

**SPECIFICATION INITIALE DES BESOINS DE QUALITE EN ACV :  
ÉTAT DE L'ART, RETOUR D'EXPERIENCES ET RECOMMANDATIONS**

**Guide à l'attention des praticiens et commanditaires d'ACV -**

Avril 2015

**Responsables scientifiques :**

**- Isabelle DESCOS, Bernard DE CAEVEL**  
RDC Environment - 57, avenue Gustave Demey  
1160 Bruxelles



Avec la participation de Maki consulting

L'association SCORE LCA est une structure d'étude et de recherche dédiée aux travaux relatifs à l'Analyse du Cycle de Vie (ACV) et à la quantification environnementale. Elle vise à promouvoir et à organiser la collaboration entre entreprises, institutionnels et scientifiques afin de favoriser une évolution partagée et reconnue, aux niveaux européen et international, de la méthode d'Analyse du Cycle de Vie et de sa mise en pratique.

- ✓ Ces travaux ont reçu le soutien de l'ADEME (Agence de l'Environnement et de la Maîtrise de l'Energie) [www.ademe.fr](http://www.ademe.fr)
- ✓ Les points de vue et recommandations exprimés dans ce document n'engagent que les auteurs et ne traduisent pas nécessairement, sauf mention contraire, l'opinion de l'ensemble des membres de SCORE LCA.
- ✓ Les informations et les conclusions présentées dans le présent document ont été établies au vu des données scientifiques et techniques et d'un cadre réglementaire et normatif en vigueur à la date de l'édition des documents.

**Comité de suivi pour SCORE LCA** : Jean-paul Cazalets – Total, Daniel Dunet – Veolia R&I, Jade Garcia – SCORE LCA, Denis Le Boulch – EDF, Stéphane Morel – Renault, Philippe Osset – SCORE LCA, Anne Prieur – GDF SUEZ, Olivier Réthoré – ADEME, Christèle Wojewodka – Saint Gobain, Lauro Cimolino – Socotec.

## Table des matières

Executive Summary .....	6
Résumé .....	8
Lexique .....	11
I. Introduction.....	13
II. La qualité en ACV -fondements .....	14
II.1. Définition .....	14
II.2. Champ de la qualité.....	14
II.3. Enjeux de la qualité .....	15
II.4. Une étape dans le processus itératif d'ACV .....	16
II.5. Les critères de qualité.....	17
II.5.1. Deux grandes notions : précision et justesse de la représentation .....	17
II.5.2. Les critères de qualité d'une donnée ACV .....	19
III. La qualité initiale dans les textes de référence .....	20
IV. Les étapes à suivre pour définir les exigences de la qualité initiale.....	22
IV.1. Approche générale .....	22
IV.1.1. L'identification de l'incertitude maximale acceptable .....	23
IV.1.2. La limitation de l'incertitude .....	23
IV.2. Les étapes à suivre.....	25
IV.2.1. Les étapes à suivre pour la détermination de l'incertitude maximale .....	26
IV.2.2. Les étapes à suivre pour Déterminer les moyens de réduire l'incertitude .....	34
IV.3. Synthèse des recommandations .....	36
V. Supports et outils disponibles pour bien définir la qualité.....	38
V.1. Outils pouvant servir de supports à la définition de la qualité des données .....	38
V.1.1. Recherche dans les bases de données (Bdd).....	38
V.1.1. La conduite d'une « screening » ACV ou d'une première iteration .....	39
V.1.2. L'analyse des contributions.....	39
V.1.3. L'analyse bibliographique ACV .....	39
V.1.4. Analyse de documents sectoriels .....	41
V.1.5. La recherche d'expertise .....	41
V.2. Fonctionnalités des logiciels ACV permettant une aide à la définition de la qualité initiale.....	42
V.2.1. Visualisation des informations des bases de données .....	42
V.2.2. Analyse des contributions.....	43
V.2.3. Qualification de la qualité des données pour sa propre étude .....	45

V.2.4. Visualisation des parts de données de bonne, moyenne et mauvaise qualité	46
V.2.5. Analyse de l'incertitude .....	48
VI. Conclusions, Limites et Perspectives .....	49
VI.1. La spécification initiale des besoins en qualité : identifier où mettre les efforts dans une démarche ACV.....	49
VI.2. La spécification initiale des besoins en qualité : un besoin d'expertise .....	49
VI.3. La recherche d'objectivité pousse à essayer d'obtenir un score de qualité, pas toujours pertinent.....	49
VI.4. Une trop grande attention portée à l'incertitude par rapport à l'erreur de modélisation .....	50
VI.5. Le manque d'interopérabilité entre les bases de données .....	50
VI.6. L'apport possible de la revue critique dans l'évaluation initiale de la qualité .....	51
VII. Annexes .....	52
VII.1. Annexe 1 : définition des principaux critères de qualité .....	52
VII.1.1. Représentativité.....	52
VII.1.2. Complétude.....	56
VII.1.3. Incertitude .....	58
VII.2. Annexe 2 : ce que disent les textes de références sur la qualité initiale des données et la qualité.....	59
VII.2.1. Les normes ISO 14040-44.....	59
VII.2.2. L'ILCD Handbook .....	63
VII.2.3. La PEF Methodology (2014) .....	67
VII.3. Annexe 3 : Les informations des bases de données sur la qualité .....	74
VII.3.1. Description de la qualité dans les bases de données .....	74
VII.4. Annexe 4 : Informations utiles dans les bases de données pour la spécification initiale de la qualité .....	90
VII.5. Annexe 5 : Liste des personnes interviewées sur la qualité initiale des données	96

## Table des figures

Figure 1 : Illustration de l'approche itérative pour la qualité des données .....	16
Figure 2 : Expression des deux grandes composantes de la qualité (précision intrinsèque et précision de modélisation).....	19
Figure 3 : Les deux grandes étapes pour définir la qualité initiale des données.....	22
Figure 4 : Chronologie des tâches à effectuer par le commanditaire et l'auteur pour déterminer l'incertitude maximale acceptable.....	26
Figure 5 : Etapes à suivre pour déterminer les moyens de réduire l'incertitude.....	34
Figure 6 : Schéma des critères clés pour sélectionner les affinages d'incertitudes .....	36
Figure 7 : Logigramme pour accompagner les praticiens en matière de qualité initiale des données ACV .....	37
Figure 8 : Exemple de l'ancienne fonctionnalité de SimpaPro permettant la visualisation des informations des bases de données.....	43
Figure 9 : Diagramme avec différentes flèches et graduations pour montrer les contributions des procédés .....	44
Figure 10: Restitution de la « Weak point analysis » (GaBi) permettant de visualiser les contributions par procédé et par flux élémentaire .....	45
Figure 11 : Fonctionnalité SimaPro visant à renseigner la Pedigree matrix en qualifiant soi-même ses données pour une étude spécifique .....	46
Figure 12 : Visualisation des pourcentages de données de bonne ou moins bonne qualité par indicateur de catégorie d'impact (GaBi) .....	47
Figure 13 : Illustration de la qualité .....	52
Figure 14 : Illustration de la complétude .....	56
Figure 15 : Exemple d'analyse de la qualité pour une étude d'ACV screening selon la méthode PEF- Etape d'identification des données couvrant au moins 90 % des contributions aux catégories d'impacts.....	70
Figure 16 : Exemple du traitement de l'amélioration de la qualité pour une étude d'ACV screening selon la méthode PEF- Etape de recherche de données meilleures pour les mêmes procédés ou de procédés de meilleure qualité .....	71
Figure 17: Formule de calcul de l'incertitude globale dans ecoinvent .....	85
Figure 18 : Exemple de champs pertinents à consulter pour la base de données ELCD ..	91
Figure 19 : Exemple de visualisation du niveau de qualité pour un inventaire donné dans GaBi .....	93
Figure 20 : Exemple de qualité des données disponible au niveau d'un point de données (flux élémentaire).....	94

## Table des tableaux

Tableau 1 : Récapitulatif des critères de qualité couverts par les documents de référence et modes d'évaluation recommandés .....	20
Tableau 2 : Proposition de RDC Environment sur les types ordre de grandeurs d'incertitudes acceptables selon les types d'études et les systèmes étudiés .....	33
Tableau 3 : Critères de qualité défini par la Méthodologie PEF .....	67
Tableau 4 : Calcul de la qualité globale (méthode semi-quantitative) utilisé par la méthodologie PEF.....	68
Tableau 5 : Classification des notes de qualité et appréciation selon la méthodologie PEF .....	69
Tableau 6 : Niveau d'exigence sur la qualité des données en fonction de la contribution des procédés pour une étude complète .....	69
Tableau 7 : Synthèse des exigences de qualité des données de la méthodologie PEF selon les types d'études.....	73
Tableau 8 : Grille de notation pour différents critères de l'ILCD Handbook .....	75
Tableau 9 : Grille de notation globale de la qualité.....	76
Tableau 10 : Grille de notation pour différents critères dans GaBi .....	79
Tableau 11 : Matrice pedigree d'ecoinvent .....	83
Tableau 12 : Valeurs par défaut pour les incertitudes de différents flux dans ecoinvent (sur base du jugement d'expert) .....	84
Tableau 13 : Correspondance entre les incertitudes liées aux différents critères et les notes de la matrice pedigree dans ecoinvent.....	85
Tableau 14 : Présentation des méthodes d'évaluation de la qualité dans la base de données ELCD.....	92

## Executive Summary

In LCA, data quality management is crucial at each step. At the beginning of a study it helps to estimate the required time and budget. It further allows to be more focused on the main contributing steps and to minimize the effort and the time of increasing quality of non-essential data.

Quality assessment tools are often designed for a post inventory evaluation and are fundamentally incomplete (only the intrinsic precision of the data is taken into account, not the appropriateness to the goal and scope).

SCORELCA asked RDC Environment to write a guide to help LCA practitioners to better assess initial requirements of data quality.

This guide is divided into 4 parts:

### 1. The definition of quality in LCA

This part reminds the two main pillars for quality:

- Precision on raw and calculated data: generally assessed by the generator of the data and can be easily integrated into a database
- Appropriateness, i.e. consistency between the real system to be represented and the LCA model (accuracy): specific to each LCA study and has to be assessed by the practitioner

### 2. A summary on the way the reference texts deal with the specification of data quality at the beginning of a LCA study

### 3. The steps to follow in order to define the initial specification

As the goal of a LCA is to draw conclusions related to the comparison of environmental impacts of systems or auto-references, it is important to get a lower uncertainty than the difference between the systems. So the initial specification of data quality has to determine the quality level allowing to sufficiently reduce the uncertainty.

Two steps are therefore necessary:

- The **identification of the maximum acceptable uncertainty**: it is either the expressed requirement of the commissioner or derived from the goal of the study by the expert
- The **determination on how to limit this uncertainty**: This stage aims at minimizing efforts for data collection. The confidentiality or the willingness of the commissioner to collect data can favour or penalize this reduction of uncertainty. The reduction will focus on data that have high contributions to the results, are uncertain and are easily accessible.

Depending on the available time for this step two procedures are proposed: a light process and a full process. For the light process, the focus will be on the reduction of uncertainty through a better data representativeness and a good modelling of the hotspots.

