



Análise do intercâmbio de datasets entre diferentes softwares de ACV e o banco de dados nacional SICV Brasil: principais limitações e soluções

Diogo A. Lopes Silva¹, Carla G. Lutosa², Thiago O. Rodrigues^{2 e 3}, Tiago E. N. Braga^{2, 3}

¹Departamento de Engenharia de Produção, Universidade Federal de São Carlos, campus de Sorocaba/SP, diogo.apls@ufscar.br

²Instituto Brasileiro de Informação em Ciência e Tecnologia, Brasília/DF

³Universidade de Brasília, Brasília/DF

Resumo. Para subsidiar a condução de estudos de Avaliação do Ciclo de Vida (ACV) existem diversos bancos de dados públicos e privados, os quais, e em geral, podem ser acessados via o uso de softwares de ACV. Esses softwares dispõem de centenas de datasets em seus bancos de dados (BDs), contudo, os datasets podem apresentar diferentes formatos de saída de dados, sendo que os dois principais formatos mundialmente adotados são o ecoSpold e o ILCD. No caso do Brasil, atualmente encontra-se disponível o Banco Nacional de Inventários do Ciclo de Vida - SICV Brasil. Assim, o presente trabalho tem como objetivo apresentar os resultados de um estudo sobre a análise do intercâmbio de datasets quando importados ao SICV Brasil nos formatos ecoSpold e ILCD. O ILCD é formato padrão para importação e exportação de dados utilizado pelo SICV Brasil, assim, para os testes realizados, os datasets em formato ecoSpold precisaram passar por um processo prévio de conversão antes de sua importação no SICV Brasil. Foi selecionada uma amostra de 10 datasets para o estudo, e os softwares escolhidos foram GaBi, ecoEditor, OpenLCA e SimaPro. Foi observada a possível perda de informações dos datasets armazenados no SICV Brasil comparativamente a versão original dos mesmos quando gerados pelos respectivos softwares de ACV. Um checklist foi desenvolvido para permitir a avaliação da perda das informações de cada dataset. Os resultados mostraram que a perda de informações dos datasets variou de 4% a 43%. Datasets oriundos de softwares que não exportam diretamente no formato ILCD apresentaram maior nível de perda de informações relativamente aos datasets já exportados diretamente no formato ILCD. Assim, foi sugerido que todos os softwares de ACV disponibilizem como opção de importação/exportação de dados o formato ILCD, evitando a necessidade de conversões intermediárias no processo. Também, foi sugerida a utilização de ferramentas computacionais para a verificação do formato ILCD dos datasets encaminhados ao SICV Brasil, o que poderia facilitar o processo de análise da perda de informações. E por fim, foi proposto que seja feita uma análise detalhada para verificar se o formato ILCD adotado tanto pelos softwares de ACV avaliados quanto pelo SICV Brasil são os mesmos.

Palavras-chave. SICV Brasil, banco de dados, inventário do ciclo de vida, metadados, software de ACV.

Introdução

O Inventário do Ciclo de Vida (ICV) é a etapa da metodologia de Avaliação do Ciclo de Vida (ACV) onde ocorre a aferição quantitativa dos aspectos ambientais do ciclo de vida de um produto (ABNT, 2009), sendo que uma das limitações básicas deste processo são o acesso e a disponibilidade de dados de inventário confiáveis e com alta qualidade (UNEP, 2007). Deve-se criar mecanismos que facilitem o acesso às informações de ICV dos mais diversos produtos, já que empresas interessadas em realizarem estudos de ACV necessitarão quase sempre de dados que não estão sob sua alçada (CHERUBINI e RIBEIRO, 2015). Por exemplo, uma empresa pode precisar de dados de ICV da produção das matérias-primas utilizadas na manufatura de seus produtos, os quais são dados normalmente de responsabilidade de fornecedores externos.

Existe uma demanda mundial no tocante ao desenvolvimento de bancos de dados (BDs) para subsidiar estudos de ACV com dados consistentes de acordo com as realidades geopolíticas de cada país. E, no caso do Brasil, recentemente foi lançado o Banco Nacional de Inventários do Ciclo de Vida – SICV Brasil.

