

COMPLEMENTACIÓN DE LAS CATEGORÍAS DE IMPACTO EMPLEADAS EN ANÁLISIS DE CICLO DE VIDA, UTILIZANDO CLASIFICADORES BAYESIANOS DE REDES NEURONALES

USE OF BAYESIAN NEURAL NETWORK CLASSIFICATION SYSTEM EMPLOYED IN LIFE CYCLE ANALYSIS FOR COMPLETING THE IMPACT CATEGORIES

BERLAN RODRÍGUEZ PÉREZ¹

RECIBIDO: 12-01-11
ACEPTADO 07-05-11

¹ Universidad Cienfuegos Carlos Rafael Rodríguez, Facultad de Ciencias Económicas y Empresariales. Cienfuegos, Cuba. E-mail: reducf@ucf.edu.cu

RESUMEN

El presente trabajo surge a partir de la aplicación de varios métodos de Análisis de Ciclo de Vida (ACV), para la comparación de los *Impactos Potenciales*, a un mismo juego de datos. A partir de dicha comparación, se ponen de manifiesto varias diferencias significativas en el cálculo de los impactos. Luego de un estudio posterior, se identificaron que las diferencias provenían de las exclusiones por algunos métodos de sustancias significativas para los impactos en otros. Esta situación motiva al estudio del estado de inclusión o no de sustancias por los diferentes métodos. Para conocer las coincidencias y divergencias que tenían los diferentes métodos, se listan por cada categoría, todas las sustancias que se incluyen y luego se comparan, con las contenidas en los demás métodos para cada una de las categorías equivalentes. Encontrándose que en algunos métodos, las sustancias empleadas representaban un bajo por ciento de todas las posibles. En este sentido se procede a estudiar la forma de incluir todas las sustancias posibles en un mismo método. Al realizarse un gran número de estimaciones, se optó por seleccionar el Clasificador Bayesiano de Redes Neuronales, para seleccionar el estimador más adecuado en cada caso. Los resultados del presente trabajo consistieron en la caracterización de los principales métodos de ACV en cuanto a su amplitud y completamiento.

Palabras clave: Impactos ambientales, comparación, métodos, Análisis de Ciclo de Vida.

SUMMARY

Present research has a foundation of applying various comparison methods of Life Cycle Analysis (LCA) for Impact Potentials using the same set of data. In comparisons we made, some significant differences in calculated results of the impacts were obtained. Later study showed that the differences came from excluding critical factors in some methods. These results led us to further conducting analysis by either including or excluding some factors in different methods used. To elucidate some coincidences or divergences observed among different methods, we listed and analyzed all the factors included in each category. We then compared the factors of each method in equivalent categories. We found that in some methods, certain factors used represented low coverage of all possibilities. We therefore proceeded to include all possibilities of factors within the same method. Upon obtaining a large number of estimates, we decided to opt for Bayesian Neural Network for selecting the best estimator in each case. Results thus obtained in our research entail that characterizing and using the principal methods of LCA in terms of their extent and completeness is of utmost importance.

Key words: populated areas, seismic vulnerability, constructive typology, macro-seismic scale, seismic scenario, project evaluation.

1. INTRODUCCIÓN

Las categorías de impacto empleadas en Análisis de Ciclo de Vida (ACV), cuentan con una serie de sustancias las cuales están relacionadas entre sí, a través de las respectivas equivalencias a las Sustancias Base de la categoría. De esta forma se convierten las cantidades másicas de todas las sustancias incluidas en el estudio a cantidades de sustancias equivalentes, es quizás el ejemplo más clásico, el empleado por los expertos del Panel Intergubernamental para el Cambio Climático (*Intergovernmental Panel on Climate – IPCC*), el de que cada molécula de metano (CH_4) que se emite a la atmósfera, es equivalente a 21 moléculas de dióxido de carbono (CO_2), es por ello que en esta categoría el factor de conversión del metano a la sustancia base, el dióxido de carbono, es 21. Un caso similar ocurre con las demás categorías, solo que se emplean otras sustancias como bases (Brakkee. y Huijbregts, 2008; Reap *et al.*, 2008).