#### **4. Tools that can help**

- LCA database search to assess if the data quality is good enough or not
- Screening LCA to point out the hotspots of the LCA results
- Contribution analysis that will show both processes and flows that contribute the most to the environmental balance
- LCA literature analysis to point out main steps of the life cycle and key parameters
- Sectorial documents analysis to identify the kind of technologies, their representativeness...and by this way the key points analysis that will influence the environmental balance
- Sectorial and LCA expertise to define some processes, provide main trends, define some uncertainty range for some data
- Use of functionalities of LCA softwares, which provide some information on data quality and allow performing a contribution analysis.

This guide was tested by SCORELCA members.

This guide perspectives are:

- To provide a structured procedure that helps the practitioner with the initial definition of data quality
- To show that an external expertise can be used but that many of these steps rely as well on common sense
- To remind that up to now, nothing special is asked to be documented in LCA reports on this first step of initial specification of data quality
- To present some ways of improvements to better use data quality: interoperability of databases (that would avoid to use only one database) and the development of critical reviews at each step of LCA that would help to validate this first step of data quality assessment.

## Résumé

En ACV, la gestion de la qualité des données est cruciale à chaque étape. La spécification initiale de cette qualité des données permet d'estimer au mieux le budget et le temps nécessaire à la réalisation d'une étude. Elle permet également de se focaliser sur les étapes les plus contributrices pour réduire l'effort et le temps de recherche d'amélioration de la qualité pour des données non essentielles.

Les outils d'évaluation de la qualité existant sont généralement mis en place pour une évaluation au stade de l'inventaire de cycle de vue et sont de ce fait en général incomplets (seule la précision de l'étude est prise en compte mais pas le caractère approprié des données au but et champ de l'étude).

SCORELCA a demandé à RDC Environment d'écrire un guide pour aider les praticiens ACV à mieux évaluer les exigences initiales de la qualité des données.

Le guide est divisé en 4 parties :

### **1. La définition de la qualité en ACV**

Les 2 piliers de la qualité sont rappelés :

- La précision intrinsèque des données brutes et calculées, qui est généralement évaluée par le générateur de la donnée et peut être aisément intégrée dans une base de données
- Le caractère approprié de la donnée entre le système réel à modéliser et le modèle ACV : cette notion est spécifique à chaque étude ACV et doit être évaluée par le praticien

### **2. Un résumé de la façon dont les textes de référence traitent de la définition initiale de la qualité des données**

### **3. Les étapes à suivre pour définir la qualité initiale des données.**

Le but d'une ACV est de dresser des conclusions relatives à une comparaison (entre systèmes ou via un autoréférencement), il faut donc que les résultats soient discriminants c'est-à-dire que l'incertitude soit plus faible que la différence de résultats entre les systèmes étudiés.

La spécification initiale de la qualité des données doit donc déterminer le niveau de qualité qui permette de réduire suffisamment l'incertitude.

Deux étapes sont donc nécessaires :

- L'identification de l'incertitude maximale acceptable : elle peut soit être la volonté directement exprimée du commanditaire de l'étude, soit résulter

indirectement du but de l'étude, étant évaluée sur base d'une expertise ACV

- La détermination de comment limiter cette incertitude

Cette étape vise à réduire les efforts de collecte de données. La confidentialité ou la volonté du commanditaire de l'étude de collecter des données peuvent pénaliser ou favoriser la réduction de cette incertitude. La réduction se focalisera sur les données dont les contributions sont les plus fortes, les plus incertaines et les plus accessibles.

En fonction du temps disponible pour cette étape, deux procédures sont proposées dans le guide : une procédure allégée et une procédure complète. Pour la procédure allégée, l'approche est plus qualitative que quantitative et vise à réduire l'incertitude des étapes majeures par une bonne représentativité et une bonne modélisation.

- **4. Une description des outils qui peuvent aider**

- • La recherche dans les bases de données ACV pour évaluer si les données sont de suffisamment bonne qualité
- • Une ACV screening pour identifier les principaux enjeux des résultats ACV
- • Une analyse des contributions qui vont montrer les procédés et les flux les plus contributeurs au bilan environnemental
- • L'analyse de la littérature ACV pour identifier les principales étapes contributrices et les paramètres influents
- • L'analyse de documents sectoriels pour identifier les différences de technologies, leur représentativité géographique et temporelle et ainsi les points clés qui influenceront le bilan environnemental
- • L'expertise sectorielle et ACV qui va fournir des informations sur les procédés en jeu, les grandes tendances et les données où l'incertitude est élevée
- • L'utilisation de fonctionnalités de logiciels ACV qui fournissent des informations sur la qualité et disposent de possibilité d'analyse de contributions.

Ce guide a été testé par les membres de SCORELCA.

Les perspectives de ce guide sont :

- De fournir un cadre structuré pour définir initialement la qualité des données en ACV avec des étapes et des outils
- De montrer qu'une expertise peut être sollicitée mais que beaucoup d'approches font également appel au bon sens
- De rappeler qu'à ce jour aucun élément complémentaire n'est demandé dans les rapports ACV sur ce sujet, mais que cette étape permet de gagner du temps
- D'identifier des éléments de facilitation de l'utilisation de données de meilleure qualité : l'interopérabilité des bases de données (pour éviter de n'utiliser, par facilité, qu'une seule base de données) et le développement des revues critiques en continu qui permettrait de valider cette analyse initiale.

## Lexique

Terminologie		Définition (source ISO 14044 : 2006)
Française	Anglaise	
<b>Qualité des données</b>	Data quality	caractéristiques des données reposant sur leur capacité à répondre aux exigences requises
<b>Revue critique</b>	critical review	processus destiné à s'assurer de la cohérence entre une étude d'analyse du cycle de vie et les principes et exigences spécifiés par les Normes internationales traitant de l'analyse du cycle de vie NOTE 1 Les principes sont décrits dans l'ISO 14040:2006. NOTE 2 Les exigences sont décrites dans l'ISO 14044 :2006
<b>Contrôle de complétude</b>	Completeness check	procédé permettant de vérifier si les informations des phases précédentes d'une analyse du cycle de vie suffisent pour arriver à des conclusions, conformément à la définition des objectifs et du champ de l'étude
<b>Contrôle de sensibilité</b>	Sensitivity check	procédé permettant de vérifier que les informations obtenues à partir d'une analyse de sensibilité sont pertinentes pour établir des conclusions et donner des recommandations
<b>Contrôle de cohérence</b>	Consistency check	procédé mis en œuvre avant d'arriver aux conclusions, permettant de vérifier que les hypothèses, les méthodes et les données sont appliquées de manière cohérente tout au long de l'étude, et conformément à la définition des objectifs et du champ de l'étude
<b>Analyse de sensibilité</b>	Sensitivity analysis	procédure systématique pour estimer les effets sur les résultats d'une étude des choix concernant les méthodes et les données
<b>Données d'activité</b>		Données collectées sur site
<b>Données d'inventaire</b>		Données issues d'un inventaire de cycle de vie, restituées comme telles sous une forme « agrégée » par flux élémentaire.
<b>Flux élémentaire</b>	Elementary flow	Flux de matière ou énergie : - entrant dans le système étudié, qui a été puisé dans l'environnement sans transformation humaine préalable, - sortant du système étudié, qui est rejeté dans l'environnement sans transformation humaine ultérieure
<b>Inventaire de cycle de vie</b>	Life Cycle Inventory	Inventaire qui catalogue les flux élémentaires traversant les frontières du système et fournit le point de départ pour l'évaluation de l'impact du cycle de vie
<b>Méthode/modèle de caractérisation</b>	characterization model	Ensemble de facteurs qui sont utilisés pour convertir les résultats de l'inventaire du cycle de vie en unité commune d'indicateur de catégorie. Les modèles de caractérisation traduisent les mécanismes environnementaux
<b>Catégorie d'impact</b>	impact category	Classe représentant les points environnementaux étudiés à laquelle les résultats de l'inventaire du cycle de vie peuvent être affectés. Exemple : l'effet de serre, l'acidification de l'air
<b>Processus élémentaire</b>	Unit process	plus petite partie prise en compte dans l'inventaire du cycle de vie pour laquelle les données d'entrant et de sortant sont quantifiées
<b>Système de produits</b>	Product system	ensemble de processus élémentaires comportant des flux de produits et des flux élémentaires, remplissant une ou plusieurs fonctions définies, qui sert de modèle au cycle de vie d'un produit
<b>Système d'arrière-</b>	<b>Background</b>	Les données provenant d'un tel système sont issues d'un mix

<b>plan</b>	<b>system</b>	de fournisseurs sur le marché. Ces données sont donc des données moyennes de marché. (sur base de l'ILCD Handbook)
<b>Système de premier plan</b>	<b>Foreground system</b>	Les données provenant du système de premier plan sont les données spécifiques d'un fournisseur. On peut utiliser une donnée « background » pour modéliser une donnée foreground.
<b>ISO</b> <i>International Organization for Standardization</i>		Organisation internationale de normalisation qui établit et publie des normes internationales dont les normes relatives à l'ACV ; ex : l'ISO 140 40 : 2006 ou encore l'ISO 14 025 : 2006
<b>PEF</b> <i>Product Environmental Footprint</i>		Méthodologie d'évaluation de l'empreinte environnementale d'un produit développée par la Commission Européenne. Elle est basée sur la méthodologie ACV. Elle permet de développer des déclarations environnementales de Type III de la norme ISO 14025 :2006 dans un cadre méthodologique spécifique. En 2014, cette méthode est en cours d'expérimentation, aussi la méthode pourra être amenée à évoluer.

## I. Introduction

Dans une ACV, la gestion de qualité est cruciale à chaque étape et en particulier au démarrage de l'étude.

Si généralement la qualité est abordée lors de la présentation des résultats comme un élément permettant de caractériser la fiabilité des données et la robustesse des interprétations et des conclusions, il n'en reste pas moins qu'il est nécessaire, selon les normes ISO 14040 et 14044, **d'établir dès la définition des objectifs et du champ de l'étude**, des exigences initiales sur la qualité des données.

Pour les praticiens ACV, le besoin en qualité est souvent difficile à caractériser en amont du projet car:

- Le besoin varie selon :
  - **Le but et le champ de l'étude**, le niveau de qualité ne sera pas le même ;
  - Les **types de données** : données d'activité, flux élémentaires (émissions, consommations), données d'inventaires, facteurs de caractérisation, résultats par indicateur de catégorie d'impacts.  
  
C'est pourquoi les critères pour caractériser cette qualité peuvent être différents.
- Les outils de mesure de la qualité existants sont incomplets et leur résultats différents :
  - Incomplets : La mesure de la qualité des données issue des bases de données n'apporte d'élément que sur la qualité intrinsèque (précision) de la donnée, et non sur son caractère approprié au but et champ de l'étude (cohérence géographique, temporelle, méthodologique et technologique)
  - Incohérents : Les critères de qualité varient selon les outils. Leur méthode d'évaluation de la qualité (approches qualitative, quantitative ou semi quantitative) et les résultats inhérents sont donc également variables.
- Le temps dont dispose le praticien pour cette tâche est souvent court.