Existem BDs públicos e privados, e em geral, eles podem ser acessados por meio de softwares de ACV. Cada *dataset* disponível em um BD representa um conjunto de dados envolvendo todas as informações de caráter qualitativo (ou metadados) e quantitativo do ciclo de vida de um produto. Os *datasets* podem apresentar diferentes formatos de saída de dados, sendo que os dois principais adotados são o ecoSpold e o ILCD. Esses

formatos foram desenvolvidos como padrões para o armazenamento virtual de *datasets* nos BDs. No caso do SICV Brasil, o formato padrão adotado é o ILCD, e a ideia é que o banco brasileiro seja acessível via *web* e *webservices*, e que os *datasets* existentes ali sejam provenientes de softwares de ACV.

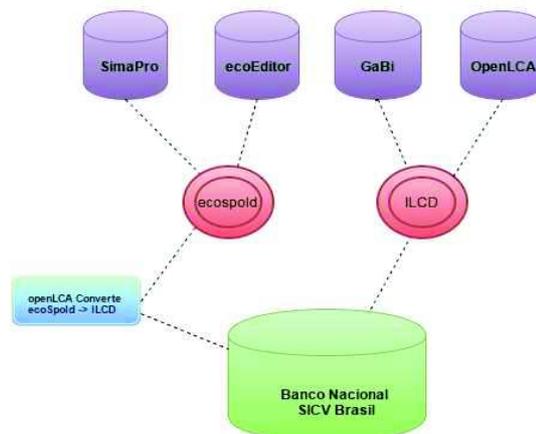
Assim, este artigo objetiva analisar o nível de perda de informações de *datasets* em processos de intercâmbio de dados entre softwares de ACV e o SICV Brasil.

Metodologia

Primeiramente foi planejado um conjunto de testes de intercâmbio de *datasets*, conforme esquematicamente ilustrado na Figura 1. Para isso, foram selecionados diferentes softwares de ACV, sendo: **ecoEditor**, **GaBi**, **openLCA**, e **SimaPro**. A escolha destes softwares se justificou pelo fato deles estarem entre os principais utilizados no Brasil e no mundo. Além disso, os responsáveis pelo desenvolvimento destes *softwares* participam de iniciativas internacionais ligadas a ACV, tais como o *Life Cycle Initiative* e a *Global LCA Data Access*, o que indica que estarão alinhados com os futuros desenvolvimentos do SICV Brasil, também participante de tais iniciativas. Também, se destaca o fato das ferramentas openLCA e do ecoEditor serem de acesso livre, enquanto que o GaBi e o SimaPro serem softwares fechados.

- **ecoEditor**: desenvolvido para a geração de *datasets* visando uma posterior publicação no BD ecoinvent. O software permite salvar *datasets* em formatos ecoSpold v1 e v2. Neste trabalho, foi utilizada a versão 3.3.65.10842 do software, e foram gerados arquivos ecoSpold v1 para posterior conversão para formato ILCD;
- **GaBi**: foi utilizado o software versão 6.4.0.3 e o seu BD versão profissional 6.107 como referência. Esta ferramenta possibilita a geração de *datasets* em formatos ecoSpold v1 e ILCD;
- **OpenLCA**: foi utilizada sua versão 1.4.0 e o BD ELCD v.3 como referência. Esta ferramenta possibilita a geração de *datasets* em formatos ecoSpold v1 e v2, e ILCD. Além disso, o OpenLCA também foi utilizado como conversor dos *datasets* em formato ecoSpold para ILCD (ver Figura 1);
- **SimaPro**: foi utilizada a versão 7.3.3 do software e o BD Industry Data 2.0, PlasticsEurope. Esta ferramenta possibilita a geração de *datasets* em formato ecoSpold v1.