Luego de seleccionar un método de Evaluación de Impacto Ambiental en el ACV, se procede a calcular las diferentes categorías que tiene incluidas, es aquí donde imperceptiblemente para el autor del presente estudio, se pueden cometer errores de estimación de los impactos potenciales, pues una sustancia que no esté incluida en la metodología, no se considera, provocando que se dejen de estimar sus efectos al medio ambiente.

En el contexto del ACV, se tienen varios métodos reportados, cada uno de ellos incluye cierta cantidad de sustancias, las cuales tienen sus propios factores de caracterización calculados para cada una de las categorías de impacto (Heijungs, 1992). En los métodos incluidos en el estudio, se identificaron las categorías equivalentes, es decir, categorías que a pesar de ser de diferentes métodos y tener distintos factores de caracterización, se emplean para evaluar los mismo o similares problemas ambientales.

En el caso del presente trabajo, se partió de la metodología USETox, ya que se ha seleccionado para incluirla en el método de Evaluación de Análisis de Ciclo de Vida ECO-SPEED, el cual diseñado por la línea de investigación de medio ambiente de la Facultad de Química y Farmacia de la Universidad Central Marta Abreu de las Villas, Cuba. En la metodología original de USETox se encontraban la mayor cantidad de sustancias, pero algunas que incluían otros métodos, como por ejemplo Impact 2002 y Eco Indicador 99, no se encontraban dentro de estas sustancias incluidas.

Como se tenían varios posibles pivotes, el problema de la estimación se volvió un problema de incertidumbre, ya que los valores calculados con cada uno de los pivotes diferían para cada sustancia.

En las situaciones como la descrita, donde se presenta la incertidumbre, se deberá hacer uso de técnicas estadísticas que permitan discernir en la práctica cuales serán las estimaciones más probables de los factores de caracterización. Para ello se empleó el Clasificador Bayesiano de Redes Neuronales (Malagón, 2003; Marín, 2005), el cual a partir de los datos de las sustancias pivotes, las que serán utilizadas como ejemplos para entrenar la red, será capaz de calcular los factores de las sustancias no incluidas. De ahí, que la aplicación de redes neuronales como herramienta para trabajar la incertidumbre en un estudio de Análisis de Ciclo de Vida, es una novedad, y en la bibliografía consultada, se pudo constatar que existen disímiles aplicaciones en otras ramas, sin embargo en esta rama de la ciencia no se había dado el caso.

Los resultados de la presente investigación son los factores de caracterización de una gran cantidad de sustancias no presentes en USETox (Rosenbaum y Bachmann, 2008; USETox TEAM, 2010), el cual se convierte en una nueva versión, que los autores de este trabajo le llamaron USETOX plus, ya que incorpora en este caso más

sustancias al análisis que nunca antes. Estos factores fueron listados en un documento de Microsoft Excel e incluidos dentro del método de evaluación ECO-SPEED.

2. MATERIALES Y MÉTODOS

Los métodos analizados en la presente investigación fueron: Eco-Indicador 99 (en su versión igualitaria, individualista y jerárquica), Impact 2002+, EDIP, TRACY y CML. De estos métodos se listaron las categorías de impacto que contienen y su posible vinculación a las categorías del método ECO-SPEED en desarrollo por el autor de este trabajo. Luego de este análisis se constató que las categorías equivalentes en los métodos estaban relacionadas con:

- Toxicidad para los humanos (dentro de la cual se incluyen las clasificaciones de sustancias carcinogénicas y no carcinogénicas, así como su compartimiento en aire, suelo y agua).
- Eco-toxicidad (dentro de la cual se hacen referencia a los compartimientos de agua, suelo y aire)
- Ocupación o uso de la tierra.
- Acidificación y eutrofización.
- Uso de recursos minerales y energéticos.
- Calentamiento global.
- Capa de ozono.

Para ilustrar con un ejemplo la implicación de las no inclusiones en los métodos de Valoración a Ciclo de Vida, supongamos una sustancia X, la cual para su fabricación, emite por cada kilogramo (kg), 1 kg material particulado de hollín de diesel, 1kg de DDT y 1kg de tetracloruro de carbono. Luego se evalúa su impacto utilizando los varios métodos, con el software SIMAPRO 7.1, y se comparan los resultados del aporte de cada una de ellas a la puntuación total del impacto. Los resultados de estas evaluaciones se listaron en el cuadro 1.