Pour guider les praticiens dans ce besoin de définition initial de la qualité des données, SCORELCA souhaite disposer d'un guide permettant de :

- Comprendre les définitions et critères qui sous-tendent la qualité des données d'une ACV
- Disposer d'une méthode pour définir la qualité initiale (illustrée de cas pratiques)
- Disposer de l'inventaire des outils disponibles ou à mobiliser pour définir ce niveau de qualité initiale.

## II. La qualité en ACV -fondements

### II.1. Définition

- La qualité des données est définie, selon les normes ISO 14040 et 14044 comme « **les caractéristiques des données reposant sur leur capacité à répondre aux exigences requises** » (ISO 14040 et ISO 14044).
- Les exigences de qualité des données doivent être définies **pour répondre aux objectifs et au champ de l'étude.**

La définition des exigences de qualité initiale des données est donc reconnue comme un requis de l'ISO 14040 car si la qualité n'est pas suffisante, et ce risque est plus grand si les exigences de qualité sont mal évaluées au départ, le but de l'étude peut ne pas être atteint.

### II.2. Champ de la qualité

Dans une ACV, l'analyse de la qualité s'applique **sur toutes les données ACV** :

- Les données d'activité ou données primaires (ex : des consommations d'énergie, des km de transport, émissions...)
- Les données d'inventaires (flux élémentaires) :
  - Données secondaires tirées des bases de données ACV
  - Données calculées dans le cadre de l'ACV
- Les données finales, résultat du calcul ACV par indicateur de catégories d'impacts, dont la qualité dépend :
  - Des données secondaires utilisées,
  - De certains aspects méthodologiques (ex : allocation) et de modélisation (ex : frontières du système)
  - Des méthodes de caractérisation retenues,

**Les exigences de qualité portent intrinsèquement sur les résultats (données finales)** et il faut les traduire, par anticipation des contributions à l'incertitude sur les résultats, en **besoins de qualité des données d'inventaire et d'activité.**

Remarque : Le champ de notre guide porte uniquement sur les données d'inventaire et d'activité, bien que théoriquement, des exigences de qualité peuvent également porter sur les méthodes (facteurs) de caractérisation. Une autre étude de SCORELCA porte sur l'évaluation de différentes méthodes de caractérisation et le JRC a également évalué cette qualité des méthodes.

## II.3. Enjeux de la qualité

La description de la qualité des données est importante pour mesurer la fiabilité des résultats de l'étude et garantir une bonne interprétation.

Par ailleurs, il existe 2 types d'ACV :

- Les ACV comparatives, qui comparent plusieurs moyens de remplir une fonction, que ce soit des produits, procédés, systèmes différents
- Les ACV non-comparatives, qui procèdent à un auto-référencement (amélioration de l'existant, comparaison de variantes).

**Les exigences de qualité seront très variables selon le type d'étude.** En effet, on attend de l'ACV une réponse discriminante, c'est-à-dire que les intervalles d'incertitude ne se superposent pas.

Dans le cas des ACV comparatives, les systèmes peuvent être très différents et il est nécessaire d'avoir une bonne fiabilité des données pour les procédés et flux élémentaires qui contribuent le plus aux impacts globaux.

En revanche, dans les ACV de type « auto-référencement », les systèmes comparés ont une grosse partie en commun. L'incertitude qui porte sur ce tronc commun n'affecte pas la fiabilité des résultats comparatifs.

*Ainsi par exemple, une ACV qui vise à estimer les impacts environnementaux d'une amélioration du rendement de pneus peut se satisfaire de données médiocres concernant la modélisation des autres pièces.*

Une mauvaise qualité des données peut empêcher l'identification des différences significatives entre systèmes du fait notamment :

- D'une mauvaise qualité des modèles de caractérisation
- D'un manque de complétude de l'inventaire du système étudié en raison de coupures et de lacunes des données
- D'une qualité inadéquate des données d'inventaire en raisons :
  - De l'incertitude intrinsèque
  - De l'incohérence temporelle, géographique ou technologique
  - De différences de choix méthodologiques : règles d'affectation, d'agrégation, de coupure, périmètres (ex : prise en compte ou non des infrastructures...).

**Au final, la définition initiale des besoins en qualité permet de faire des choix et de déterminer là où on mettra et où on ne mettra pas les efforts pour obtenir une étude de qualité en limitant le travail.**

Elle permet ainsi de réduire le temps de collecte de certaines données, voire de diminuer le nombre d'itérations.

## II.4. Une étape dans le processus itératif d'ACV

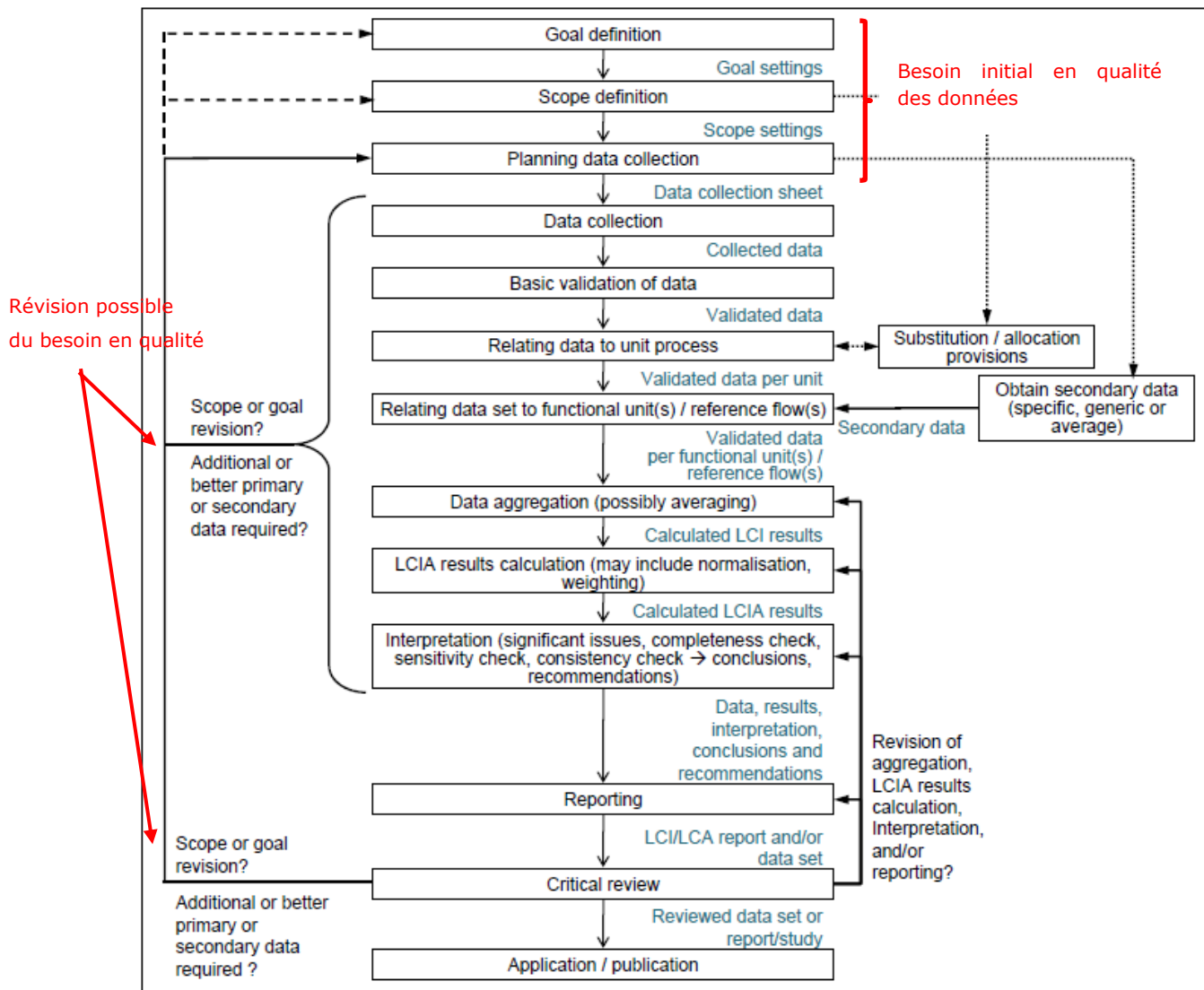
Notons que si la spécification initiale des besoins en qualité doit se faire dès l'**amont de l'étude, elle peut intervenir jusqu'après une première itération (ou des calculs sommaires)**, ce qui permet d'identifier ces enjeux.

Par ailleurs, il est possible de revenir, en cours d'étude, sur les exigences initiales si cela est justifié par les résultats et approprié par rapport au but et champ de l'étude.

**En cela, la spécification des besoins en qualité s'inscrit totalement dans l'approche itérative de l'ACV.**

Le schéma ci-dessous présente l'approche itérative de l'ACV et en rouge, les possibilités d'amélioration de la qualité des données via une approche itérative :

**Figure 1 : Illustration de l'approche itérative pour la qualité des données**



Source : ILCD Handbook, General Guide for Life Cycle Assessment  
 – Detailed guidance, figure 5 p 27

## II.5. Les critères de qualité

Différents critères sont proposés pour caractériser la qualité. Ces critères varient selon les documents de référence.

### II.5.1. DEUX GRANDES NOTIONS : PRÉCISION ET JUSTESSE DE LA REPRÉSENTATION

On peut cependant regrouper ces critères de qualité au sein de deux grandes notions :

- **❶ La précision intrinsèque des données brutes et calculées**

Elle est affectée par l'imprécision des mesures proprement dites (précision des appareils) et par le manque de représentativité des mesures (ex : mesure sur procédé stable alors qu'il est instable 40% du temps avec un rendement nettement moindre).

Elle peut aussi être affectée par des erreurs humaines : erreurs de transcription (ex : décalage d'une ligne lors d'une copie dans un fichier Excel), de calcul (poids sec ou humide), d'unités (kWh ou MJ).

Enfin, l'imprécision peut provenir d'un modèle permettant d'établir des données (par exemple un modèle permettant de calculer les émissions de pesticides à partir de la consommation d'engrais).

L'imprécision intrinsèque doit être évaluée **par le créateur de la donnée. Elle peut être aisément intégrée dans une base de données.** C'est la raison pour laquelle il sera indispensable de décrire la façon dont la qualité des données est évaluée dans les différentes bases de données ACV.

Note : L'imprécision à laquelle il est fait référence ne porte aucunement sur le modèle ACV qui est réalisé.

- **❷ La correspondance entre l'objet à représenter (système réel) et sa représentation (modélisation ACV)**

**La correspondance entre le but et champ de l'étude et la modélisation (*appropriateness*) est spécifique à l'ACV réalisée.** L'ampleur de l'incertitude dépend des difficultés de cohérence entre la donnée utilisée et la donnée « réelle » :

- Les périodes de temps :
  - Données anciennes pour modéliser un procédé actuel,
  - Données actuelles pour évaluer un procédé du futur...
- Les technologies :
  - Degré d'avancement (technologie de pointe, moyenne, ancienne)
  - La justesse de sélection pour représenter correctement la technologie (ex : modéliser la production de PP en l'approchant avec du HDPE est-il acceptable ?)