Figura 1: Representação esquemática dos testes delineados para o intercâmbio de *datasets*



Na Figura 1, todos os *datasets* formato ecoSpold foram gerados com base na versão v1. Além disso, é importante destacar que os softwares analisados também possibilitam a importação de arquivos em formatos diferentes do ecoSpold e do ILCD. Por exemplo, o SimaPro possibilita também a geração de arquivos em

formato Excel, e o GaBi, permite a geração de arquivos em formato GPR. Porém, esses e os demais formatos pré-disponíveis nos softwares analisados se mostraram fora do escopo do presente estudo.

Ao todo foram selecionados 10 *datasets* (ver Tabela 1), sendo que cada um foi importado ao SICV Brasil e a perda de informações durante o intercâmbio de dados foi avaliada conforme o *checklist* da Tabela 2.

O *checklist* foi desenvolvido para servir de referência para a organização, mapeamento do fluxo de informações e padronização dos resultados da análise do processo de intercâmbio de dados. Ao todo, o *checklist* apresenta 38 itens distribuídos pelas categorias: “*Process information*”, “*Modelling and validation*”, “*Administrative information*”, “*Inputs/Outputs*”, e “*Outros*”. Tais categorias foram organizadas a partir dos requisitos do formato ILCD, conforme consta nos documentos EC/JRC/IES (2010; 2011).

Por fim, é importante mencionar que o *checklist* da Tabela 2 não apresenta todos os 38 itens estabelecidos devido a limitações de espaço no texto, mas é mostrada uma síntese com os seus principais tópicos e conteúdo.

Tabela 1: Amostra de *datasets* selecionados para o estudo

Descrição do <i>dataset</i>	Software de ACV	Banco de dados	Formato original do <i>dataset</i>
GLO: Diesel Production, low sulfur	ecoEditor	ecoinvent version 3	ecoSpold v1
EU-27: Landfill of plastic waste	GaBi	GaBi Professional DB 6.107	ecoSpold v1 e ILCD
EU-25: Corrugated board boxes	OpenLCA	ELCD v.3	ILCD
GLO: PVC pipe E	SimaPro	Industry Data 2.0, Plastic Europe	ecoSpold v1
CH: Cement (average) PE	GaBi	GaBi Professional DB 6.107	EcoSpold v1 e ILCD
GLO: Natural gas	SimaPro	Industry Data	ecoSpold v1
GLO: Planks	SimaPro	Industry Data	ecoSpold v1
RER: Acetone dimethylcetone	Gabi	GaBi Professional DB 6.107	ILCD
GLO: Excavator PE	Gabi	GaBi Professional DB 6.107	EcoSpold v1
CN: Elevator average utilization	Gabi	GaBi Professional DB 6.107	EcoSpold v1

Tabela 2: *Checklist* utilizado para a avaliação da perda de informações durante o intercâmbio de dados

TÓPICOS	ITEM	Critério de avaliação	Notas			Software de ACV e BD escolhido		Nome original do <i>dataset</i>
			NÃO ATENDE	ATENDE PARCIALMENTE	OK	De formato	Para formato	Responsável
<i>Process information</i>		<p>NÃO ATENDE: O item não foi identificado.</p> <p>ATENDE PARCIALMENTE: O item não foi totalmente atendido.</p> <p>OK: O item foi atendido satisfatoriamente.</p>						
		<i>Key data set information</i>				Comentários		
	1	<i>Location</i>						
	2	<i>Geographical representativeness description</i>						
	3	<i>Reference year</i>						
						
		<i>Quantitative reference</i>						
8	<i>Reference flow(s)</i>							



	<i>Time representativeness</i>					
	9	<i>Data set valid until:</i>				
	10	<i>Time representativeness description</i>				
				
TOTAL DE NOTAS NO TÓPICO						
<i>Modelling and validation</i>	<i>LCI method and allocation</i>					
	13	<i>Type of data set</i>				
	14	<i>LCI method principle</i>				
				
	<i>Data sources, treatment, and representativeness</i>					
	18	<i>Data cut-off and completeness principles</i>				
				
