR-O se les asigna a todas las entidades eficientes una puntuación igual a 1, y a las ineficientes un valor menor.

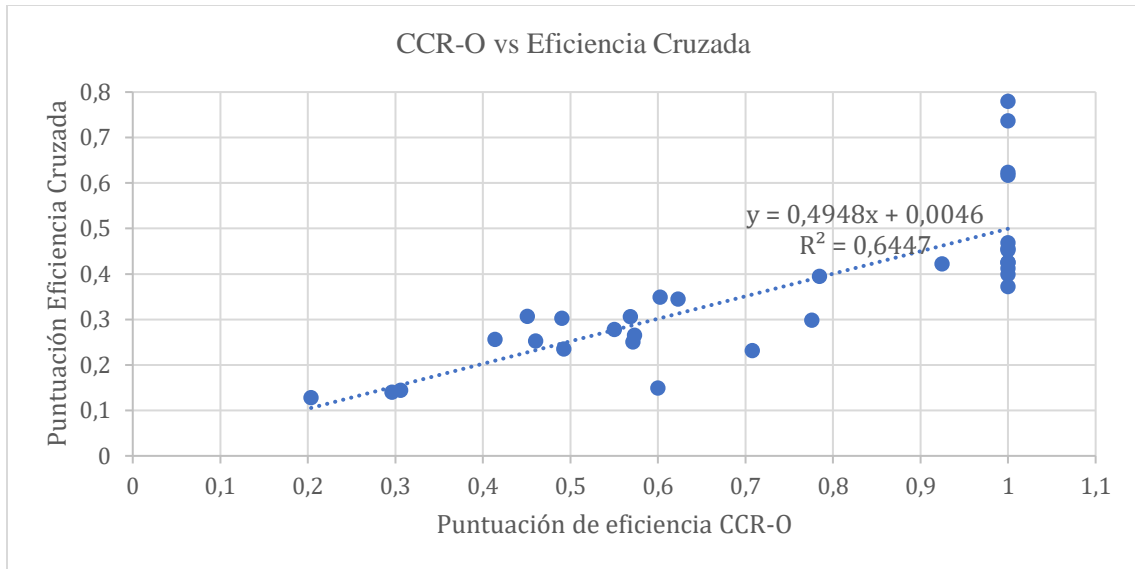


Figura 28. Dispersión CCR-O vs. Eficiencia Cruzada.

En cambio, para observar la relación existente entre los puntajes asignados por el modelo CCR-O y Súper Eficiencia se optó por no considerar las observaciones correspondientes a los grupos GEOMÁTICA y GPH pues, como se analizó anteriormente, toman valores atípicos. En la Figura 29 también es posible observar una fuerte correlación positiva, muy similar a la anterior, esto es un coeficiente de correlación de Pearson de 80,6%.

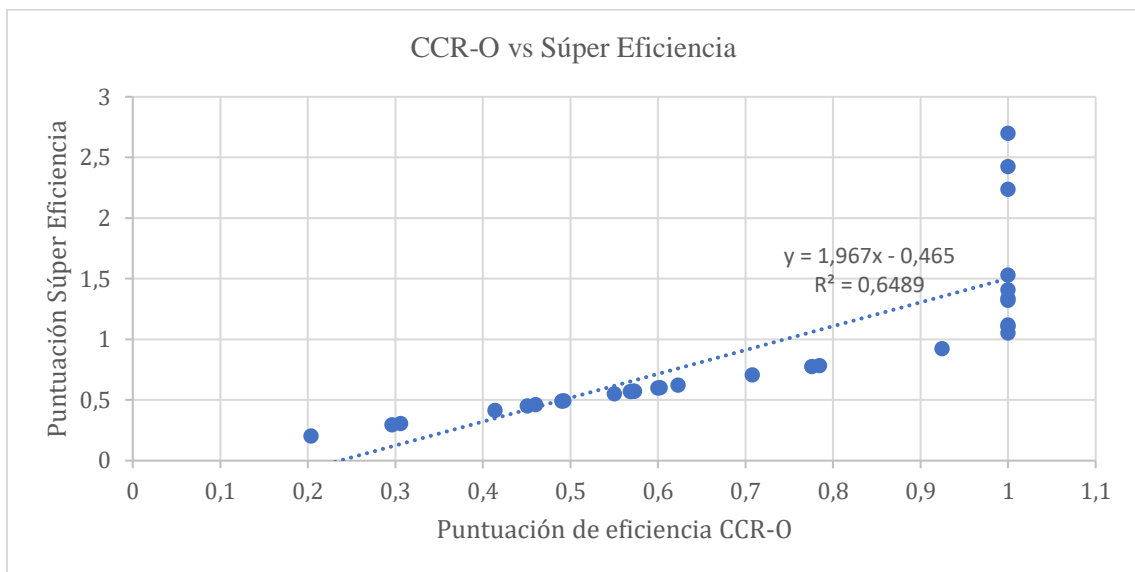


Figura 29. Dispersión CCR-O vs. Súper Eficiencia.

9.2.3.1 Análisis Clúster según puntuación de Eficiencia. En la literatura es común encontrar el uso de los resultados obtenidos a partir de la metodología DEA en conjunto con algunas técnicas estadísticas univariadas y/o multivariadas complementarias. En este caso se utilizó el Análisis Clúster Jerárquico, con distancia euclidiana al cuadrado, para clasificar los grupos en diferentes categorías según la similitud de los mismos. Los datos utilizados para tal fin corresponden a los resultados obtenidos por los modelos CCR-O, Súper Eficiencia y Eficiencia Cruzada.

En primer lugar, se va a efectuar un análisis entre los resultados originales, obtenidos del modelo CCR-O, y los arrojados por el método de Eficiencia Cruzada. Luego se va a proceder de la misma manera con los resultados obtenido por el método de Súper Eficiencia y los de Eficiencia Cruzada. No resultaría dicente la comparación de los resultados del modelo CCR-O y Súper Eficiencia puesto que estos dos arrojan las mismas puntuaciones de eficiencia para las entidades ineficientes.

La Figura 30 corresponde al Dendrograma resultante del Clúster Jerárquico entre las puntuaciones de eficiencia asignadas por el modelo CCR-O y las obtenidas a partir de Eficiencia Cruzada. Si se decidiera realizar una agrupación en 5 clústeres, lo que tendría que hacerse sería trazar una línea vertical que cortase 5 uniones en el diagrama (línea roja), de este modo se obtendría una distribución como la que se muestra en la Tabla 24.

Tabla 24
Análisis Clúster según resultados de CCR-O y Eficiencia Cruzada

Clúster 1	Clúster 2	Clúster 3	Clúster 4	Clúster 5
CAGE	GEOMÁTICA	CIDLIS	DICBoT	GIEMA
OPALO	GIGHC	GISEL	GRM	GIMBA
GEPS	CIDES	GIGBA	GIMAT	GIEP
GPH	HDSP		INTERFAZ	
	CEMOS		GIIB	
	GIP		F&M	
	INTERFASE		CPS	
	STI		INNOTEC	
	GMPH		GIC	
			SIMON	
			INME	
			RadioGIS	

Nótese que de los 12 grupos *Eficientes* 4 se encuentran en el clúster 1, y los 8 restantes en el segundo clúster.