Es decir, si se analizan los resultados obtenidos, tres de los métodos darían como el único problema a tener en cuenta, la emisión de DDT; dos de ellos, tendrían como único problema las emisiones de hollín de diesel y tetracloruro de carbono respectivamente y por último, uno de los métodos arrojaría que no existen problemas con las emisiones.

En todos los casos los métodos comparados presentaban distintos factores de caracterización, calculados de formas diferentes, ya que cada uno de los autores utilizó diferentes vías para su cálculo, pero todos seguían el mismo objetivo, poder comparar las sustancias incluidas en cada categoría a través de los factores calculados. Para identificar las inclusiones que tiene cada método se presentan todas las sustancias que contienen entre todos, luego se cuentan la cantidad de sustancias que incluye cada método, como se muestra en el cuadro 2.

CUADRO 1.

Comparación de los resultados obtenidos por los diferentes métodos para la sustancia X. Fuente: Elaboración propia.

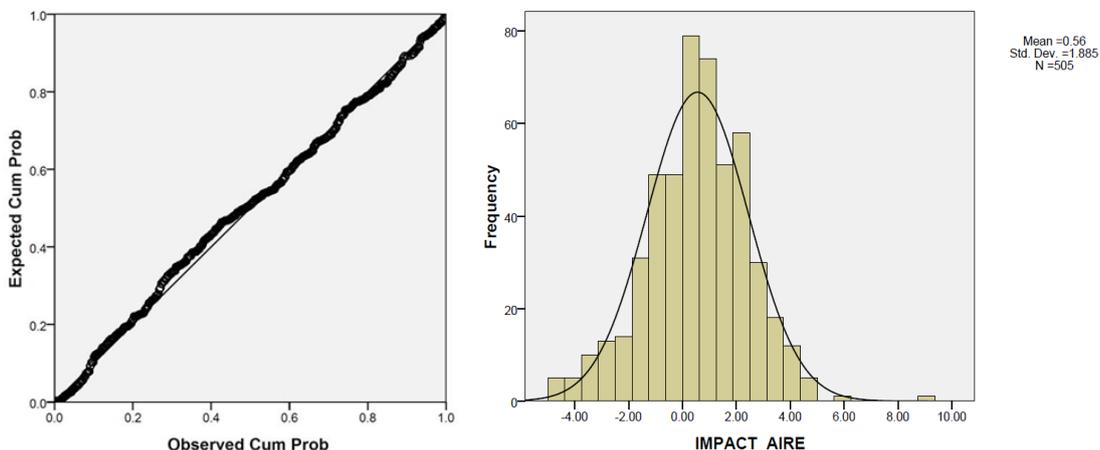
Emisiones	USETox	Eco 99	Impact	EDIP	CML	TRACY
Hollín Diesel	-	18,4	-	-	-	-
DDT	-	-	4,24	-	1,51E-10	1,2E+4
CCl ₄	1,79E-9	-	-	-	-	-

Los cálculos de los factores de caracterización de las sustancias no incluidas, se realizaron utilizando como pivotes algunas de las sustancias incluidas en todos los métodos que se analizaron. De esta forma una sustancia que tiene factores en ambos, como estos factores están dados en unidades comparativas, puede ser utilizada como pivote, para estimar los factores de las demás, sustancias que componen la categoría. Primeramente, para que una sustancia se considere pivote, se analiza si las categorías de impacto tienen comportamiento semejante en cuanto a su función de densidad para los impactos potenciales (Mitchell, 1997). Para ello se emplean las pruebas de hipótesis correspondientes al ajuste de las distribuciones, de estas

CUADRO 2. Totales de sustancias incluidas en cada método para la categoría carcinogénesis en el aire, agua y suelo.
Fuente: Elaboración propia.