	22	<i>Data collection period</i>				
	<i>Completeness</i>					
23	<i>Completeness product model</i>					
TOTAL DE NOTAS NO TÓPICO						
<i>Administrative information</i>	<i>Commissioner and goal</i>					
				
	25	<i>Data set generator / modeler</i>				
	<i>Data entry by</i>					
	26	<i>Time stamp (last saved)</i>				
	27	<i>Data set format(s) (source data set)</i>				
	<i>Publication and ownership</i>					
				
33	<i>Access and use restrictions</i>					
TOTAL DE NOTAS NO TÓPICO						
<i>Inputs/outputs</i>	<i>Inputs</i>					
	34	<i>Todos os campos existentes no SICV Brasil foram preenchidos corretamente (type of flow, classification, flow, variable, etc.) ?</i>				
	<i>Outputs</i>					
	35	<i>Todos os campos existentes no SICV Brasil foram preenchidos corretamente (type of flow, classification, flow, variable, etc.) ?</i>				
				
TOTAL DE NOTAS NO TÓPICO						
TOTAL GERAL DE NOTAS						

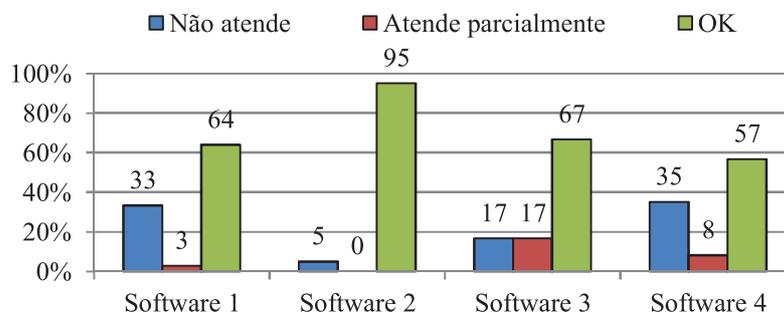
Para efeito da análise da perda de informações dos *datasets*, ao final da avaliação de cada tópico do *checklist*, foi realizada a contagem dos itens atendidos, parcialmente atendidos e não atendidos. Quanto maior a quantidade de itens enquadrados nos critérios “**não atende**” e/ou “**atende parcialmente**”, maior foi o nível de perda de

informações do *dataset* importado ao SICV Brasil. O nível de 100% (sem perda de informações) foi utilizado como referência para o *dataset* que viesse a atender a todos os 38 itens do *checklist*.

Resultados e Discussão

Na Figura 2 é mostrado um gráfico com os resultados gerais da análise de perda de informações dos 10 *datasets*, incluindo o percentual de itens atendidos, parcialmente atendidos e não atendidos. Para preservar a marca dos softwares de ACV avaliados, seus nomes oficiais foram substituídos por Software 1, 2, 3 e 4.

Figura 2: Nível da perda de informações dos *datasets* importados para o SICV Brasil.



Pela Figura 2, percebe-se que os *datasets* apresentaram um nível de perda de informações variando de 5 a 43% (ou 57 a 95% de itens atendidos), sendo que os *datasets* mais completos foram obtidos quando o arquivo original foi gerado a partir do software 2. A maior perda de informações foi obtida ao se utilizar o software 4. Já o software 1 e software 3 apresentaram performances similares.

Sobre os itens não atendidos ou apenas parcialmente atendidos pelos *datasets* analisados, se ressalta principalmente a perda de informações para as categorias “*Process information*” e “*Modelling and validation*”. Assim, a seguir, são discutidos os principais destaques durante a análise de intercâmbio dos *datasets*.