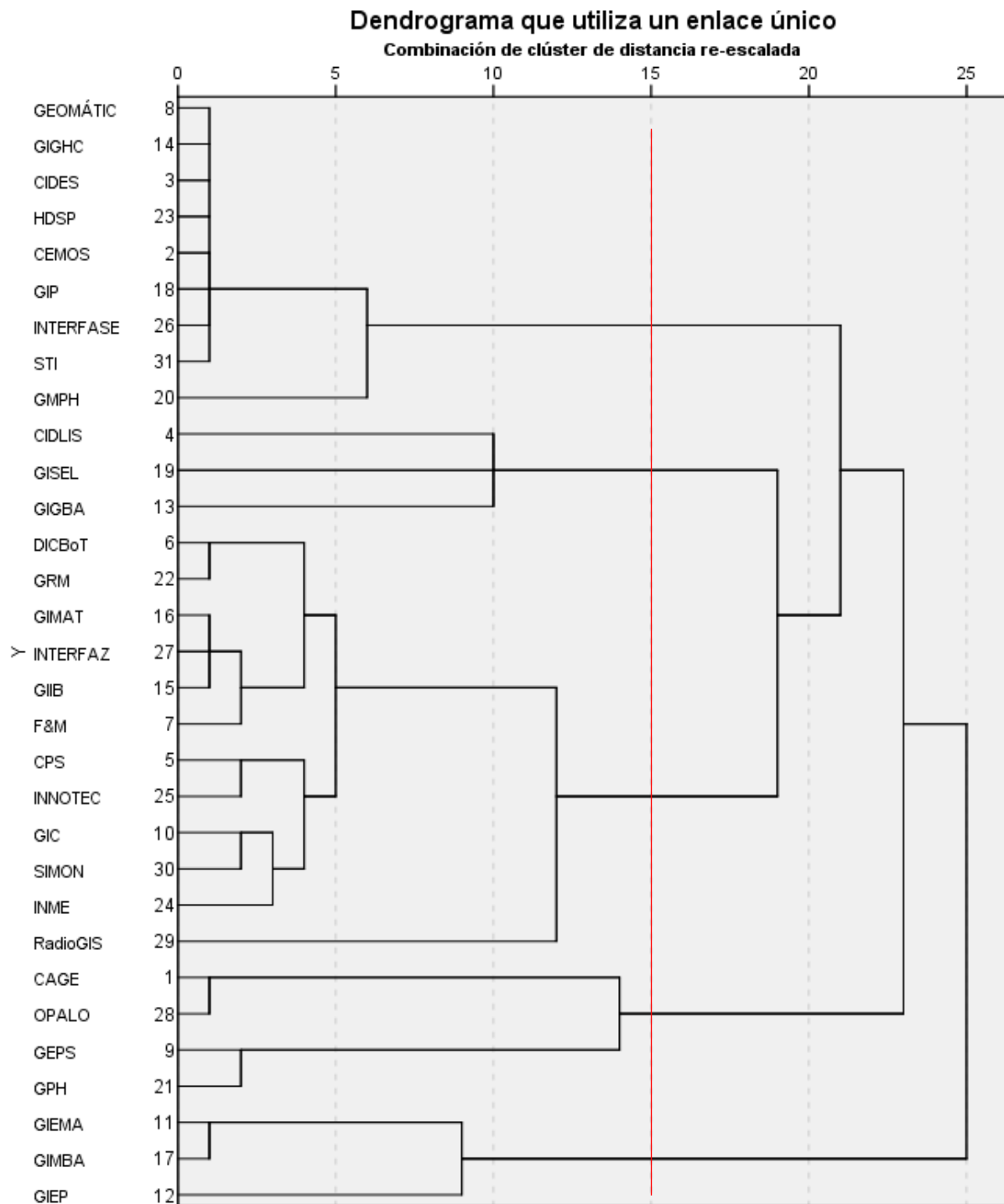


Figura 30. Dendrograma - Clúster CCR-O vs. Eficiencia Cruzada.

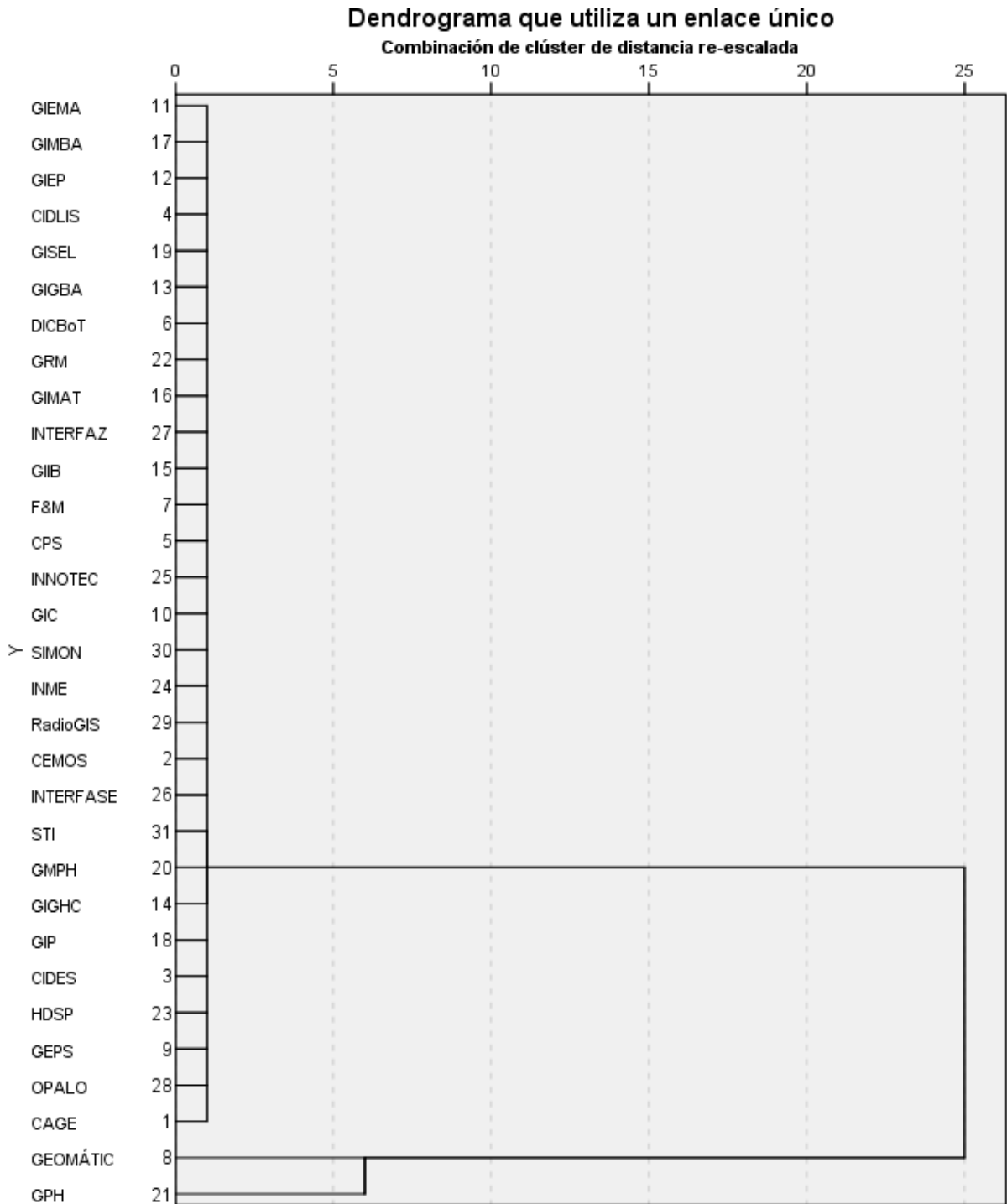


Figura 31. Dendrograma - Súper Eficiencia vs. Clúster Eficiencia Cruzada.

En cuanto al Dendrograma obtenido del Clúster Jerárquico entre las puntuaciones de eficiencia obtenidas en los modelos de Súper Eficiencia y Eficiencia Cruzada (Figura 31), podría decirse que muestra una gran similitud entre los grupos en términos de sus valores de eficiencia, es decir, las

puntuaciones en general son bastante homogéneas. Sin embargo, y como se pudo corroborar en el análisis de correlación anterior, los grupos de investigación GEOMÁTICA y GPH son los que presentan las observaciones más distantes, identificándose de este modo como valores atípicos.

Entonces, con el fin de conocer cómo se agruparían los GI según los resultados de los métodos de Súper Eficiencia y Eficiencia Cruzada, se realizó un análisis clúster de la siguiente manera: Inicialmente se separaron las entidades eficientes de las ineficientes, para luego realizar el análisis con cada grupo de entidades. En cuanto a los *eficientes*, se han omitido los grupos con observaciones atípicas (GEOMÁTICA y GPH) y se ha obtenido de este modo la agrupación reflejada en la Figura 32.