- Les lieux : la correspondance géographique est en partie redondante avec la correspondance des technologies, mais elle doit aussi prendre en compte :
  - Les impacts en amont (fournisseurs, mix électrique...),
  - Les impacts locaux fonction des conditions locales (rareté de l'eau, type de traitement des déchets, sensibilité des écosystèmes impactés...)
  - Le mode d'exploitation des procédés (qualité de l'entretien, contrôle des émissions...)
- Les aspects méthodologiques :
  - Frontières de système
  - Critères de coupure
  - Règle d'affectation
  - Choix des catégories d'impacts.

Ces deux notions peuvent se représenter graphiquement au travers du schéma de la figure ci-dessous.

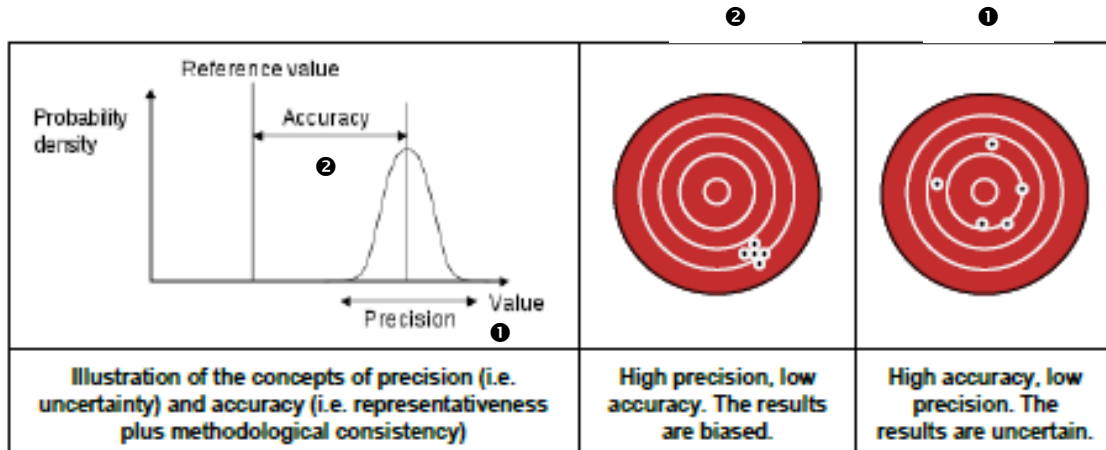
- La précision intrinsèque représente la dispersion autour de la valeur modélisée : on peut également parler **d'incertitude**.
- La correspondance entre objet de la mesure et objet de la modélisation (appropriateness) représente l'écart entre la valeur « réelle » et la valeur modélisée. Cet écart représente un **biais**. Il est difficile de quantifier ce biais dans la mesure où l'on ne connaît jamais parfaitement le cas « réel ». Il est cependant possible d'identifier certaines approximations, hypothèses retenues dans le cadre de la modélisation.

Au final, c'est souvent l'incertitude globale, qui agrège ces deux sources d'incertitude, qui va qualifier la qualité d'une donnée.

Note : dans le vocabulaire de l'ILCD ces deux notions correspondent à « precision » (incertitude intrinsèque des données) et « accuracy » (incertitude liée à la représentation de l'objet étudié).

Le schéma ci-dessous illustre ces deux notions, de deux façons visuelles différentes :

**Figure 2 : Expression des deux grandes composantes de la qualité (précision intrinsèque et précision de modélisation)**



Source : ILCD Handbook, General Guide for Life Cycle Assessment  
 – Detailed guidance, figure 28 p 326

### II.5.2. LES CRITÈRES DE QUALITÉ D’UNE DONNÉE ACV

En résumé, bien qu’il existe des différences entre documents de référence sur la manière de les nommer et de les évaluer, on peut citer comme critères communs permettant de définir la qualité d’une donnée :

- La représentativité
  - Technologique
  - Géographique
  - Temporelle
- La complétude
- La cohérence méthodologique
- L’incertitude

Caractère approprié de la modélisation par rapport au système réel

Incertitude liée à la dispersion des valeurs

De manière générale, une idée de **l’incertitude globale** est fournie par la suite, qui recouvre ces deux notions.

Note : il n’est pas toujours aisé d’identifier si les outils / bases de données parlent d’incertitude globale ou d’incertitude liée à l’une des deux composantes.

Des détails sur ces différents critères et les enjeux de qualité sont précisés en annexe 1 de ce document.

### III. La qualité initiale dans les textes de référence

Ce paragraphe synthétise les conclusions d’une analyse approfondie de certains documents de référence :

- L’ensemble des documents de référence s’accorde sur l’importance de la spécification initiale de la qualité des données.
- Les critères de qualité sont similaires

Le tableau ci-dessous récapitule ces critères et aborde la façon dont ils sont évalués par les principaux documents de référence. Trois approches sont possibles :

- Une approche qualitative : approche descriptive, non standardisée
- Une approche quantitative (caractérisée par un chiffre pour le critère en question ou une valeur d’écart-type)
- Une approche semi-quantitative : attribution d’une note sur base d’une grille de notation.

**Tableau 1 : Récapitulatif des critères de qualité couverts par les documents de référence et modes d’évaluation recommandés**

Termes en FR	Termes en EN	ISO 14044 (2006)		ILCD Handbook (2010)		PEF methodology (2013) <sup>1</sup>	
		Présence du critère	Nature de l’évaluation	Présence du critère	Nature de l’évaluation	Présence du critère	Nature de l’évaluation
<b>Caractère approprié de la modélisation / Appropriateness</b>							
Représentativité temporelle	Time-related representativeness	X (f)	Qualitatif	x	Semi-quantitatif	x	Semi-quantitatif / qualitatif
Représentativité géographique	Geography representativeness	X (f)	Qualitatif	x	Semi-quantitatif	x	Semi-quantitatif / qualitatif
Représentativité technologique	Technology representativeness	X (f)	Qualitatif	x	Semi-quantitatif	x	Semi-quantitatif/ qualitatif
Complétude	Completeness	X (e)	Quantitatif	x	Quantitatif & semi-quantitatif	x	Quantitatif & semi-quantitatif/ qualitatif
Cohérence méthodologique	Methodological appropriateness and consistency	X (g)	Qualitatif	X	Semi-quantitatif	x	Semi-quantitatif/ qualitatif
<b>Précision intrinsèque / Precision</b>							
Précision	Precision uncertainty /	X(d)	Quantitatif		Quantitatif & semi-quantitatif		Semi-quantitatif/ qualitatif
<b>Incertitude globale / overall uncertainty</b>							

<sup>1</sup> Ces éléments pourraient, le cas échéant, être amenés à évoluer à l’avenir. Les documents sur lesquels nous nous basons sont les plus récents à date de l’établissement du rapport (2013-2014).

	Overall uncertainty	X (j)	Quantitatif	x	Semi-quantitatif	x	Semi quantitatif
<b>Autres critères</b>							
Couverture temporelle	Time-related coverage	X (a)	descriptif				
Couverture géographique	Geographical coverage	X (b)	descriptif				
Couverture technologique	Technology coverage	X (c)	descriptif				
Reproductibilité	Reproducibility	X (h)	Qualitatif				
Source	Sources of data	X (i)	descriptif				
Documentation	Documentation (source, format transparency to ensure reproducibility)		descriptif	X		x	Format ILCD
Revue	Review aspects		descriptif	X		x	
Nomenclature	Nomenclature (structure and the elementary flows)		descriptif	X		x	Flux ILCD

- D'autres éléments additionnels visant à mieux caractériser les aspects de qualité sont importants à analyser :
  - La description de certaines informations dans la documentation
  - Les nomenclatures de flux
  - Les types de revues mises en œuvre et la compétence des réviseurs
- La mobilisation d'un jugement d'expert en première approche est mentionnée à plusieurs reprises.
- Peu de recommandations sont présentées sur la spécification de la qualité initiale des données dans les textes de référence, mais quelques éléments communs :
  - L'analyse des contributions des procédés aux flux
  - L'identification d'exigences sur base de PCR ou EPD
  - L'analyse de données issues d'études antérieures en prenant notamment en compte les analyses de gravité, de sensibilité, ou d'incertitude
  - Eventuellement, une première itération.
- Il n'est nulle part mentionné que les rapports ACV doivent inclure cette définition initiale. En effet, dans la mesure où elle fait partie du processus itératif, c'est la qualité finale des données qui est consignée dans le rapport d'ACV.

Des détails sur les différents documents de référence sont présentés en annexe 2.

## IV. Les étapes à suivre pour définir les exigences de la qualité initiale

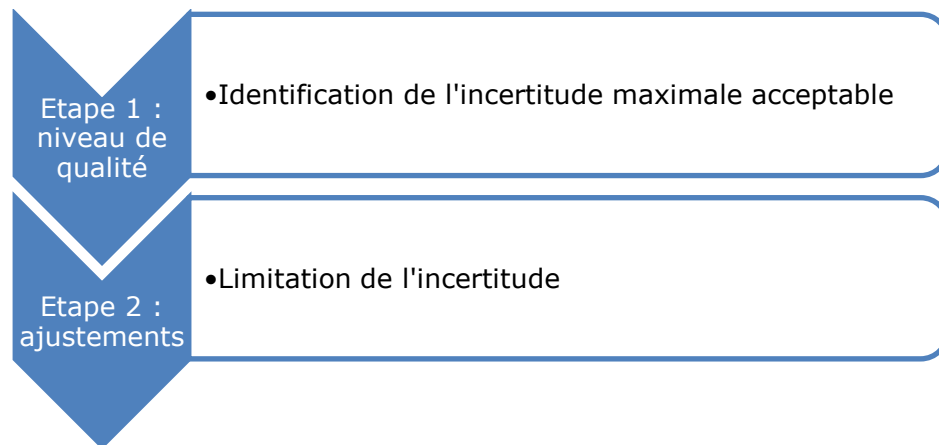
### IV.1. Approche générale

Une ACV doit permettre de tirer des conclusions quant à la comparaison d'impacts environnementaux de systèmes, que ce soit dans une étude comparative (systèmes différents) ou une étude d'évolution (changements de systèmes, typiquement pour de l'écoconception). Pour cela, il faut que l'incertitude sur les résultats soit plus petite que la différence entre les résultats, autrement dit que les résultats soient discriminants. **La spécification initiale des besoins en qualité cherche à estimer a priori le besoin de finesse des données qui permettra de réduire l'incertitude au moins jusqu'à ce niveau.**

Si les différences attendues de résultats sont grandes (systèmes très différents) ou si seules les grandes différences sont pertinentes (décisions stratégiques), l'incertitude sur les résultats peut être plus grande.

Il existe donc deux grandes étapes pour établir les exigences de la qualité initiale des données, qui seront par la suite détaillées :

**Figure 3 : Les deux grandes étapes pour définir la qualité initiale des données**



### IV.1.1. L'IDENTIFICATION DE L'INCERTITUDE MAXIMALE ACCEPTABLE

**La spécification initiale commence donc par une analyse du but** (faut-il de gros écarts ou non pour prendre une décision discriminante ?) **et une estimation d'expert sur les écarts attendus.** Ceci fournit l'intervalle d'incertitude acceptable.