Considerando o Software 1, o item que **não foi atendido** no tópico “*Process information*” foi: “*use advice for data set*”, pois este item não existe em nenhum dos campos do software. Já os itens que não foram atendidos no tópico “*Modelling and validation*” foram: “*modelling constants*”, “*data cut-off and completeness principles*”, “*data selection and combination principles*”, “*deviation from data selection and combination principles / explanations*”, “*data source(s) used for this data set (source data set)*”, e “*completeness product model*”; e no tópico “*Administrative information*”, os itens “*intended applications*”, “*permanent data set URF*”, e “*workflow and publication status*” não foram atendidos, pois todos estes itens citados não existem no software. No tópico “*Inputs/outputs*”, os seguintes itens não foram atendidos: “*data source type*” e “*data derivation type/status*”, pois tais itens estavam em branco no *dataset* importado para o SICV Brasil, ainda que na sua versão original do Software 1, os mesmos estivessem todos preenchidos. Portanto, isso caracteriza perda de informações no processo de intercâmbio, e isso pode ser explicado pelo fato dos formatos ecoSpold e ILCDC não serem totalmente compatíveis quanto aos itens que cada formato tem como requisito durante a construção de *datasets*.

Sobre o Software 2, a grande maioria dos itens do *checklist* foram atendidos ao se importar os *datasets* em formato ILCDC para o SICV Brasil. A exceção ocorreu apenas para o item “*permanent dataset URF*”, localizado no tópico “*Administrative information*”, pois este item não existe no software. A baixa perda de informações



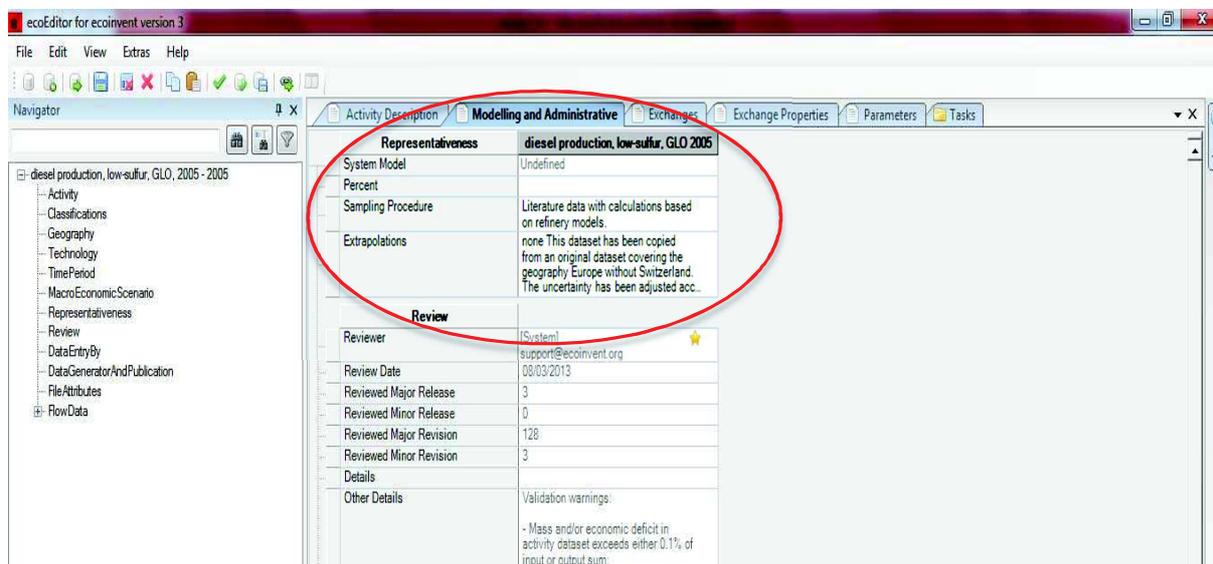
dos *datasets* importados a partir do Software 2 indica que, provavelmente, tanto o SICV Brasil quanto o Software 2 utilizam a mesma plataforma para implementação do formato ILCD.