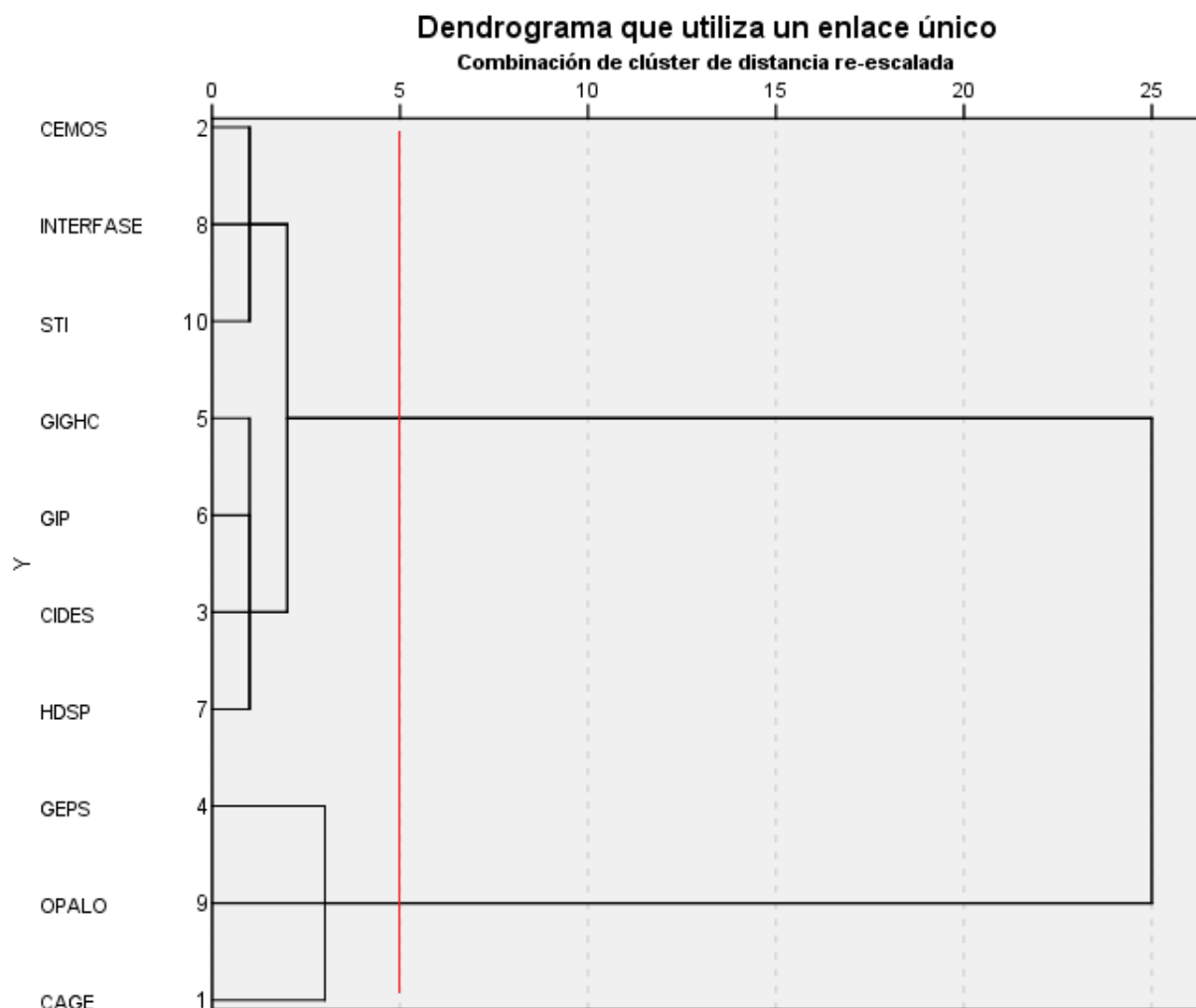


Figura 32. Dendrograma - Clúster entidades Eficientes.

Haciendo una partición en el Dendrograma anterior se obtiene la agrupación que se muestra en la Tabla 25 de los grupos eficientes.

Tabla 25
Agrupación de los grupos Eficientes

Clúster 1	Clúster 2	Clúster 3
GPH	GEPS	CEMOS
GEOMÁTICA	OPALO	INTERFASE
	CAGE	STI
		GIGHC
		GIP
		CIDES
		HDSP

En el caso de los *ineficientes* el análisis se realizó con todos los grupos ya que en este conjunto de entidades no se identificó valor atípico alguno, obteniéndose así el Dendrograma que se muestra en la Figura 33.

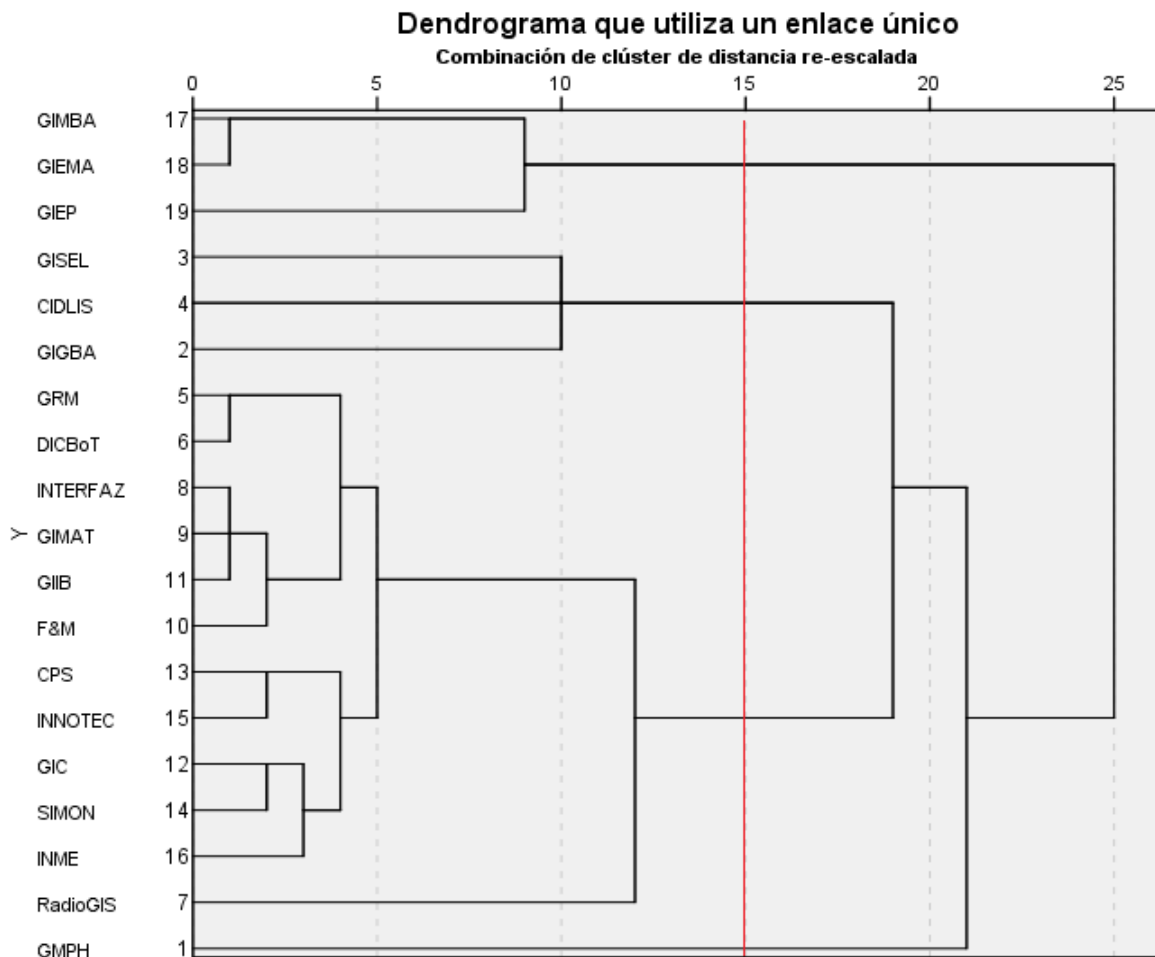


Figura 33. Dendrograma - Clúster entidades Ineficientes.