Aire	IMPACT	ECO_E	ECO_I	ECO_H	TRACI	USETOX
Cantidad total	991	991	991	991	991	991
Cantidad no incluida	486	931	979	945	844	389
Cantidad incluida	505	60	12	46	147	602
% de inclusion	50.96%	6.05%	1.21%	4.64%	14.83%	60.75%
Agua	IMPACT	ECO_E	ECO_I	ECO_H	TRACI	USETOX
Cantidad total	991	991	991	991	991	991
Cantidad no incluida	492	938	984	952	846	390
Cantidad incluida	991	991	991	991	991	991
% de inclusion	50.35%	5.35%	0.71%	3.94%	14.63%	60.65%
Suelo	IMPACT	ECO_E	ECO_I	ECO_H	TRACI	USETOX
Cantidad total	991	991	991	991	991	991
Cantidad no incluida	490	937	982	951	845	390
Cantidad incluida	991	991	991	991	991	991
% de inclusion	50.55%	5.45%	0.91%	4.04%	14.73%	60.65%

FIGURA 1. Gráficos de probabilidad normal e histograma de frecuencia de categoría carcinogénicos en el aire de la metodología IMPACT 2002.



CUADRO 3. Estadígrafos de prueba Kolmogorov-Smirnov para las categorías de carcinogénicos en aire suelo y agua para IMPACT 2002, y para carcinogénicos en el aire para Eco Indicador 99 (igualitario).

		IMPACT AIRE	IMPACT_AGUA	IMPACT_SUELO	ECO E AIRE
Most Extreme Differences	Absolute	.040	.045	.038	.125
	Positive	.026	.044	.038	.125
	Negative	-.040	-.045	-.032	-.060
Kolmogorov-Smirnov Z		.897	1.006	1.198	.971
Asymp. Sig. (2-tailed)		.398	.264	.114	.302

a. Test distribution is Normal.

pruebas se muestran los gráficos de probabilidad normal e histograma de frecuencia (Figura 1) como ejemplo a los realizados y los estadígrafos de prueba para normalidad de algunas categorías en el cuadro 2.

Esta situación motiva la duda de que si se emplean diferentes métodos, puedan cambiar los resultados de las evaluaciones realizadas y en consecuencia pudieran no obtenerse resultados del todo satisfactorios, los cuales estarán relacionados a la cantidad de sustancias incluidas en el método. A pesar de que el cálculo de los factores de caracterización puede ser realizado para todos los casos de coincidencias de sustancias en categorías equivalentes, solo se han completado las categorías de toxicidad humana, para las sustancias carcinogénicas y no carcinogénicas, en los tres compartimientos de emisiones (aire, agua y suelo).

Es por esta razón que resulta posible calcular los factores de caracterización de una sustancia no incluida en un método, a partir de las equivalencias que tiene con las sustancias pivotes (estimadores), siguiendo los pasos de cálculo descritos en la ecuación (Ec.1).

$$FC_{UseTox}(X_i) = \frac{FC_{Met_j}(X_i)}{FC_{Met_j}(Piv_k)} (FC_{UseTox}(Piv_k)) \quad (Ec.1)$$

Donde:

$FC_{UseTox}(X_i)$: Representa el factor de caracterización para la sustancia X_i para el método USETox.

$FC_{Met_j}(X_i)$: Representa el factor de caracterización para el método "j" que tiene la sustancia "i".

$FC_{Met_j}(Piv_k)$: Representa el factor de caracterización para el método "j" que tiene el pivote "k".

$FC_{USETox}(Piv_k)$: Representa el factor de caracterización para el método USETox que tiene el pivote "k".

De la Ec.1, se puede deducir que para cada sustancia "i" se tendrán tantas estimaciones de su factor de caracterización para USETox, como "k" pivotes existan entre USETox y el método "j".

Una de las consideraciones a tener en cuenta en los casos de incertidumbre es la estabilidad que tengan los diferentes estimadores utilizados, para ellos se utilizan varias pruebas de comparación, la primera se presenta en la figura 2, donde se puede apreciar el comportamiento de los pivotes seleccionados.

Para el trabajo con los datos de las diferentes categorías, se empleó la transformación logarítmica, método muy utilizado para minimizar la varianza de los datos, luego de haber probado la normalidad de los mismos, se transformaron a valores estandarizados de distribución normal, los cuales son recomendables para los estudios comparativos, como lo es el presente.