Sobre a avaliação dos *datasets* oriundos do Software 3, os itens que não foram atendidos no tópico “**Process information**” foram: “*use advice for data set*”, pois este item não existe em nenhum dos campos do software; e “*technology description including background system*”, pois não é possível incluir o diagrama de fluxo de processo, o qual representa um esquema gráfico do sistema de produto. Outros softwares como o Software 1 e Software 2 permitem a inserção de imagens, ou de esquemas gráficos, porém, essa é uma limitação no Software 3. No tópico “**Modelling and validation**” não foram atendidos os campos: “*deviation from LCI method principle/explanations*”, “*deviation from data selection and combination principles/explanations*”, e “*Completeness product model*”; no tópico “**Administrative information**”, os itens “*permanent data set URI*”, e “*workflow and publication status*” também não foram atendidos; e no tópico “**Inputs/outputs**”, os itens “*data source type*” e “*data derivation type/status*” também não foram atendidos, pois todos estes itens citados não existem no software. Portanto, verificaram-se limitações nos *datasets* em formato ILCD gerados a partir Software 3, e isso pode estar relacionado as diferenças existentes na forma como o Software 3 implementa o formato ILCD comparativamente ao SICV Brasil. Também, como destacado anteriormente, nem todos os itens contemplados pelos *datasets* em formato ecoSpold estão presentes no formato ILCD, e vice-versa, logo, perda de informações ocorrerão durante a conversão ecospold para ILCD.

Em relação ao Software 4, foi verificado no tópico “**Process information**”, que os seguintes itens não foram atendidos “*use advice for data set*”, e “*general comment on dataset*”; no tópico “**Modelling and validation**”, não foram atendidos os campos “*LCI method approaches*”, “*modelling constants*”, “*data cut-off and completeness principles*”, “*data selection and combination principles*”, “*deviation from data selection and combination principles/explanations*”, “*data treatment and extrapolations principles*”, e “*completeness product model*”; enquanto que no tópico “**Administrative information**”, os itens “*intended applications*”, “*permanent data set URI*”, e “*workflow and publication status*” não foram atendidos; e por fim, no tópico “**Inputs/outputs**”, os itens “*data source type*” e “*data derivation type/status*” não foram atendidos, pois todos estes itens citados não estão disponíveis no Software 4.

Como foi visto, a maior parte dos itens não atendidos durante a análise se deu pela não existência dos mesmos nos formatos originais dos *datasets*, previamente ao processo de importação para o SICV Brasil. A maior parte dos itens não contemplados estão associados aos metadados, ou seja, a descrição das fontes de dados, do escopo geográfico, tecnológico, e temporal dos *datasets*, e informações administrativas. Todavia, a partir do exemplo da Figura 3, verifica-se que a não existência dos itens destacados não implica que os mesmos, necessariamente, não tenham sido contemplados de alguma forma por cada *dataset* em seu formato original.

Figura 3: Itens destacados no software ecoEditor para o tópico “Representativeness”.



Por exemplo, para a categoria “**Modelling and validation**” do *checklist*, itens como “*data cut-off and completeness principles*”, e “*data source(s) used for this data set (source data set)*” podem ser contemplados de formas diferentes segundo o formato ecoSpold do software ecoEditor. Assim, a Figura 3 mostra o caso do *dataset* intitulado “*GLO: Diesel Production, low sulfur*”, e nele, tais itens do *checklist* foram incluídos no *dataset* a partir do preenchimento dos itens do tópico “*Representativeness*”. Contudo, esses itens da Figura 3 não existem no *checklist* da Tabela 2, e por isso que quando ocorreu a conversão do *dataset* para formato ILCD que ocorreram essas perdas de informações no processo de intercâmbio. Além disso, é importante destacar que esta mesma problemática também foi verificada para os demais softwares de ACV estudados.

Conclusões

Foi realizada a análise do intercâmbio de *datasets* importados para o SICV Brasil a partir de alguns dos mais representativos softwares de ACV existentes, e os resultados mostraram que entre todos os softwares avaliados, aquele que apresentou o menor nível de perda de informações foram os *datasets* gerados a partir do Software 2. Os *datasets* importados gerados a partir do Software 4, foram aqueles que apresentaram os maiores níveis de perda de informações, enquanto que os *datasets* gerados a partir do software 1 e do software 3, apresentaram resultados similares e intermediários de perda de informações.