Notons qu'il y a **un intervalle différent pour chaque catégorie d'impact.** Les catégories qui méritent le plus d'attention en matière de réduction des incertitudes sont celles :

- **Où les impacts globaux sont jugés les plus importants,**
  - Soit intrinsèquement (changement climatique vs. écotoxicité)
  - Soit en raison d'impacts normalisés élevés<sup>2</sup> (eutrophisation liée aux émissions de nitrates pour des systèmes agricoles)
- **Où les écarts attendus sont les plus réduits** et où la discrimination sera donc plus difficile à obtenir

### IV.1.2. LA LIMITATION DE L'INCERTITUDE

Il faut ensuite déterminer comment réduire de manière efficiente l'incertitude au(x) niveau(x) requis. **Le critère central est la diminution de l'incertitude par unité de temps** (de l'auteur) **ou de coût** (s'il faut acheter des données ou consulter un expert externe). L'affinage des données sera d'autant plus prioritaire que :

- Les données sont une grosse source d'incertitude sur les résultats. Cette incertitude est le produit de :
  - La contribution de la donnée aux résultats comparés<sup>3</sup> (le flux élémentaire X contribue à raison de 25% à la catégorie d'impacts)
  - L'incertitude relative de la donnée<sup>4</sup> (intervalle de +/-Y% autour de la valeur centrale)
- Le temps (coût) d'acquisition des données est court. Il est court si :

---

<sup>2</sup> Cette identification des impacts normalisés se base sur un jugement d'expert ou, si elle est disponible, sur la base de la littérature.

<sup>3</sup> En fait, la contribution de la donnée à la différence entre les résultats des différents systèmes. En effet, s'il y a un tronc commun, la contribution des procédés du tronc commun n'est pas à considérer.

<sup>4</sup> Cette incertitude relative peut être estimée à partir d'informations fournies par les bases de données. Si la base de données renseigne la donnée comme intrinsèquement non fiable, on sait directement que l'incertitude relative est forte (voir l'Annexe 4 pour des exemples d'informations disponibles). Si la base de données ne fournit que des métadonnées (géographie, temps, technologie, nombre de sources), l'utilisation d'outils quantitatifs est à proscrire car trop consommateurs de temps. C'est toujours à l'expert d'estimer l'incertitude relative sur base de ces métadonnées, dont les influences combinées changent du tout au tout selon les cas.

- Le fournisseur de données dispose déjà des données et qu'elles ont le bon format
- Le fournisseur de donnée est identifié et motivé (c'est le cas typique des commanditaires et de certains fournisseurs).

Par ailleurs, des critères complémentaires peuvent également favoriser / pénaliser cette réduction d'incertitude:

- **La volonté du commanditaire** de connaître plus précisément le système qui le concerne. On peut donc collecter des données primaires pour des motivations autres que de rendre les résultats discriminants.
- **La contrainte de confidentialité** : Le commanditaire ne veut pas que des concurrents sachent qu'il réalise une étude et ne veut donc pas leur demander des données.

Dans ces deux cas, ce sont les données des systèmes du commanditaire qui sont plus affinées que ceux requis sur la base du critère central d'efficience.

## IV.2. Les étapes à suivre

### Remarque pour la prise en compte des éléments de recommandation :

Il est fondamental, lors du démarrage de l'étude, **d'identifier le niveau de consentement du commanditaire à passer du temps** sur cette spécification initiale de la qualité.

- **S'il n'existe pas de demande sur la qualité des données** ou sur l'incertitude, et/ou que le temps d'ACV est restreint (***1 à 2 j de travail pour le praticien***), alors **la procédure allégée** à suivre est la suivante :
  - Identifier les enjeux environnementaux clés et sélectionner les catégories d'impacts adéquates (sur base d'une réflexion « interne » et d'expertise)
  - Rechercher les « hotspots » : étapes / procédés les plus contributeurs (via les bases de données ou la littérature)
  - Identifier si parmi les données accessibles rapidement (bases de données, données primaires ou secondaires), il est possible de réduire l'incertitude en adaptant au but et champ de l'étude concerné:
    - Les aspects temporels
    - Les aspects géographiques
    - La modélisation
- **Si l'ACV a un budget suffisant et un délai suffisant** (*10-30 j*), alors la procédure proposée ci-dessous convient

Par ailleurs, les commanditaires peuvent parfois être peu expérimentés en matière d'ACV et d'importance des questionnements sur la qualité des données.

Dans ce cas, le guide peut simplement constituer une aide pédagogique invitant le commanditaire à se poser certaines questions, et ainsi le faire progresser en ACV.

Il convient également de rappeler que le commanditaire n'étant pas un expert de l'ACV, il peut tout à fait travailler en collaboration avec le praticien ACV (expert) ou un expert sectoriel pour assurer cette définition de la qualité initiale de l'ACV. L'expert aura ainsi un rôle didactique auprès du commanditaire.

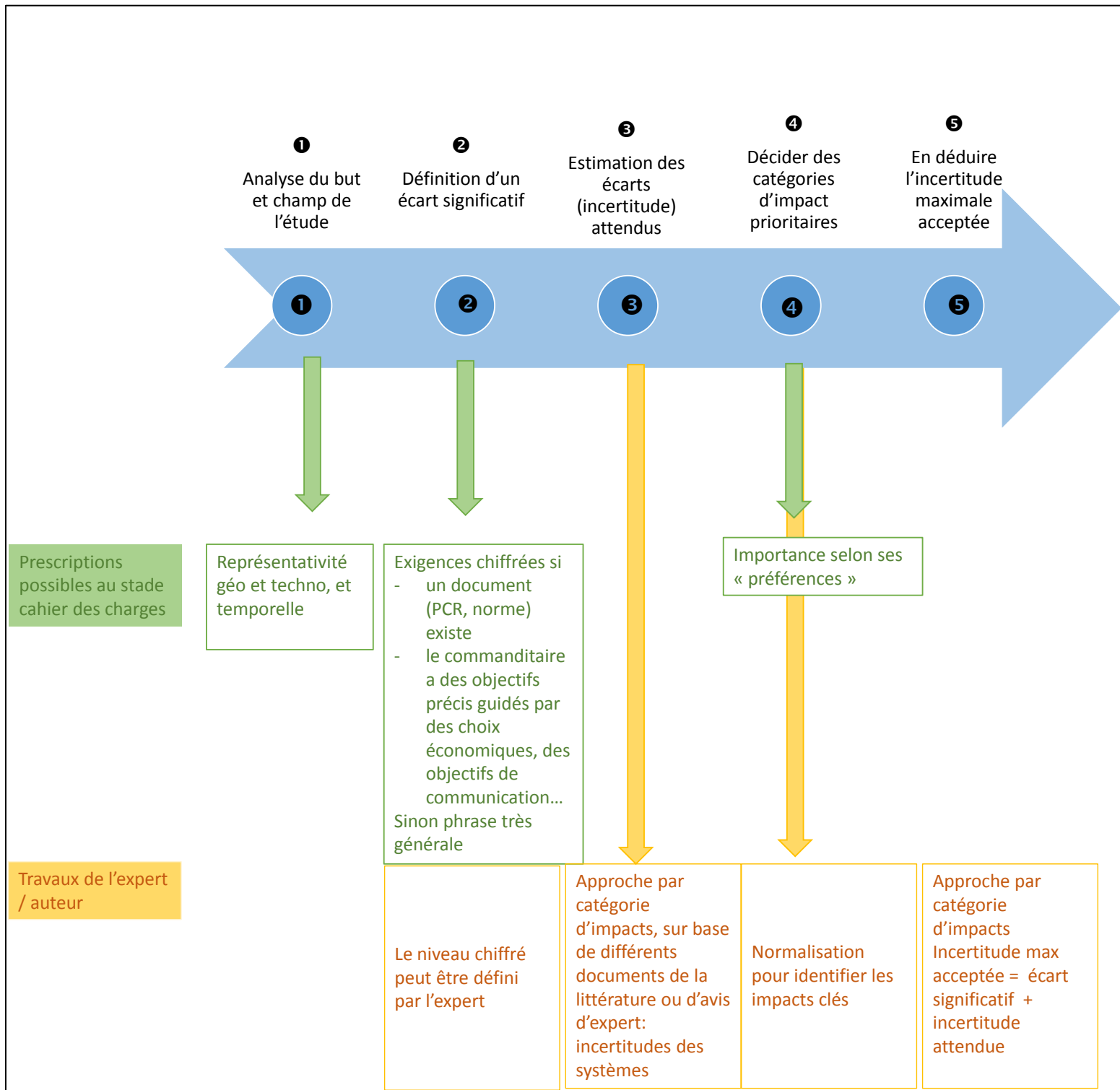
Pour chacune des deux étapes présentées ci-dessus, nous proposons un cheminement chronologique. Les tâches sont distinguées par couleur, en vert celles pour le **commanditaire**, en orange celles pour **l'auteur/ l'expert**.

Note : Parfois les rôles commanditaires et praticiens peuvent être confondus (cas par exemple d'un ingénieur ACV au sein d'une entreprise).

#### **IV.2.1. LES ÉTAPES À SUIVRE POUR LA DÉTERMINATION DE L'INCERTITUDE MAXIMALE**

Le schéma ci-dessous présente les différentes tâches :

**Figure 4 : Chronologie des tâches à effectuer par le commanditaire et l'auteur pour déterminer l'incertitude maximale acceptable**



- **1 Analyser le but et champ de l'étude**

En règle générale, cette étape a déjà été définie par le commanditaire qui a un but bien précis pour conduire une étude ACV. En fonction de ce but, il peut déjà invoquer, dans son cahier des charges, des éléments sur un certain nombre d'exigences notamment en matière de représentativité géographique, temporelle et technologique.

Exemples de phrases à insérer dans le cahier des charges :

Niveau léger :

*Le soumissionnaire proposera dans son offre, les données qu'il jugera les plus représentatives des scénarios à modéliser, en ce qui concerne l'âge des données, leur représentativité technologique et géographique.*

*Un effort sera porté pour disposer de données récentes. La justification de la représentativité des sources de données choisies sera explicitée.*

Niveau plus précis : exemple de l'ACV de la méthanisation en France au regard des autres voies de valorisation organique – Scénarios actuels et prospectifs

*Les données utilisées, devront, pour ce qui concerne la situation actuelle, se baser sur les données les plus récentes en matière de représentativité des différentes technologies existantes en France, tant pour la méthanisation que pour les autres voies de valorisation organique.*

*Ainsi si des données n'existaient pas de façon directe dans des bases de données, il conviendra de prendre en considération des données réelles d'entreprises de méthanisation.*

*Pour ce qui concerne le scénario prospectif, l'auteur de l'étude proposera, sur base d'analyse des tendances du marché, des hypothèses d'évolutions des données jugées clé, et mettra en évidence les niveaux d'incertitude.*

Remarque : Pour l'analyse conséquentielle, les enjeux clés vont porter sur la substitution et sur la bonne définition du périmètre.