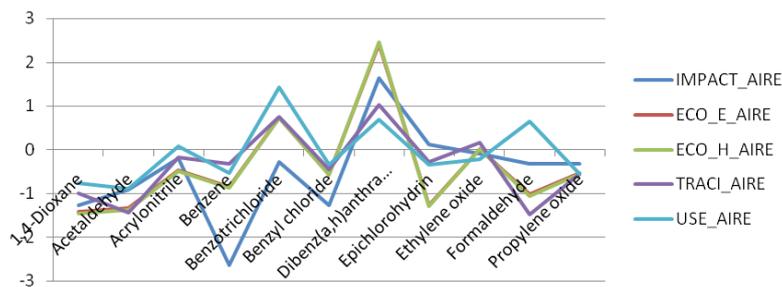


FIGURA 2. Dispersión de los pivotes seleccionados.

3. RESULTADOS Y DISCUSIÓN

Los pivotes a utilizar fueron seleccionados primeramente a partir de la inclusión de una sustancia en los métodos estudiados, resultante de esta selección preliminar queda un conjunto preliminar de 11 posibles, de los cuales se seleccionan 7, teniendo en cuenta las pruebas que se les realizaron de igualdad de medias y rangos, como se muestra en el cuadro 3, existían 7 estimadores dentro de un grupo homogéneo, los que se encuentran representados en la figura 3, utilizando un gráfico de cajas y bigotes, para mostrar la dispersión de los estimadores.

CUADRO 3. Pruebas de Múltiple Rangos para la selección de estimadores. Método: 95.0 porcentaje LSD.

	Casos	Media	Grupos Homogéneos
Acetaldehyde	5	-1.18987	X
_4_Dioxane	5	-1.1814	X
Benzene	5	-1.0421	X
Formaldehyde	5	-0.645192	XX
Benzyl chloride	5	-0.637742	XX
Epichlorohydrin	5	-0.617292	XX
Propylene oxide	5	-0.507019	XX
Acrylonitrile	5	-0.250429	X
Ethylene oxide	5	-0.0246475	XX
Benzotrithloride	5	0.669274	X
Dibenz_a_h_anthracene	5	1.63654	X

Gráfico Caja y Bigotes

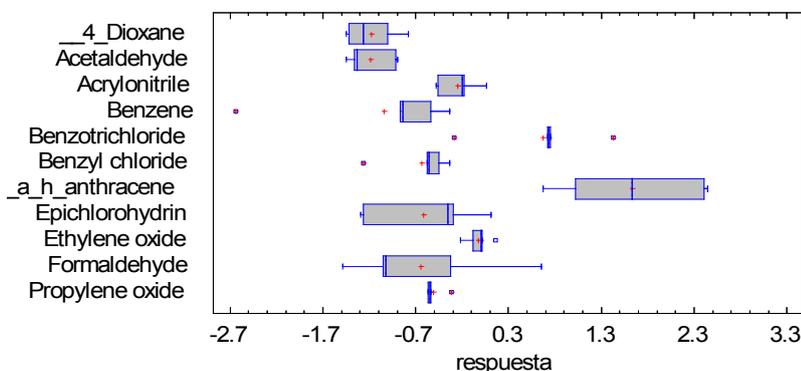


FIGURA 3. Gráfico de Cajas y Bigotes para la dispersión de los estimadores.

Luego de tener seleccionados los estimadores a utilizar se procede a realizar los cálculos con Ec.1, para cada una de las “i” sustancias, se obtendrán $i \cdot k$ estimaciones de su factor de caracterización. Luego estas estimaciones hay que analizarlas y de ellas obtener el factor a utilizar, para este fin, se recurrió al Clasificador Bayesiano de Red Neuronal (PRN o PNN por sus siglas en inglés) (Beamonte y Lacruz, 2009), un método no paramétrico para clasificar observaciones, basándose en “k” variables cuantitativas observadas, que serán los pivotes.

En vez de hacer alguna suposición acerca de la naturaleza de la distribución de las variables dentro de cada grupo, el procedimiento construye un estimador no paramétrico de la

función de densidad de cada grupo en una localización deseada, basada en observaciones de ese grupo. El estimador se construye usando una medida de dispersión (Parzen), que pondera las observaciones de cada grupo de acuerdo a su distancia desde la localización especificada.

Para la clasificación, se asignan las observaciones a los diferentes grupos, basándose en el producto de 3 factores:

1. La *función de densidad* estimada en el vecindario del punto;
2. Las *probabilidades anteriores de pertenecer a cada grupo* (son utilizadas por el método para determinar la probabilidad de pertenencia a un grupo, antes de que los datos sean examinados);

- Los costos de *mal-clasificar* casos que pertenecen a un grupo dado.