Para realizar una agrupación en cuatro clústeres, se trazó una línea vertical que cortase 4 uniones en el diagrama (línea roja en el Dendrograma, Figura 33). Aunque GMPH no se adhiere a algún conglomerado al hacer el corte en ese punto, no indica que se trate de un grupo con observaciones atípicas como en el caso de los eficientes, sino que quiere decir que es el grupo cuya ineficiencia es mucho menor en comparación a los demás. La Tabla 26 resume el resultado de la agrupación de los GI ineficientes.

Tabla 26
Agrupación de los grupos Ineficientes

Clúster 1	Clúster 2	Clúster 3	Clúster 4
GIMBA	GISEL	GRM	GMPH
DIEMA	CIDLIS	DICBoT	
GIEP	GIGBA	INTERFAZ	
		GIMAT	
		GIIB	
		F&M	
		CPS	
		INNOTEC	
		GIC	
		SIMON	
		INME	
		RadioGIS	

9.2.4 Clasificación por Escuelas. Una vez obtenida una clasificación a nivel general de todos los grupos de investigación, es interesante analizarlos a nivel de escuelas, a fin de conocer cuántos fueron catalogados como eficientes y/o ineficientes dentro de éstas.

Para efectos de una mejor visualización de esta idea, en las Tablas 27 y 28 se han dividido las escuelas según la facultad, y dentro de ellas se especifican los grupos eficientes e ineficientes con su respectiva posición obtenida en la clasificación dada por el modelo de Súper Eficiencia y Eficiencia Cruzada.

En primera instancia, se analizará la clasificación obtenida según el modelo de Súper Eficiencia. Aquí es importante señalar que las dos primeras posiciones de la clasificación general se encuentran en la escuela de Ingeniería Civil, la cual cuenta con un total de 3 grupos de

investigación, dos de ellos eficientes (GPH y GEOMÁTICA) y uno (INME) que no alcanzó la frontera de eficiencia ocupando el puesto N° 28.

Tabla 27
Clasificación por Escuela (Facultad de Ingenierías Fisicomecánicas)

	Súper Eficiencia		Eficiencia Cruzada		
	Eficiente	Ineficiente	Eficiente	Ineficiente	
Fisicomecánicas	EIC	1 GPH	28 INME	2 GPH	23 INME
		2 GEOMÁTICA		9 GEOMÁTICA	
	EISI	3 CAGE	23 GIIB	3 CAGE	21 GIIB
		6 HDSP	26 SIMON	12 HDSP	24 SIMON
		12 STI		14 STI	
	EEIE	4 OPALO	22 F&M	4 OPALO	18 F&M
			27 INNOTECH		17 INNOTECH
	EDI	5 GEPS	20 INTERFAZ	1 GEPS	22 INTERFAZ
	E3T	11 CEMOS	15 GISEL	7 CEMOS	20 GISEL
			16 CIDLIS		27 CIDLIS
	EIM		18 DICBoT		15 DICBoT
			30 GIEMA		30 GIEMA

Seguando con la Facultad de Fisicomecánicas, se observa que la E3T es la escuela que presenta la mayor cantidad de grupos ineficientes, de un total de 5 GI sólo uno (CEMOS) resultó ser eficiente obteniendo el puesto N°11. En contraste, la Escuela de Ingeniería de Sistemas, que presenta la misma cantidad de grupos que la E3T, tuvo 3 eficientes y 2 ineficientes, además de ello, allí se encuentra la posición N°3 del ranking general (CAGE).

A su vez, las Escuelas de Ingeniería Industrial y Mecánica exhiben una cantidad de 2 GI ineficientes cada una, sin embargo, la EEIE cuenta con un grupo eficiente en la posición N°4 (OPALO) mientras que en la EIM ninguno es eficiente.

A pesar de que las seis primeras posiciones del ranking general dado por Súper Eficiencia son ocupadas por grupos de la Facultad de Fisicomecánicas, en la Facultad de ingenierías Fisicoquímicas la EIQ presenta todos sus grupos como eficientes en las posiciones 7, 9 y 10, siendo la única escuela 100% eficiente en términos de investigación (Ver Tabla 28). Así mismo, es de

resaltar el hecho de que las escuelas de Ingeniería de Petróleos y Metalúrgica, cada una con 3 grupos de investigación, presenten todos sus GI por debajo de la frontera eficiente.

Tabla 28
Clasificación por Escuela (Facultad de Ingenierías Físicoquímicas)

		Súper Eficiencia				Eficiencia Cruzada			
		Eficiente		Ineficiente		Eficiente		Ineficiente	
Físicoquímicas	EIQ	7	CIDES			11	CIDES		
		9	GIP			6	GIP		
		10	INTERFASE			5	INTERFASE		
	ESGEO	8	GIGHC	14	GIGBA	8	GIGHC	13	GIGBA
	EIP			13	GMPH			10	GMPH
				17	GRM			16	GRM
				31	GIEP			31	GIEP
				21	GIMAT			25	GIMAT
	EIMT			24	GIC			26	GIC
				29	GIMBA			29	GIMBA

En contraste con la clasificación asignada mediante el método de Súper Eficiencia, en el ranking dado por Eficiencia Cruzada se observa un cambio bastante marcado en las posiciones de los grupos ineficientes, más que en la de los eficientes, hasta con un movimiento de 11 puestos.

En cuanto a los grupos eficientes, el primer puesto es ocupado por GEPS, perteneciente a la escuela de Diseño Industrial, que asciende a la posición N° 1 desde el puesto 5 dado por Súper Eficiencia, así como INTERFASE, de la EIQ, que tuvo una transición del puesto 10 al 5. Igualmente es importante resaltar el caso de GEOMÁTICA, el cual descendió 7 puestos llegando a la posición N° 9 según este método, y de HDSP que pasó del puesto 6 al puesto 12. Por su parte, los grupos CAGE, OPALO y GIGHC conservaron la posición dada por Súper Eficiencia, esto podría significar que son grupos que presentan una eficiencia más sólida.

Por otra parte, en los ineficientes se observa que el grupo INNOTECH, de la EEIE, pasa de estar en la posición 27, del ranking por Súper Eficiencia, al puesto 17 en el ranking de Eficiencia Cruzada. En cambio, CIDLIS pasa de estar en la posición 16 a la 27, mientras que INME y GISEL

sólo se mueven 5 posiciones respecto a la anterior clasificación, ascendiendo y descendiendo respectivamente.

Nótese que las posiciones 11, 12 y 14, según Eficiencia Cruzada, se encuentran en la columna de los Eficientes, mientras que los puestos 10 y 13 en la columna de Ineficientes. Así, HDSP y STI de la EISI, junto con CIDES de la EIQ, fueron clasificados por debajo de GMPH (puesto 10) que es un grupo de investigación ineficiente. Adicionalmente, el grupo STI es superado en posición por GIGBA (puesto 13) que también fue catalogado como ineficiente.

9.2.5 DEA vs Colciencias. Con el propósito de hacer una breve comparación entre los resultados de las dos últimas convocatorias de Colciencias y los obtenidos según los métodos expuestos anteriormente, se resume en la Tabla 29 las posiciones conseguidas por cada grupo según los modelos Súper Eficiencia y Eficiencia Cruzada, y la clasificación de cada GI según las convocatorias 737 de 2015 y 781 de 2017 de Colciencias.

En primera instancia si se observa el grupo GPH, que obtuvo la primera posición según Súper Eficiencia y la segunda conforme Eficiencia Cruzada, no fue asignado a las categorías más altas (A1 o A) de Colciencias, en ambas convocatorias obtuvo una clasificación C. Mientras que GEOMÁTICA, con la segunda posición según Súper Eficiencia y la primera en Eficiencia Cruzada, se posicionó en la categoría D en el año 2015 y pasó a ser un grupo Reconocido en 2017. Por su parte, CAGE, el grupo que ocupó en tercer lugar en el ranking según los dos modelos anteriormente mencionados, pasó de una categoría D en 2015 a C en 2017.