- Exemple pour définir la bonne substitution :

*Pour la modélisation de l'énergie éolienne versus le photovoltaïque, il s'agit de se poser la question de la substitution : à quoi se substitue l'énergie renouvelable produite ?*

- *Au niveau de l'éolien :*

- *Si l'on est dans une approche où l'on répond aux objectifs réglementaires (de l'UE), alors l'éolien supplémentaire va venir se substituer à de l'éolien (ou à une autre énergie renouvelable) ; **l'effet de substitution sera faible.***

- *Si non, alors l'éolien vient substituer le mix électrique marginal du pays ; l'effet de substitution sera substantiel*
- *Il faut par ailleurs tenir compte des effets indirects sur la stabilité du réseau (besoin de flexibilité des modes traditionnels s'il y a plus d'éolien)*
- *Au niveau du photovoltaïque : il y a une meilleure corrélation entre l'offre et la demande (plus de production pendant les heures ouvrables) et donc une moindre augmentation du besoin de flexibilité qu'au niveau de l'éolien*

Conclusion pour le cahier des charges :

- Cas d'une étude prospective sur le développement de l'énergie éolienne versus le photovoltaïque

*L'auteur de l'étude devra notamment prendre en compte les enjeux de substitution dans le cas qui nous concerne. Ainsi les incertitudes liées au type de mix énergétique substitué (si plusieurs alternatives sont possibles) devront être caractérisées afin de permettre une bonne comparaison entre les scénarios développement de l'éolien et développement du photovoltaïque. De même, il faut tenir compte des effets indirects sur la stabilité du réseau (besoin de flexibilité des modes traditionnels s'il y a plus d'éolien).*

- Exemple pour définir le choix du périmètre :

*Pour la modélisation de la production de bioplastiques, il est important de connaître l'impact sur les surfaces cultivées : est-ce que la production de plastiques issus d'agro-ressources va venir concurrencer la production d'autres produits (comme par exemple le biodiesel, l'alimentation humaine, l'alimentation animale...) et va engendrer une augmentation des surfaces agricoles au détriment d'autres écosystèmes, forestiers ou autres? En fonction de la réponse, il convient d'adapter le périmètre et de l'étendre aux impacts liés à la réduction des surfaces cultivées pour l'autre produit (exemple : remplacement du biodiesel par du pétrole).*

Conclusion pour le cahier des charges :

- Cas d'une étude prospective sur le développement de la production de bioplastiques à l'échelle mondiale

*L'auteur de l'étude devra notamment prendre en compte les conséquences du développement de cultures dédiées aux bioplastiques et des éventuelles tensions et compétition des terres entre plusieurs usages. Si ces compétitions sont telles qu'elles venaient à pénaliser d'autres cultures, le périmètre de l'étude devra s'étendre aux conséquences engendrées (par exemple plus fort recours à la production de pétrole plutôt qu'au biodiesel). Les éventuelles incertitudes associées à l'étude de ces conséquences devront être prises en compte.*

- ② Décider ce qu'est un écart significatif (lié au but de l'étude)

Cet écart significatif doit être défini par le commanditaire.

Il peut être issu de plusieurs approches :

- **Le commanditaire vise à calculer l'impact environnemental de son produit en conformité avec un PCR développé pour l'affichage environnemental.**

Le PCR a généralement déjà défini un travail d'exigence de qualité, qui peut se présenter de la façon suivante :

- Exigence sur le niveau de qualité des données en fonction d'une grille de scores
- Etablissement de critères de coupure pré-définis et de procédés inclus et non-inclus dans le périmètre de l'étude
- Règles méthodologiques de modélisation
- Définition du niveau de « granulométrie » des données : données spécifiques à un producteur ou génériques à un secteur ou encore maximisantes.

Note : Si le but de l'utilisation du PCR est différent du but initial (qui vise à discriminer, au sein d'un groupe de produits, les produits les moins impactants), il est nécessaire de rester critique sur les exigences définies par le PCR.

Exemples de phrases à insérer dans le cahier des charges :

*Le soumissionnaire suivra les exigences de qualité des données définies par le PCR « X », à savoir : ... **(citer les exigences en matière de qualité des données et la source du référentiel)***

*Si pour une raison liée au but de l'étude le soumissionnaire considère que certaines exigences sont à revoir, elles seront mentionnées explicitement dans son offre.*

- **Il peut exister des exigences liées à des critères économiques**

*Exemple : un industriel souhaite investir dans un four à arc électrique nouvelle génération qui consomme moins. Les calculs de taux de retour sur investissement ont estimé qu'il fallait que par rapport à l'ancien four, le gain en matière d'impact environnemental soit de plus de 20 % par rapport à la situation initiale.*

- **Les systèmes modélisés peuvent avoir un gros tronc commun**

*Exemple : Dans le cas d'une approche en écoconception, si le produit est bien maîtrisé pour toutes les étapes, donc avec des données peu incertaines, et que les efforts portent sur une fraction du produit (exemple : on améliore la climatisation d'une voiture) l'entreprise peut exiger un niveau très faible d'incertitude (moins de 5 % d'écart sur certains impacts environnementaux comme les émissions de GES). En*

*revanche pour certaines catégories d'impacts aux méthodes de caractérisation plus incertaines, l'incertitude sera plus élevée et il conviendra de prendre en compte des écarts plus grands.*

- **Il peut exister des sources de forte incertitude pressenties**

*Par exemple dans le cas d'une étude prospective, certaines hypothèses engendrent des incertitudes fortes. De ce fait, l'incertitude générée par les autres données pourra être plus élevée.*

Exemples de phrases à insérer dans le cahier des charges :

Cas où le commanditaire est en mesure de définir ce seuil

*Nous estimons, sur la base de nos propres besoins, que les incertitudes acceptables sont de : **X** % pour toutes les catégories d'impacts (**ou bien détailler par catégorie d'impacts**)*

Cas où le commanditaire n'est pas en mesure de définir ce seuil

*Le prestataire définira le niveau acceptable de l'incertitude pour rendre discriminants les résultats des systèmes étudiés.*

Remarque :

Si l'objectif n'est pas directement une comparaison immédiate, l'enjeu est de réduire l'incertitude.

Si l'objectif est de reproduire, à des pas de temps réguliers, une analyse, pour mesurer les différences (objectif de comparaison historique), alors l'un des points sensible sera de mesurer en premier lieu la qualité des données que le commanditaire peut directement maîtriser.

- **3 Pré-estimer les écarts attendus par catégorie d'impacts**

Cette pré-estimation peut se faire au travers de différentes actions listées ci-dessous. Cette liste n'est pas exhaustive, mais fournit des suggestions hiérarchisé dans l'ordre le plus simple d'accès aux informations. La mise en œuvre de ces actions doit être légère (maximum quelques heures de travail).

- Recherche de bibliographie qui estime cette incertitude (via la littérature ACV comparant les systèmes) ;
- Recherche éventuelle des incertitudes individuelles des systèmes étudiés via la littérature ACV spécifique à l'un des systèmes mais aussi via une littérature sur les aspects environnementaux d'un produit : BREF qui mentionnent des données d'émissions, document d'écolabel, rapports environnement ou développement durable d'entreprise... ;
- Recherche des incertitudes des catégories d'impacts ;

- Recherche dans les bases de données des procédés reflétant chacun des systèmes et indiquant ainsi les incertitudes (globales ou par catégories d'impacts ou par flux) ;
- Premiers travaux d'ACV simplifiée (ACV screening) ;
- Faire appel à l'expertise externe est aussi toujours possible.

Les outils mobilisables seront détaillés par la suite (cf. §V.1 p.38).

• ④ Décider des catégories d'impacts prioritaires

▪ Décider de l'importance intrinsèque

L'importance intrinsèque est liée au choix du commanditaire et aux enjeux environnementaux du système étudié.

*Par exemple, pour les meubles, la question des consommations de ressources, renouvelables comme non renouvelables, de l'effet de serre (enjeu de captation de carbone des meubles en bois et utilisation d'énergie) ou la production d'oxydants photochimiques sont fondamentaux.*

Des choix spécifiques au commanditaire peuvent également entrer en ligne de compte. Ainsi, pour certains commanditaires, le plus gros enjeu portera sur les impacts sur l'effet de serre, et non sur l'éco-toxicité, pour laquelle les facteurs de caractérisation sont de faible qualité.

Dans ce cas le commanditaire peut pré-lister, dès le stade du cahier des charges, les catégories d'impacts qu'il juge prioritaires.

Exemple :

*Compte-tenu de nos attentes et des enjeux environnementaux du système étudié, il est demandé d'étudier les catégories suivantes : « XXX », « YYY », « ZZZ ».*

▪ Déterminer l'importance des impacts normalisés

Pour identifier les catégories d'impacts prioritaires, une approche objective peut consister à comparer l'impact moyen du système étudié avec les impacts totaux d'une zone géographique (un pays, un continent).

*Par exemple, le poste alimentation pèse lourd dans les impacts GES de la consommation des ménages et par conséquent, il convient de considérer les GES ; de même, la consommation d'eau utilisée pour produire et laver un vêtement, peut sembler peser relativement lourd dans les impacts totaux d'un européen moyen.*

Cette approche peut également se baser sur le jugement d'expert, sans calculs.

- ⑤ En déduire l'incertitude maximale acceptable (=écart minimal attendu pour chaque catégorie d'impact prioritaire)

L'incertitude acceptable calculée sera donc la différence entre l'écart attendu et l'écart significatif souhaité par le commanditaire et de l'incertitude attendue.

*Exemple :*

*Ecart significatif attendu par le commanditaire : 10 % sur l'effet de serre (si l'écart est moindre, l'investissement n'est pas jugé intéressant)*

*Ecart attendu entre les 2 systèmes : 30 % pour l'effet de serre*

→ *Incertaince acceptable sur la comparaison des systèmes : 20 %*

Note : l'incertitude a une composante absolue (erreur maximale) mais aussi une composante probabiliste. Une incertitude forte peut être acceptable si la probabilité d'erreur est faible. Par exemple, si différents fournisseurs utilisent des technologies très différentes, avec un seul dont la technologie a beaucoup plus d'impacts que les autres, l'incertitude est grande mais la probabilité que ce soit ce fournisseur-là est faible. On pourrait donc quand-même considérer que l'écart entre les systèmes est discriminant.