La esfera de influencia de la función de ponderación de Parzen, puede ser optimizada reteniendo los valores del grupo de entrenamiento uno a la vez y determinando la distancia al vecino más cercano, basándose en el porcentaje de veces que el punto retenido es correctamente clasificado.

Para la obtención de los factores de clasificación de cada una de las categorías de impacto, se construyó una red bayesiana de clasificación por cada metodología, ya que se completaron los valores de los factores a partir de las coincidencias de los métodos con USETox.

Es decir, para completar los factores de las sustancias de USETox que no se encontraban por ejemplo en Eco 99, se construyó una red con los valores de los factores de las sustancias que coincidían en los dos métodos, los que se utilizaron para obtener una matriz de valores posibles, los cuales fueron clasificados por la red y para seleccionar el más probable. Como USETox tiene 6 categorías de impacto para la

categoría de daño Salud Humana (carcinogénesis y no carcinogénesis, para los compartimientos aire, agua y suelo) y estos valores se querían completar de 5 métodos, el resultado fue el completamiento de estas categorías, a través de la confección de 6 redes de clasificación.

Para la confección de la red se proveyeron probabilidades previas, pero no costos de error los que se tomaron igual para todos los casos. En la figura 4 se representa una Red Neuronal de Clasificación Típica. Como se puede observar la red cuenta con la capa de captura, patrón y sumatoria y salida. La capa de patrón fue entrenada a partir de los valores ya calculados en cada método, los cuales son utilizados como conjunto de entrenamiento, para adaptar la red a las particularidades del caso en cuestión.

Los resultados de utilizar la red neuronal entrenada para clasificar en cada una de las filas de cada combinación de métodos, cual fue el estimador que proporcionaba el valor más probable. Al seleccionar el mejor pivote, se utilizan los resultados obtenidos por el mismo como factores de caracterización para la sustancia clasificada.

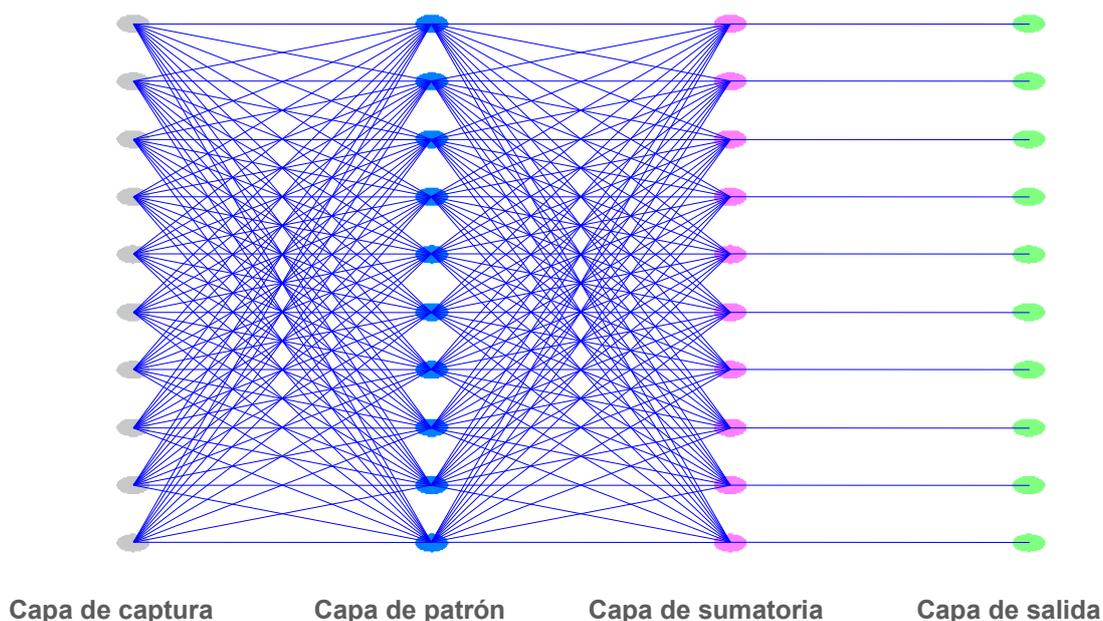


FIGURA 4. Red Bayesiana de Clasificación Típica.