De forma similar, al observar en primer lugar los grupos según la clasificación dada por Colciencias, se encontró que CEMOS, GIMBA, GIC e INTERFASE mantuvieron su posición en la categoría más alta (A1) en las dos convocatorias. Sin embargo, estos grupos obtuvieron posiciones por debajo de la décima y quinta posición en los modelos de Súper Eficiencia y

Eficiencia Cruzada respectivamente. Es de resaltar el caso del grupo GIMBA, que se mantuvo en el puesto 29 con los dos modelos anteriores, aunque fue catalogado A1 por Colciencias.

Tabla 29
Resultados de la clasificación mediante DEA y Colciencias

Grupo de investigación	Súper Eficiencia	Eficiencia Cruzada	Colciencias Convocatoria 737 de 2015	Colciencias Convocatoria 781 de 2017
CAGE	3	3	D	C
CEMOS	11	7	A1	A1
CIDES	7	11	A1	R*
CIDLIS	16	27	C	NR*
CPS	25	19	A	A
DICBoT	18	15	B	C
F&M	22	18	B	B
GEOMÁTICA	2	9	D	R
GEPS	5	1	A	B
GIC	24	26	A1	A1
GIEMA	30	30	A	A
GIEP	31	31	B	B
GIGBA	14	13	C	A
GIGHC	8	8	C	C
GIIB	23	21	B	B
GIMAT	21	25	A1	A
GIMBA	29	29	A1	A1
GIP	9	6	C	C
GISEL	15	20	A	A
GMPH	13	10	C	C
GPH	1	2	C	C
GRM	17	16	C	A
HDSP	6	12	C	A1
INME	28	23	B	R
INNOTECH	27	17	A	A
INTERFASE	10	5	A1	A1
INTERFAZ	20	22	C	B
OPALO	4	4	B	B
RadioGIS	19	28	D	NR
SIMON	26	24	B	B
STI	12	14	B	B

Nota: *R y NR hace referencia a grupos de investigación Reconocidos y No Reconocidos.

Como resulta evidente en la Tabla 29, los métodos de Eficiencia Cruzada y Súper Eficiencia ordenan los 31 grupos de mayor a menor puntuación de eficiencia, mientras que Colciencias los agrupa en categorías según el valor obtenido en el índice ScientiCol correspondiente a su modelo de medición, así que para corroborar detalladamente las diferencias en las posiciones del ranking dado por Colciencias y el asignado por los métodos DEA se hace necesario que ambos manejen el mismo tipo de clasificación. Para ello, en base al análisis clúster del Apartado 9.2.3.1, se van a agrupar los GI en 5 categorías basadas en las puntuaciones de eficiencia, asignadas por los métodos DEA, para que puedan compararse con las cinco categorías que asigna la convocatoria 737 de 2015 de Colciencias.

En ese sentido, si el nivel más alto de Colciencias es el A1 entonces los grupos que se encuentren en esa categoría serán los primeros en el ranking, es decir, estarán agrupados en el puesto 1, seguidos de los que pertenecen al nivel A (puesto 2), y en el tercer puesto los de la categoría B; los grupos que se encuentren en la categoría C van a pertenecer al puesto 4 y los de la categoría D al puesto 5.

Los 5 clústeres que van a representar la clasificación dada por los modelos DEA serán los que se muestran en la Tabla 30, y que han resultado del análisis clúster realizado en el Apartado 9.2.3.1. En estas agrupaciones el nivel de eficiencia irá descendiendo desde el primer clúster hasta el número 5, de modo que los 12 grupos calificados como eficientes con la metodología DEA van a ubicarse en los clústeres 1 y 2, y los 19 grupos ineficientes de ahí en adelante hasta el clúster 5.

Tabla 30
Categorías de clasificación modelos DEA

Clúster 1	Clúster 2	Clúster 3	Clúster 4	Clúster 5
CAGE	GEOMÁTICA	CIDLIS	DICBoT	GIEMA
OPALO	GIGHC	GISEL	GRM	GIMBA
GEPS	CIDES	GIGBA	GIMAT	GIEP
GPH	HDSP		INTERFAZ	
	CEMOS		GIIB	
	GIP		F&M	
	INTERFASE		CPS	
	STI		INNOTEC	
	GMPH		GIC	
			SIMON	
			INME	
			RadioGIS	

Para realizar la respectiva comparación entre las posiciones, en la primera columna de la Tabla 31 se han ubicado los grupos de investigación en orden alfabético, en la segunda columna se ha registrado la posición de acuerdo al clúster generado en base a los modelos DEA, de modo que en el puesto 1 estarán los grupos pertenecientes al clúster 1, en el puesto 2 los que estén en el clúster 2, y así sucesivamente; y en la tercer columna las posiciones según la convocatoria 737 de Colciencias, en términos numéricos como se planteó anteriormente. En la cuarta columna se ha calculado la diferencia de las posiciones en valores absolutos, a fin de observar en cuántos de los grupos ha coincidido la clasificación y en cuántos ha variado.

Así pues, se ha obtenido que sólo los grupos INTERFAZ y GRM han coincidido en la posición de la clasificación dada por el clúster de los modelos DEA y por la convocatoria 737 de Colciencias, mientras que 14 (45%) de los 31 grupos de investigación tuvieron una diferencia de 1, es decir, variaron su clasificación en un puesto. Los 15 grupos restantes tuvieron una variación de al menos dos posiciones.

Tabla 31
Diferencias entre clasificaciones.

Grupo de Investigación	Clúster modelos DEA	Colciencias Convocatoria 737 de 2015	Diferencia
CAGE	1	5	4
CEMOS	2	1	1
CIDES	2	1	1
CIDLIS	3	4	1
CPS	4	2	2
DICBoT	4	3	1
F&M	4	3	1
GEOMÁTICA	2	5	3
GEPS	1	2	1
GIC	4	1	3
GIEMA	5	2	3
GIEP	5	3	2
GIGBA	3	4	1
GIGHC	2	4	2
GIIB	4	3	1
GIMAT	4	1	3
GIMBA	5	1	4
GIP	2	4	2
GISEL	3	2	1
GMPH	2	4	2
GPH	1	4	3
GRM	4	4	0
HDSP	2	4	2
INME	4	3	1
INNOTEC	4	2	2
INTERFASE	2	1	1
INTERFAZ	4	4	0
OPALO	1	3	2
RadioGIS	4	5	1
SIMON	4	3	1
STI	2	3	1

Si bien el 94% de los grupos presentó variaciones en la clasificación de un método a otro, no es posible afirmar plenamente que los resultados obtenidos de aplicar la metodología DEA discrepan radicalmente de los resultados publicados en la convocatoria 737 de Colciencias, pues 16 (52%) de los 31 grupos analizados tuvieron variaciones en la clasificación de no más de un puesto.

Con el fin de corroborar si las clasificaciones que se muestran en la Tabla 31 se parecen, se aplicó la prueba U de Mann-Whitney, una prueba estadística no paramétrica que permite comparar dos grupos independientes.

Para realizar esta prueba se utilizó el paquete estadístico SPSS de IBM, al cual se ingresó la información correspondiente a la variable independiente categórica *Método* conformada por dos niveles: Ranking DEA y convocatoria 737 de 2015 de Colciencias; y la variable dependiente ordinal, correspondiente a las clasificaciones en términos cuantitativos de dichos métodos, a partir de la cual se van a comparar los grupos de la variable independiente (véase Tabla 31).

Así bien, la Hipótesis nula que la prueba U plantea es:

$H_0 =$ la distribución de clasificación es la misma entre las categorías de *Método*

$H_1 =$ la distribución de clasificación difiere entre las categorías de *Método*

Si la significancia asintótica (prueba bilateral) obtenida es menor al nivel de significancia establecido del 5% se rechaza la Hipótesis nula, de lo contrario se acepta.

Los resultados de la prueba arrojan un nivel de significancia de 74,6%, que indica que la decisión es aceptar la Hipótesis nula y concluir que no hay diferencias significativas entre la clasificación asignada por los dos métodos.