Foi destacado que a maior parte das informações perdidas durante o processo de intercâmbio de dados foi relacionada aos metadados, e isto resulta em limitações de conteúdo especialmente no escopo dos *datasets* convertidos, podendo inclusive comprometer o uso desses *datasets* em aplicações práticas, como em estudos de ACV realizados pela indústria e academia.

Como a maior parte dos itens não atendidos durante a análise de perda de informações dos *datasets* se deu devido a não existência dos mesmos no formato ILCD utilizado pelo SICV Brasil, como ação, se sugere realizar uma comparação entre as formas de implementação do formato ILCD adotados tanto pelo SICV Brasil quanto pelos softwares de ACV. Assim, será possível averiguar com mais detalhes se as diferenças nos resultados apresentam, de fato, dependência da forma como o formato ILCD foi implementado.

Outra sugestão seria a utilização de uma ferramenta para verificação automática do formato ILCD dos *datasets* previamente a importação definitiva deles ao SICV Brasil. Tal ferramenta poderia ser o ILCD Editor, disponível



V Congresso Brasileiro em Gestão do Ciclo de Vida

19 a 22 de Setembro 2016 - Fortaleza, CE

em EU/JRC/EPLCA (2016). Assim, tal verificação poderia ser feita pelos gestores do SICV Brasil, e no caso de haver inconsistências no *dataset* gerado, o usuário (*data provider*) poderia ser informado sobre as informações faltantes/incompletas no *dataset* submetido e sobre as necessidades de melhoria no mesmo.

Por fim, seria também importante que todos os softwares de ACV apresentassem a possibilidade de exportar diretamente os *datasets* para formato ILCD, sem a necessidade de realizar conversões intermediárias (e.g., de ecoSpold para ILCD), pois isso auxiliaria na redução de parte do problema de perda de informações durante o intercâmbio de dados.

Agradecimentos

Os autores agradecem ao apoio financeiro fornecido pelo “Conselho Nacional de Desenvolvimento Científico e Tecnológico” – CNPq, processo N° 170101/2015-0. Diogo A. Lopes Silva também agradece ao Instituto Brasileiro de Informação em Ciência e Tecnologia – IBICT pelo suporte técnico durante à pesquisa conduzida, e à “Fundação de Amparo à Pesquisa do Estado de São Paulo” – FAPESP, processo N° 2013/06736-9.

Referências Bibliográficas

ASSOCIAÇÃO BRASILEIRA DE NORMAS TÉCNICAS. (2009). NBR ISO 14044: Gestão ambiental – avaliação do ciclo de vida – Requisitos e orientações. Rio de Janeiro

CHERUBINI, E.; RIBEIRO, P. T. (2015). Diálogos Setoriais Brasil e União Europeia: desafios e soluções para o fortalecimento da ACV no Brasil. Instituto Brasileiro de Informação em Ciência e Tecnologia, Brasília: 2015. ISBN: 9788570131058.

EUROPEAN COMMISSION – JOINT RESEARCH COUNCIL. INTERNATIONAL – INSTITUTE FOR ENVIRONMENT SUSTAINABILITY. (2010). International reference life cycle data system (ILCD) Handbook – nomenclature and other conventions. Luxembourg: European Commission – Joint Research Centre – Institute for Environmental and Sustainability.

_____. (2011). International reference life cycle data system (ILCD) – Documentation of LCA datasets. Version 1.1 Beta. Luxembourg: European Commission – Joint Research Centre – Institute for Environmental and Sustainability.

EUROPEAN COMMISSION – JOINT RESEARCH COUNCIL. INTERNATIONAL – EUROPEAN PLATFORM ON LIFE CYCLE ASSESSMENT. (2016). Life cycle inventory datasets related tools. Disponível em: http://eplca.jrc.ec.europa.eu/?page_id=140. Acesso em 11 Abr 2016.

UNITED NATIONS ENVIRONMENT PROGRAMME. (2007). Life cycle management - a business guide to sustainability, Paris: UNEP/SETAC Life Cycle Initiative.