CUADRO 3. Cantidad de estimaciones realizadas para el completamiento de las categorías de impacto de toxicidad humana en los diferentes métodos.

Aire	IMPACT	ECO_E	ECO_I	ECO_H	TRACI	USETOX
Cantidad total	991	991	991	991	991	991
Cantidad Estimada	486	931	979	945	844	389
Cantidad incluida	991	991	991	991	991	991
% de inclusión actual	100 %	100 %	100 %	100 %	100 %	100 %
Agua	IMPACT	ECO_E	ECO_I	ECO_H	TRACI	USETOX
Cantidad total	991	991	991	991	991	991
Cantidad Estimada	492	938	984	952	846	390
Cantidad incluida	991	991	991	991	991	991
% de inclusión actual	100 %	100 %	100 %	100 %	100 %	100 %
Suelo	IMPACT	ECO_E	ECO_I	ECO_H	TRACI	USETOX
Cantidad total	991	991	991	991	991	991
Cantidad Estimada	490	937	982	951	845	390
Cantidad incluida	991	991	991	991	991	991
% de inclusión actual	100 %	100 %	100 %	100 %	100 %	100 %

Con la utilización de los métodos descritos anteriormente se pudieron estimar los valores de los Factores de Clasificación de la mayoría de las sustancias no incluidas en uno y otro método, dando como resultado un completamiento de los métodos utilizados en Análisis de Ciclo de Vida. La búsqueda de factores de clasificación resulta en la actualidad un proceso lento y engorroso, por lo que no siempre se tienen los valores para una gran cantidad de sustancias.

La cantidad de sustancias de las cuales se calcularon sus Factores de clasificación fue el 100 % de las sustancias no incluidas en los métodos, la distribución resultante es representada en el cuadro 3.

4. CONCLUSIONES

Según la cantidad de sustancias incluidas en las categorías de impacto de Carcinogénicos y No Carcinogénicos con respecto a todas las que están al menos en una, se obtuvo que el método USEtox es el más completo, seguido de IMPACT 2002 y TRACI. Las distribuciones de repartimiento de los factores de caracterización de las categorías de impacto, a pesar de tener diferen-

tes parámetros de localización y forma, tienen distribuciones similares, las que en todos los casos, luego de su transformación logarítmica, siguen distribución normal. Las sustancias que se repiten en todos los métodos, denominadas como pivotes, se utilizaron como estimadores de las demás sustancias en cada categoría. La presencia de varias estimaciones para cada valor de los factores de caracterización, puede ser trabajado con la utilización de Redes Neuronales de Clasificación Bayesiana.

5. REFERENCIAS BIBLIOGRÁFICAS

- BEAMONTE, B. A. y B. LACRUZ. 2009. *Redes funcionales, neuronales y bayesianas en problemas de clasificación y predicción*. Universidad de Zaragoza. Zaragoza, España. 264 p.
- BRAKKEE, K.W. y M.A.J. HUIJBREGTS. 2008. Characterisation Factors for Greenhouse Gases at a Midpoint Level. *Int J LCA* 13 (3) 191-201.
- HEIJUNGS, R. 1992. *Environmental Life Cycle of Products*. B&G. Netherlands. 169 p.

- MALAGÓN, L. C. 2003. *Clasificadores bayesianos.El algoritmo Naïve Bayes*. McGraw-Hill. Madrid, España. 146 p.
- MARÍN, I. Á. 2005. *Sistemas Expertos, Redes Bayesianas y sus aplicaciones*. Universidad de Duesto. Madrid, España. 65 p.
- MITCHELL, T. 1997. *Machine Learning*. McGraw-Hill. New York, U.S.A. 284 p.
- REAP, J., F. ROMÁN y B. BRAS. 2008. *A survey of unresolved problems in life cycle assessment*. Int J Life Cycle Assess. Atlanta, U.S.A. 214 p.
- USETox TEAM. 2010. *USETox™ User manual*. January, Int LCA : 5-23.
- ROSENBAUM, R. y T.M. BACHMANN. 2008. *USETox—the UNEP-SETAC. 2008. Toxicity model: recommended characterisation factors for human toxicity and freshwater ecotoxicity in life cycle impact assessment*. Int J Life Cycle Assess (208)13:532–546.