La diferencia que se da entre el método DEA y el modelo de medición utilizado por Colciencias radica en que, por una parte, la ventana de observación (tiempo considerado) en este trabajo fue limitada por la no disponibilidad de información que solo permitió analizar un año (2016), mientras que Colciencias, en ambas convocatorias, tiene en cuenta un periodo de 7 y 5 años para *Artículos* y *Tesis de grado*²⁰ respectivamente. Así mismo, es de resaltar el hecho de que Colciencias considera, para la medición de los GI, una mayor cantidad de productos (Ver Apéndice B) en

²⁰ Solo se hace mención a la ventana de observación de estos dos productos (Artículos y Tesis) ya que son los únicos que se consideran en este trabajo.

comparación con los dos utilizados aquí; en las variables de entrada también se observa una diferencia significativa, ya que como lo afirma Rojas, (2010) “la relación entre los recursos de un grupo y su producción no es visible en el método usado por Colciencias”. Por otro lado, a diferencia de Colciencias, DEA no asigna ponderaciones a priori a cada insumo y producto tenido en cuenta en la evaluación, sino que las asigna de manera que beneficie la puntuación de eficiencia de cada entidad.

10. Conclusiones

Si bien no es posible establecer en definitiva que el Análisis Envolvente de Datos es la técnica más adecuada para evaluar la eficiencia de grupos de investigación, sí es apropiado señalar que constituye una metodología simple y sencilla para realizarlo. Los modelos DEA además de proporcionar una medida de eficiencia que es relativa a otros grupos dentro del mismo conjunto de entidades evaluadas, muestra los grupos eficientes que son referentes para aquellos que no alcanzaron la máxima eficiencia, de manera que sirvan como guía en el proceso de mejora. Más aún, esta metodología le permite al investigador seleccionar el tipo de mejora que le interesa identificar –que en este trabajo se enfocó a la maximización de los productos– para que finalmente en los resultados conozca en qué cantidades es posible incrementar la producción para alcanzar una puntuación del 100% en eficiencia.

Previo a la aplicación del modelo DEA, el análisis clúster de las variables permitió corroborar la homogeneidad de las entidades; identificar las observaciones con valores extremos en alguna de las variables de entrada y/o salida, siendo éste el caso de HDSP, que presenta el máximo valor en

Estudiantes y Artículos, CEMOS con la mayor cantidad de *Presupuesto* asignado y OPALO con la mayor dirección de *Tesis de grado*; e identificar las entidades cuyos valores en sus variables son cercanos, es decir, se identificó que 28 grupos del total presentan afinidad en los insumos *Docentes y Estudiantes*, y que en 25 grupos dicha afinidad se ve refleja principalmente en el producto *Tesis*.

De los 34 grupos de investigación que pertenecen a las Facultades de Ingeniería de la Universidad Industrial de Santander, sólo fue posible realizar el análisis con 31 de ellos, debido a que los demás no presentaron producción en el año 2016 según la información suministrada por la Vicerrectoría de Investigación y Extensión, es decir, que tanto en la variable *Artículos* como en la de *Tesis* no tuvieron producción alguna para ese año, anulando completamente las variables de salida; tal es el caso de los grupos CIDIC, GIROD y FILOEN. Ahora, aunque algunos grupos presentaron valores de cero en alguna de las variables (CIDES, GEMÁTICA, GHP y RadioGIS) sí se tuvieron en cuenta para el análisis debido a que al menos tenían valores no nulos en algunas de las variables de entrada y de salida.

Por su parte, la aplicación del modelo DEA CCR orientado a las salidas, en dos fases, arrojó que 12 (38,8%) de los 31 grupos evaluados fueron catalogado como eficientes, es decir, obtuvieron una puntuación de eficiencia del 100%; mientras que, de los 19 restantes que no alcanzaron la máxima calificación, 11 fueron puntuados por encima del 55%, lo que indica que la mayoría de los grupos ineficientes obtuvieron un nivel de eficiencia medio.

Así mismo, los resultados de esta aplicación han permitido identificar los grupos de investigación eficientes que aparecen un mayor número de veces como referentes de grupos ineficientes: GEPS ocupa el primer lugar con un total de 16 veces, seguido por GPH y OPALO con 10 y 8 respectivamente. Lo anterior permite considerar tales grupos como aquellos con la tecnología de producción más común y cuya eficiencia sería más sólida.

Adicionalmente, los resultados han permitido determinar las mejoras en la producción que cada grupo de investigación calificado por debajo del 100% podría realizar para alcanzar la puntuación máxima de eficiencia, por ejemplo, uno de los grupos que tendría que efectuar un mayor porcentaje de mejora es GIEP (79,6%), que debe producir, por ejemplo, 2 artículos A2, 1 B y dirigir adicionalmente 2 tesis doctorales para ser eficiente, siendo éstas únicamente mejoras radiales dado que este grupo no presenta valores de holgura. Es de aclarar que estas mejoras son una combinación de las tantas posibles que podría aplicar el grupo para alcanzar la eficiencia, en este trabajo se presentan aquellas que señalan la mínima cantidad de producto adicional requerido. En el caso de GIEP podría no producir 3 artículos como se mencionó anteriormente, sino producir en su lugar 4 artículos B para obtener el mismo puntaje de mejora, además de aquella que debe efectuar en la dirección de Tesis.

Otra de las mejoras más destacadas se observa en el grupo GIEMA, que tendría que producir mínimo 6 artículos adicionales (3 A1 y 3 A2) y además dirigir 6 tesis (1 en maestría y 5 en doctorado), para alcanzar un puntaje de mejora de 5,39 en Artículos y 5,35 en Tesis; o bien, producir 5 artículos A1 y 2 tipo C además de dirigir 4 tesis de pregrado y 9 de maestría para obtener esos mismos puntajes.

Para la clasificación de los GI, se optó por aplicar los dos métodos que frecuentemente son utilizados como complemento a la metodología DEA, tal es el caso de Súper Eficiencia y Eficiencia Cruzada. Se decidió aplicar ambos modelos teniendo en cuenta que la particularidad de cada uno trae consigo ventajas significativas.

En el caso de Súper Eficiencia, los cinco primeros puestos fueron ocupados por GPH, GEMOMÁTICA, CAGE, OPALO y GEPS, en ese orden. Siendo GEPS, GPH y OPALO los grupos que aparecen en un mayor número de conjuntos de referencia. Téngase en cuenta que Súper

Eficiencia no solo ordena las entidades eficientes de la más a la menos eficiente, sino que adicional a ello permite identificar aquellas consideradas como datos atípicos, es decir, que tienen un comportamiento marcadamente diferente en comparación con los otros y que por tal razón son mejores que las demás.

Por su parte, Eficiencia Cruzada arrojó una clasificación en la cual los primeros puestos fueron ocupados por GEPS, GPH, CAGE, OPALO e INTERFASE, dejando en claro una fuerte relación con el orden asignado según el número de veces que aparecen los grupos eficientes en los conjuntos de referencia. Una característica importante de esta técnica de clasificación es que no solo toma en cuenta las entidades catalogadas como eficientes según el DEA, sino que compara y ordena tanto las eficientes como las no eficientes y permite conocer a través del ranking, aquellas entidades que, aunque no alcanzaron la frontera eficiente, no están lejos de serlo, tal es el caso de GMPH y GIGBA, los cuales están por encima de algunos grupos eficientes en el ranking.

Para comparar los resultados obtenidos entre ambas metodologías de clasificación, y entre éstas y los resultados originales del modelo DEA, se realizó un análisis de correlación. El primero de ellos realizado entre las puntuaciones obtenidas por los modelos de Súper Eficiencia y Eficiencia Cruzada, señaló una relación moderada entre ellos, con un coeficiente de correlación de Pearson $R = 61,51\%$, sin embargo, este resultado se debe a la presencia de grupos con observaciones atípicas (GEOMÁTICA y GPH) que al ser omitidos del análisis de dispersión dicho coeficiente manifiesta una fuerte correlación positiva ($R=90,1\%$). Por otra parte, del análisis de correlación entre los resultados originales del modelo DEA (CCR-O) y los obtenidos mediante Eficiencia Cruzada, se obtuvo que estos dos modelos se correlacionan altamente de forma positiva, con un valor R de $80,3\%$. De igual forma ocurre en el análisis entre los resultados originales y los obtenidos por Súper Eficiencia que arroja un coeficiente de correlación de Pearson (R) de $80,6\%$.