En pratique, il peut être difficile d'établir a priori, à l'aide des étapes précédentes, l'incertitude acceptable pour les catégories d'impacts considérées comme majeures ; aussi, nous proposons dans le tableau ci-dessous des repères de valeur en fonction des types d'études :

**Tableau 2 : Proposition de RDC Environment sur les types d'ordre de grandeurs d'incertitudes acceptables<sup>5</sup> selon les types d'études et les systèmes étudiés**

Type ACV	Stratégie, Long terme et Court terme, Eco-conception
Différence entre systèmes	Stratégie, prospectif
Grande différence	50 %
Faible différence	20%

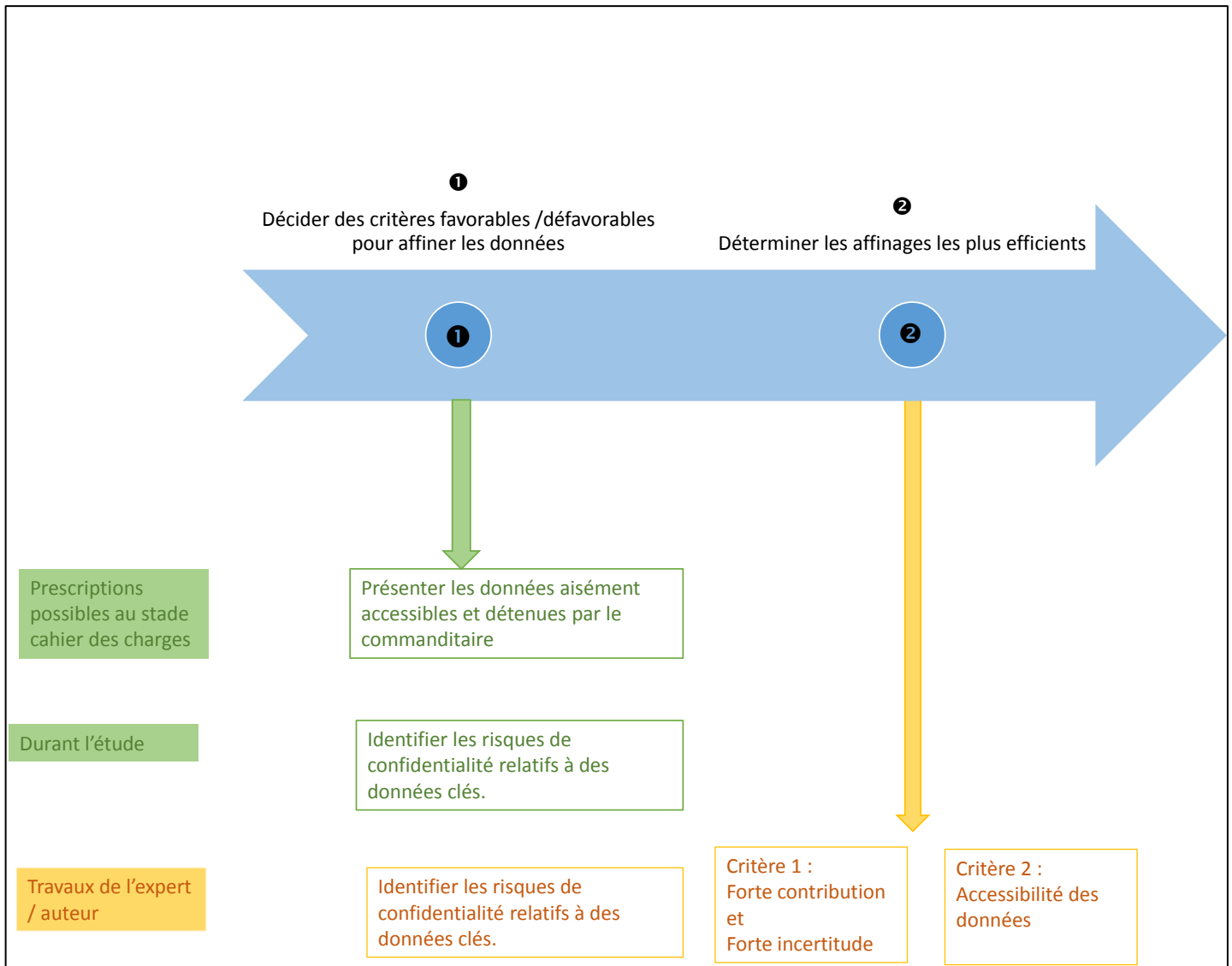
Cette approche allégée est cependant à mettre en regard des écarts que le commanditaire juge significatifs.

<sup>5</sup> L'incertitude entre les deux systèmes se calcule de la façon suivante :  $diff = (S2-S1) / S1$  où diff est la différence acceptable en % entre les 2 systèmes, S2 est le résultat d'impact du système 2, S1 le résultat d'impacts du système 1.

## IV.2.2. LES ETAPES À SUIVRE POUR DÉTERMINER LES MOYENS DE RÉDUIRE L'INCERTITUDE

Pour réduire l'incertitude, les étapes à suivre sont présentées et détaillées ci-dessous :

**Figure 5 : Etapes à suivre pour déterminer les moyens de réduire l'incertitude**



- **1** Décider des critères complémentaires favorables ou défavorables à la collecte de données plus fines

- Volonté de disposer / fournir des données de grande qualité sur le système que le commanditaire maîtrise.

Un niveau minima de qualité est attendu par le commanditaire, en particulier pour les données qu'il maîtrise ; ainsi si le praticien utilise des données génériques sans même les confronter à celles du commanditaire, ce dernier risque de considérer le niveau de qualité insuffisant. L'expert va donc surpondérer l'importance des données maîtrisée par le

commanditaire. Deux autres raisons motivent la surreprésentation des données des commanditaires :

- Leur facilité d'accès
- Le besoin de cohérence dans les documents émanant du commanditaire (données citées par ailleurs)

Notons toutefois que cela peut être une perte de temps si aucune de ces deux conditions n'est remplie, dans la mesure où certaines données maîtrisées par le commanditaire peuvent avoir une influence très faible sur les résultats. Le niveau de qualité pour ces données est alors supérieur à celui nécessaire pour l'objectif global.

Note : si l'incertitude est asymétrique (la réduction de l'incertitude va dans le sens d'une diminution des impacts calculés pour l'un des deux systèmes), l'interprétation de l'incertitude (plus forte pour le système concurrent) doit en tenir compte.

Ces éléments peuvent être transcrits dans le cahier des charges, par exemple en mentionnant précisément les données que le commanditaire mettra à disposition pour un niveau relativement fin.

- Confidentialité : certaines données, quoique disponibles, ne seront pas mises à disposition de l'auteur, car le commanditaire (ou le détenteur des données) les juge trop sensibles. De ce fait, si des données pour lesquelles le niveau de confidentialité se pose devaient être affinées, il est vraisemblable que le niveau de qualité sera plus faible.
- **2 Déterminer les affinages les plus efficaces**

On va favoriser la recherche de :

- Données à forte contribution et à forte incertitude relative sur la base
  - De la littérature
  - D'une estimation d'expert
  - D'un screening.
- Données accessibles à effort (temps, coût) réduit, sur la base :
  - D'une estimation d'expert
  - De la consultation de bases de données
  - Eventuellement de la consultation d'une source externe clé pré-identifiée.

Ces questions peuvent se résumer au petit schéma ci-dessous : plus les données sont influentes et accessibles, plus il sera nécessaire et aisé de réduire les incertitudes.

**Figure 6 : Schéma des critères clés pour sélectionner les affinages d’incertitudes**

		Accessibilité (coût de la recherche de données)	
		<b>facile</b> →	<b>difficile</b>
Influence sur les résultats	+	Enjeu qualité haut Réduction de l’incertitude aisée	Enjeu qualité haut Réduction de l’incertitude difficile
	-	Enjeu qualité faible Réduction de l’incertitude aisée	Enjeu qualité faible Réduction de l’incertitude difficile

Le cas le plus évident pour réduire l’incertitude est en vert ; les cas en orange sont à étudier au cas par cas. Le cas en rouge représente le cas où aucune action n’est nécessaire pour réduire l’incertitude.

Pour l’identification des données sur lesquelles réduire l’incertitude, deux types d’experts peuvent intervenir :

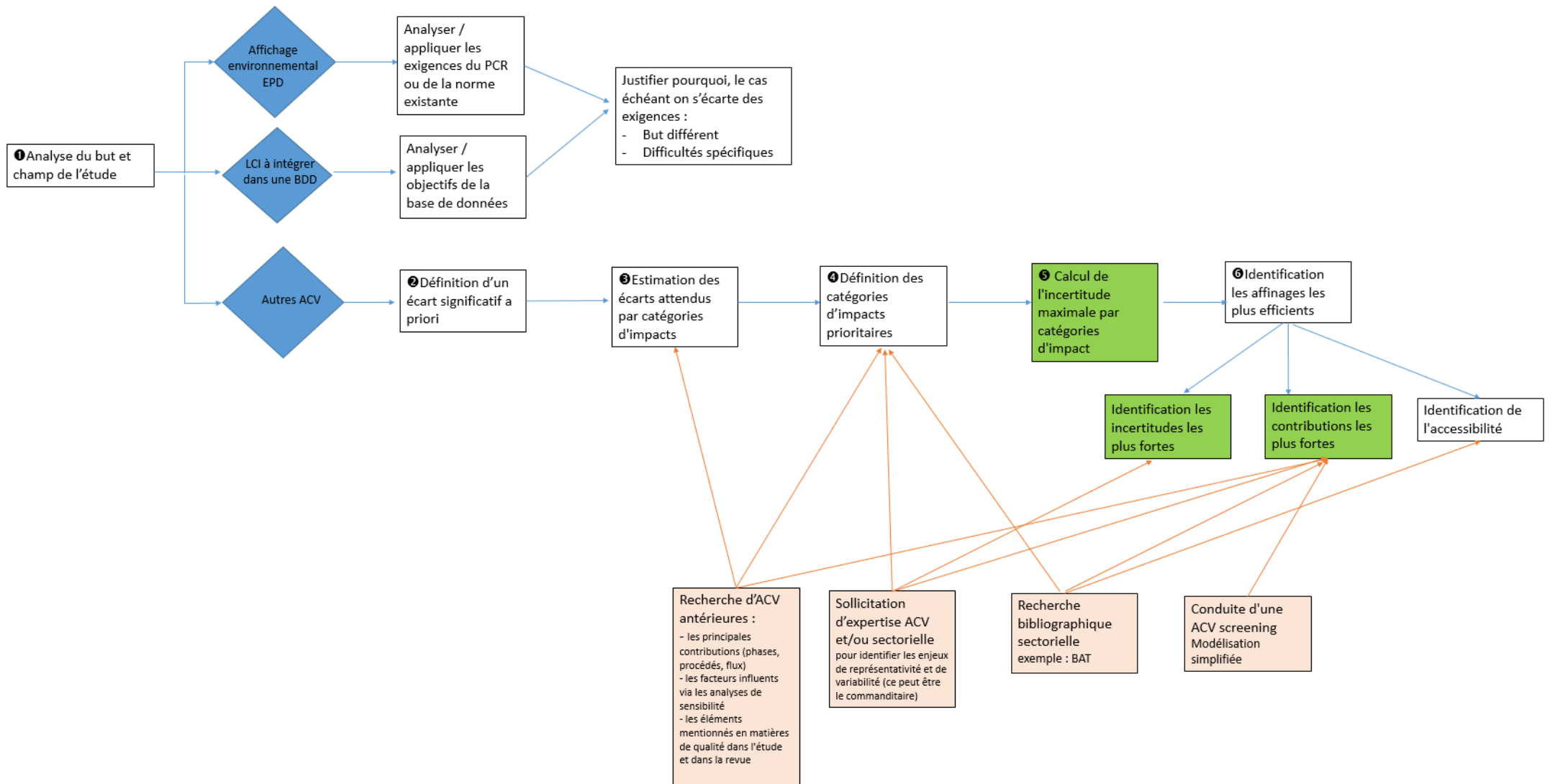
- Les experts ACV, pour les contributions relatives et, dans certains cas, pour les incertitudes relatives
- Les experts techniques, pour les incertitudes relatives difficiles à appréhender par l’expert ACV. Ces experts techniques doivent être interrogés par l’expert ACV. L’expert technique peut juger de l’incertitude générée par les incohérences temporelles, technologiques et géographiques et par l’étroitesse de l’échantillon (données d’un seul site ou de plusieurs).

### IV.3. Synthèse des recommandations

Cette partie consiste en la présentation des pistes les plus utiles afin de spécifier les besoins en ACV.

Nous proposons ci-dessous un logigramme qui vise à se poser les bonnes questions afin de déterminer de façon initiale la qualité requise pour les données. Il prend en compte les différentes recommandations citées par les documents de références et les experts.