Así mismo, se llevó a cabo un análisis clúster a partir de las puntuaciones obtenidas por los modelos de Súper Eficiencia, Eficiencia Cruzada y CCR-O, con el cual fue posible identificar los grupos cuyas puntuaciones de eficiencia son cercanas y generar agrupaciones de éstos. De este modo se obtuvieron cinco clústeres que permiten identificar a los grupos CAGE, OPALO, GEPS y GPH como los más eficientes.

Finalmente, la comparación de los resultados obtenidos en este trabajo y la clasificación dada por Colciencias muestra que, por ejemplo, los grupos CEMOS, GIMBA, GIC e INTERFASE obtuvieron una categoría A1 en las convocatorias de 2015 y 2017, y fueron clasificados como *Ineficientes* por el modelo DEA. Así que para asegurar que las dos metodologías (DEA y Colciencias) arrojan resultados similares o difieren entre ellas, se realizó un estudio de las distancias entre las posiciones que ocupan los grupos en la clasificación dada por ambos métodos, y se encontró que 16 de los 31 grupos analizados (52%) no tuvieron variaciones por encima de un puesto, razón por la cual no es posible afirmar con certeza que los resultados obtenidos al aplicar la metodología DEA discrepan radicalmente de los resultados publicados en la convocatoria 737 de Colciencias. Para verificar lo anterior se realizó la prueba no paramétrica de Mann-Whitney entre las clasificaciones y se encontró un nivel de significancia del 74,6%, lo que indica que no existen diferencias significativas entre ambas clasificaciones.

11. Recomendaciones

La aplicación de esta metodología requiere de bases de datos completas y confiables pues la calidad de los resultados depende en su totalidad de la calidad de la información que se ingresa.

Así pues, es recomendable llevar un registro actualizado, continuo y detallado de las actividades semestrales de cada grupo de investigación, donde se incluyan los integrantes y el periodo de tiempo que llevan activos en el grupo, su nivel de formación en cada periodo académico, las horas dedicadas a la investigación de cada uno de ellos, los proyectos desarrollados con su respectiva duración y producción, además de toda la producción realizada adicionalmente; es igualmente importante el registro de los ingresos económicos tipificados.

La medida de la eficiencia de los grupos de investigación de las facultades de ingeniería, se debería realizar regularmente a manera de seguimiento de la actividad investigativa en la Universidad. El presente trabajo sienta las bases para la creación de un indicador que permita conocer periódicamente el desempeño de los grupos de investigación. Debido a que las clasificaciones de eficiencia no son absolutas, sólo podrá catalogarse a un grupo de investigación como *eficiente*, si a lo largo de estimaciones periódicas de esta medida, se ha mantenido en posiciones de eficiencia.

Sería interesante pensar en analizar todos los grupos de investigación de la universidad con el fin de revisar, a nivel interno, su comportamiento en términos investigativos. Si bien el presente trabajo es un punto de partida, en éste no se toman en cuenta la totalidad de los grupos de investigación de la institución, sino solo una parte de ellos. Lo anterior, debido a las condiciones que exige la metodología DEA en cuanto a la homogeneidad de las unidades a evaluar. Sin embargo, podría pensarse en realizar trabajos futuros donde se cataloguen los GI por ramas de conocimiento, o como Experimentales y No Experimentales y de esta forma poder evaluarlos mediante el DEA.

Referencias Bibliográficas

- Abbott, M., & Doucouliagos, C. (2003). The efficiency of Australian universities-a data envelopment analysis. *Economics of Education Review*, 22, 89-97.
- Aigner, D. J., & Chu, S. F. (1968). On Estimating the Industry Production Function. *American Economic Review*, 58, 826-839.
- Aigner, D. J., Lovell, C., & Schmidt, P. (1977). Formulation and Estimation of Stochastic Frontier Production Models. 6, 21-37.
- Alberto, C. L., Carignano, C., & Ercole, R. (2010). Análisis de eficiencia de las universidades públicas en Argentina mediante métodos no paramétrico. *Educere et Educare*, 5(10).
- Ali, A. I., & Seiford, L. (1993). The mathematical programming approach to efficiency analysis. En H. Fried, S. Schmidt, & K. Lovell, *Measurement of Productive Efficiency : Techniques and Applications* (págs. 120-159). Oxford University Press. Obtenido de <https://ebookcentral.proquest.com/lib/bibliouis-ebooks/detail.action?docID=241626>
- Álvarez, A. (2013). *La medición de la eficiencia y la productividad*. Difusora Larousse - Ediciones Pirámide.
doi:<https://ebookcentral.proquest.com/lib/bibliouissp/detail.action?docID=3428776>
- Andersen, P., & Petersen, N. C. (Octubre de 1993). A Procedure for Ranking Efficient Units in Data Envelopment Analysis. *Management Science*, 39(10), 1261-1264.
- Arenas, W., Soto, J., & Rivera, O. (2004). La evaluación de los grupos de investigación según los indicadores de eficiencia de colciencias versus su evaluación según el análisis envolvente de datos-DEA. *SCIENTIA ET TECHNICA*, 189-194.

- Aristizábal, D., Castro, M., Echeverri, P., & Valencia, J. (Septiembre de 2017). Assessing the efficiency of science, technology and innovation using Data Envelopment Analysis (DEA): The case of Colombia. *DYNA*, 84(202), 215-220. doi:<http://dx.doi.org/10.15446/dyna.v84n202.59554>
- Banker, R., & Chang, H. (2006). The super-efficiency procedure for outlier identification, not for ranking efficient units. *European Journal of Operational Research*, 175(2), 1311-1320.
- Banker, R., Charnes, A., & Cooper, W. (1984). Some Models for Estimating Technical and Scale Inefficiencies in Data Envelopment Analysis. *Management Science*, 30(9), 1078-1092.
- Barbosa, S. (2010). *Evaluación de la eficiencia de las escuelas de la Universidad Industrial de Santander aplicando Análisis Envolvente de Datos (DEA)*. Bucaramanga.
- Boussofiane, A., Dyson, R. G., & Thanassoulis, E. (1991). Applied data envelopment analysis. *European Journal of Operational Research*, 52, 1-15.
- Centro Virtual de Noticias - CVN. (2005). ¿Qué tanto promueven las investigación las universidades de Colombia? *Revista Credencial*, 66-70. Obtenido de <http://www.mineduacion.gov.co/cvn/1665/fo-article-88439.pdf>
- Charnes, A., Cooper, W., & Rhodes, E. (1978). Measuring the Efficiency of Decision Making Units. *European Journal of Operational Research*, 2, 429-444.
- Charnes, A., Cooper, W., Golany, B., & Seiford, L. (1985). Foundations of data envelopment analysis for Pareto-Koopmans efficient empirical production functions. *Journal of Econometrics*, 30(1), 91-107.
- Coelli, T., Rao, D., O'Donnell, C., & Battese, G. (2005). *An introduction to efficiency and productivity analysis* (Segunda ed.). Srpinger Science + Business Media, Inc.

Colciencias. (s.f.). *¿Qué es un grupo de investigación?* Obtenido de <http://legadoweb.colciencias.gov.co/faq/qu-es-un-grupo-de-investigaci-n>

Colciencias. (2014). *Modelo de medición de grupos de investigación, desarrollo tecnológico o de innovación y de reconocimiento de investigadores del sistema nacional de ciencia, tecnología e innovación año 2014*. Departamento administrativo de ciencia, tecnología e innovación, Bogotá.

Colciencias. (2015). *Estado de la ciencia en Colombia*. Recuperado el Julio de 2017, de <https://sites.google.com/a/colciencias.gov.co/estado-de-la-ciencia-2015/mapa>

Coll, V., & Blasco, O. (2006). *Evaluación de la eficiencia mediante el Análisis Envolvente de Datos*. Universidad de Valencia, Valencia.

Cooper, W., Seiford, L., & Tone, K. (2006). *Introduction to Data Envelopment Analysis and its uses*. Springer Science+Business Media, Inc.

Daraio, C., & Simar, L. (2007). *Advanced Robust and Nonparametric Methods in Efficiency Analysis: Methodology and Applications*. Springer.

Debreu, G. (1951). The coefficient of Resource Utilization. *Econometrica*, 19(3), 273-292.

Díez, F. d. (2007). *Análisis de eficiencia de los departamentos universitarios. El caso de la universidad de Sevilla*. Madrid: DYKINSON, S.L.

Doyle, J., & Green, R. (Mayo de 1994). Efficiency and Cross-Efficiency in DEA: Derivations, Meanings and uses. *Operational Research Society Ltd.*, 45(5), 567-578.

Färe, R., & Grosskopf, S. (1994). Estimation of returns to scale using data envelopment analysis A comment. *European Journal of Operational Research*, 79, 379-382.

- Farrell, M. J. (1957). The measurement of productive efficiency. *Journal of the Royal Statistical Society*, 120(3), 253-190.
- Galache, T., Gómez, T., Pérez, F., Rivas, C., Ruiz, C., Sánchez, J., . . . Caballero, R. (2010). Análisis de la eficiencia de los grupos de investigación por ramas de conocimiento. *EconPapers*, 5, 733-744.
- Giraldo, N. (2007). *Evaluación de los grupos de investigación según los indicadores de eficiencia de colciencias versus su evaluación según el análisis envolvente de datos-DEA*. Tesis pregrado, Universidad Tecnológica de Pereira, Facultad de ingeniería, Pereira.
- Golany, B., & Roll, Y. (1989). An application procedure for DEA. *Omega*, 17(3), 237-250.
- Gutiérrez, A., & Maz, A. (2004). Cimentando un proyecto de investigación: La revisión de la literatura. *Revista EM*, 9(1), 20-37.
- Johnes, J. (2006). Data envelopment analysis and its application to the measurement of efficiency in higher education. *Economic of Education Review*, 25(3), 273-288.
- Johnes, J., & Yu, L. (2008). Measuring the research performance of Chinese higher education institutions using data envelopment analysis. *China Economic Review*, 19, 679–696. doi:10.1016/j.chieco.2008.08.004
- Katharaki, M., & Katharakis, G. (2010). A comparative assessment of Greek universities' efficiency using quantitative analysis. *International Journal of Educational Research*, 49, 115-128.
- Knorr, K., Mittermeir, R., Aichholzer, G., & Waller, G. (1979). Leadership and group performance: a positive relationship in academic research units. En F. Andrews, *Scientific Productivity: The Effectiveness of Research Groups in Six Countries* (págs. 95-120). Viena.

- Korhonen, P., Tainio, R., & Wallenius, J. (2001). Value Efficiency Analysis of Academic Research. *European Journal of Operational Research*, 130, 121-132.
- López, O., & Suárez, E. (2011). *Evaluación de la eficiencia de las instituciones educativas oficiales de Bucaramanga mediante el análisis envolvente de datos DEA*.
- Lovell, K. (1993). Production frontiers and productive efficiency. En H. Fried, S. Schmidt, & K. Lovell, *Measurement of Productive Efficiency : Techniques and Applications* (págs. 3-67). Oxford University Press. Obtenido de <https://ebookcentral.proquest.com/lib/bibliouis-ebooks/detail.action?docID=241626>
- Meeusen, W., & Van den Broeck, J. (1977). Efficiency Estimation from Cobb-Douglas Production Functions with Composed Error. *International Economic Review*, 18, 435-444.
- Ministerio de Educación Nacional. (Diciembre de 2008). *Metodología para la distribución de recursos artículo 87 de la ley*. Bogotá.
- Murias, P. (2003). *Metodología de aplicación de Análisis Envolvente de Datos: Evaluación de la eficiencia técnica en la Universidad de Santiago de Compostela*. Tesis doctoral, Universidad de Santiago de Compostela, Departamento de métodos cuantitativos para la economía y la empresa, Santiago de Compostela.
- Oral, M., & Yolalan, R. (1990). An empirical study on measuring operating efficiency and profitability of bank branches. *European Journal of Operation Research*, 46, 282-294.
- Organización para la cooperación y el Desarrollo Económico. (2006). *Manual de Oslo: Guía para la recogida e interpretación de datos sobre innovación*. Grupo Tragsa.
- Pautasso, M. (2013). Ten simple rules for writing a literature review. *Plos Computational Biology*, 9(7). doi:10.1371/journal.pcbi.1003149

- Pino, J., Solís, F., Delgado, M., & Barea, R. (2010). Evaluación de la eficiencia de grupos de investigación mediante análisis envolvente de datos (DEA). *El profesional de la información*, 19(2), 160-167. doi:DOI: 10.3145/epi.2010.mar.06
- Red Iberoamericana de indicadores de ciencia y tecnología - RICYT. (2017). *Red Iberoamericana de indicadores de ciencia y tecnología*. Obtenido de <http://dev.ricyt.org/ui/v3/comparative.html?indicator=INVESTTEJCSEPER>
- Restrepo, B. (s.f.). *Conceptos y Aplicaciones de la Investigación Formativa, y Criterios para Evaluar la Investigación científica en sentido estricto*. Obtenido de https://www.cna.gov.co/1741/articles-186502_doc_academico5.pdf
- Restrepo, M. I., & Villegas, J. G. (2007). Clasificación de grupos de investigación colombianos aplicando análisis envolvente de datos. *Facultad de ingeniería Universidad de Antioquia*(42), 105-119.
- Rojas, M. (2010). *Clasificación de los grupos de investigación de la facultad de ingeniería de la universidad nacional de Colombia, mediante la estimación de la eficiencia técnica utilizando Análisis Envolvente de Datos*. Tesis de maestría, Universidad Nacional de Colombia, Facultad de Ingeniería, Bogotá.
- Sagarra, M., Mar-Molinero, C., & Agasisti, T. (2017). Exploring the efficiency of Mexican universities. Integrating DEA and multidimensional scaling. *Omega*, 67, 123-133. Obtenido de <https://doi.org/10.1016/j.omega.2016.04.006>
- Sengupta, J. (1990). Transformations in Stochastic DEA Models. *Journal of Econometrics*, 46, 109-123.
- Sexton, T., Silkman, R., & Hogan, A. (1986). Data Envelopment Analysis: Critique and Extensions. *New Directions for Program Evaluation* , 1986(32), 73.105.

- Tone, k. (2001). A slacks-based measure of efficiency in data envelopment analysis. *European Journal of Operational Research*, 130(3), 498-509. Obtenido de [https://doi.org/10.1016/S0377-2217\(99\)00407-5](https://doi.org/10.1016/S0377-2217(99)00407-5)
- Torrico, A., Teodoro, G., Molina, J., Gómez, T., & Caballero, R. (2007). Análisis de la eficiencia de las unidades productivas de una Universidad. *Revista Electrónica de Comunicaciones y Trabajos de ASEPUMA*(8), 163-195.
- Tse, C., & Yew, K. (2011). Efficiency assessment of universities through data envelopment analysis. *Procedia Computer Science*, 3, 499-506. Obtenido de <https://doi.org/10.1016/j.procs.2010.12.084>
- Visbal, D., Martínez, M., & Guijarro, F. (2017). Assessing the Efficiency of Public Universities through DEA. A Case Study. *Sustainability*, 9(8). doi:10.3390/su9081416
- Visbal, D., Mendoza, A., & Causado, E. (2016). Eficiencia en las instituciones de educación superior públicas colombianas: una aplicación del análisis envolvente de datos. *Civilizar*, 16(30), 105-118.
- Yáñez, G., García, A., & González, M. (2015). Research Efficiency Assessment of Colombian Public Universities 2003-2012: Data Envelopment Analysis. *INGE CUC*, 11(2), 97-108. doi:<http://dx.doi.org/10.17981/ingecuc.11.2.2015.10>