Figure 7 : Logigramme pour accompagner les praticiens en matière de qualité initiale des données ACV



## V. Supports et outils disponibles pour bien définir la qualité

### V.1. Outils pouvant servir de supports à la définition de la qualité des données

Le logigramme consiste en un résumé des étapes à suivre mais aussi en un inventaire des outils pouvant aider à conduire ces différentes étapes. Ces étapes sont classées en fonction de la facilité d'accès à l'information et du temps nécessaire pour le praticien :

- Rechercher des informations dans les bases de données ACV
- Conduire une première itération
- Mener une analyse de la littérature ACV
- Mener une recherche bibliographique sectorielle
- Rechercher des informations sur les exigences de qualité via l'expertise

Note : bien évidemment **si une expertise interne** existe au sein de la structure du commanditaire, elle est à mobiliser en premier lieu.

#### V.1.1. RECHERCHE DANS LES BASES DE DONNÉES (BDD)

La recherche dans les bases de données vise à identifier si les données sur lesquelles on peut s'appuyer sont de bonne qualité, suffisante ou non.

Cette étape peut intervenir après la phase de screening, lorsque les principales contributions ont été identifiées. Elle peut également intervenir a priori, sur base d'un jugement d'expert ou d'hypothèses posées sur les procédés les plus contributeurs.

Cette recherche dans les bases de données permet donc de trancher plus rapidement, mais avec moins d'assurance que ce qui sera fait au stade de l'inventaire proprement dit, pour déterminer quelles données doivent être primaires (spécifiques) et quelles données peuvent être issues de bases de données.

Les rapports sont consultés lorsque les jeux de données ne fournissent pas assez d'informations directement sur la qualité.

Les 2 principales utilisations à ce stade de la qualité dans les Bdd :

- Identifier si les systèmes que l'on souhaite modéliser sont déjà présents directement dans une base de données.
  - Si oui identifier les incertitudes intrinsèques au système étudié
  - Si on a identifié les contributions majeures du système, on peut regarder si les procédés les plus contributeurs sont présents dans les Bdd et identifier également leur incertitude, pour mesurer comment la réduire le cas échéant.

### V.1.1. LA CONDUITE D'UNE « SCREENING » ACV OU D'UNE PREMIÈRE ITERATION

Dans les cas où des données ne sont a priori pas disponibles, mener une première itération, sur base d'un logiciel ACV ou de simples calculs rapides, à la main, peut fournir de premières informations pertinentes qui permettront d'identifier les étapes clés, notamment quand il existe peu de connaissance/ de littérature ACV sur le sujet.

Un échange avec certains fournisseurs de données (d'autant plus aisé si c'est le client) peut également permettre de collecter des données qui permettront, dès cette phase de modélisation simplifiée, de mieux caractériser le système à modéliser. Cette approche peut être faite très en amont, par une discussion sur les technologies rencontrées, la variété des procédés,... ou au démarrage de l'étude lors de l'élaboration d'un questionnaire de données. Notons que dans ce dernier cas, le risque est de solliciter de façon très intense le fournisseur de données sans encore avoir identifié quels étaient les paramètres les plus cruciaux.

### V.1.2. L'ANALYSE DES CONTRIBUTIONS

Une bonne analyse des contributions doit permettre, de proche en proche, d'identifier :

- Les contributions des différentes étapes du cycle de vie aux différentes catégories d'impact étudiées
- Les contributions, au sein de ces étapes, des procédés les plus impactants
- Le couplage de ces procédés impactants avec les flux élémentaires les plus impactants.

**Des méthodes de visualisation graphique sont présentes dans certains logiciels ACV pour pouvoir effectuer de telles analyses (cf. V.2p 42)**

Une fois ces contributions majeures identifiées, il est plus facile de déterminer là où il sera nécessaire de disposer de données de qualité.

### V.1.3. L'ANALYSE BIBLIOGRAPHIQUE ACV

L'analyse bibliographique est le premier outil aisément disponible. Les sources de données bibliographiques sont des ressources internes, des journaux spécialisés en ACV et/ou au niveau sectoriel mais également les rapports de certaines bases de données.

Il s'agit d'étudier des ACV existantes afin d'identifier les enjeux clés et notamment :

- **Les étapes du cycle de vie les plus impactantes**

Une fois que l'on a identifié les étapes les plus impactantes, on sait s'il suffit de disposer de données moyennes ou s'il faut des données spécifiques.

Savoir s'il faut aller rechercher de la donnée spécifique chez ses fournisseurs est un des enjeux des plus cruciaux pour les industriels. Cela peut cependant dépendre du positionnement de l'industriel dans la supply chain. Ainsi un fabricant et conditionneur de détergents est très attaché à cette question

tandis qu'une industrie productrice de produits chimiques sera peut-être moins sensible à ces aspects.

*Exemple :*

*Pour des produits lessiviels, les exigences de qualité sur le transport entre l'usine et le distributeur sont faibles car les enjeux sont faibles. En revanche, elles portent beaucoup sur la phase d'utilisation et par conséquent on va concentrer les efforts là-dessus.*

*Pour les couches, les enjeux sont davantage dans la composition des couches et le mode de production de ces composants. Dans ce cas par exemple, il est important de remonter à des données spécifiques de fournisseurs et non à des données moyennes.*

Ces éléments s'observent en général dans la présentation de résultats, sous la forme d'histogramme

- **Les paramètres clés** et notamment ceux influençant le plus les données : ceux-ci s'observent en général en analyses de sensibilité

Elles permettent d'identifier les paramètres influents, de façon plus fine le cas échéant qu'une simple vision par étape. Elles peuvent également permettre de disposer d'une vision des intervalles de valeurs et des valeurs moyennes.

Cela permettra d'identifier sur quelles données un niveau de qualité élevé est à obtenir.

En fonction des données, les travaux qu'il conviendra de mener en matière de collecte sont variables :

- Dans le cas des paramètres, il s'agit de données d'activité (consommations, émissions...) : il s'agira de disposer d'une collecte de données précise auprès des fournisseurs de données pour ces paramètres

Cette collecte pourra consister en un travail exhaustif (si le budget de l'étude le permet) avec un travail sur la variabilité, ou comme c'est le plus souvent le cas, en une consultation d'expert(s) visant à déterminer les fourchettes de valeur et d'attribuer, le cas échéant une distribution de probabilité.

- Dans le cas de données d'inventaires (jeux de données issus de bases de données ACV) : il s'agira de vérifier la qualité des données des inventaires retenues, tout d'abord au regard du caractère approprié au but et au champ de l'étude, mais aussi en matière de complétude, cohérence méthodologique...

Les autres éléments des analyses bibliographiques portent sur les contenus des revues critiques et éventuelles limites identifiées en termes de qualité, les sources de données... Les points d'attention sont de s'assurer que l'objectif et champ de l'étude sont assez similaires à celui de l'étude à réaliser pour éviter les extrapolations abusives.

#### V.1.4. ANALYSE DE DOCUMENTS SECTORIELS

L'analyse de la littérature technique vise davantage à mieux connaître le secteur, les technologies... Cette dernière approche peut se faire par exemple via l'analyse des BREF, l'identification et la description de technologies...

#### V.1.5. LA RECHERCHE D'EXPERTISE

Lorsque les données ACV sont manquantes, il peut être nécessaire de rechercher des expertises. Cette expertise permet d'identifier, pour les produits typiques du secteur, les principaux enjeux environnementaux, les principales contributions au cycle de vie, les technologies et leur marché, les modes d'utilisation et de fin de vie.

L'expertise est de préférence interne (l'entreprise et/ou le consultant a déjà conduit des études sur des produits / technologies du secteur) car elle est plus accessible.

Mais elle peut aussi être externe, en tout cas partiellement, centrée sur les procédés supposés clés (fortement contributeurs aux impacts globaux) pour lesquels l'expertise interne manque.

L'expertise externe se trouve auprès des parties prenantes, d'experts techniques.

Ces expertises techniques peuvent être **sectorielles** :

- Dans des documents techniques (BAT, études environnementales, approches sectorielles...) : ces documents permettent de comprendre les enjeux du secteur
- Auprès d'experts du secteur, qui vont pouvoir :
  - Décrire le processus et expliciter les principaux enjeux de consommations d'intrants, d'énergie, et d'émissions
  - Fournir les types de variation possibles des technologies et des paramètres clés, au regard du but et du champ de l'étude
  - Qualifier le cas échéant la pertinence de sélectionner certains inventaires de bases de données

Ces expertises peuvent également être **des expertises ACV** qui permettront de visualiser les principales contributions, les enjeux en matière d'incertitude et d'écart...

Selon le moment où se situe la mobilisation de l'expertise, elle peut aider :

- À mieux qualifier les enjeux environnementaux du secteur (émissions, intrants, consommations) et de ce fait à déterminer les flux élémentaires pour lesquels une imprécision forte génèrera une incertitude forte sur les résultats globaux.
- À identifier les « hot spots » de l'ACV
- À déterminer les plages de variation et niveaux d'incertitudes de certaines données primaires (ex : plage de variation des consommations d'énergie dans un secteur).

Dans ce cas, le praticien conduit une modélisation rapide, mesure les ordres de grandeur, identifie comment certaines approximations ou certains éléments négligés influencent les résultats.

*Exemple : lors de la modélisation d'une bouteille PET les commanditaires et auteurs s'interrogent souvent sur le besoin en données précises d'inventaire pour les colles des étiquettes. Cette donnée est assez peu accessible, aussi on peut choisir dans une base de données un procédé « colle » assez générique, identifier dans quelle mesure celui-ci contribue substantiellement aux résultats. On peut aussi utiliser plusieurs sources si elles sont disponibles et mesurer la sensibilité aux choix de la source. Si les sources sont très différentes (différents lieux, âge, technologie) mais avec des LCI proches, il n'est pas opportun a priori de chercher des données primaires. Inversement, si les sources ont des caractéristiques similaires mais des LCI sensiblement différents, il est opportun a priori de chercher des données primaires, sauf si la contribution à l'impact est très faible.*

L'idéal est de pouvoir faire le screening avec des procédés « conservateurs » c'est-à-dire en maximisant les impacts. De ce fait, si la contribution calculée est faible, on est certain que la donnée réelle est faible également et on peut utiliser des données de moindre qualité.

Note : la base de données Impacts de l'ADEME propose au travers des outils de modélisation, des procédés par défaut dits « grossiers » dont le but est, pour une catégorie de procédés similaires (ex : différents types de bois), de disposer d'une donnée conservatrice.

## **V.2. Fonctionnalités des logiciels ACV permettant une aide à la définition de la qualité initiale**

Les logiciels de données permettent de fournir des outils pour aider à caractériser le besoin initial en données.

Plusieurs fonctionnalités sont disponibles :

### **V.2.1. VISUALISATION DES INFORMATIONS DES BASES DE DONNÉES**

En premier lieu, les logiciels permettent une visualisation des données tirées d'une ou plusieurs bases de données.

Ces logiciels restituent les différents champs d'informations des bases de données, et permettent généralement de disposer d'informations qualitatives sur la façon dont les données sont construites. Ces données sont généralement portées dans la documentation.

Par exemple :

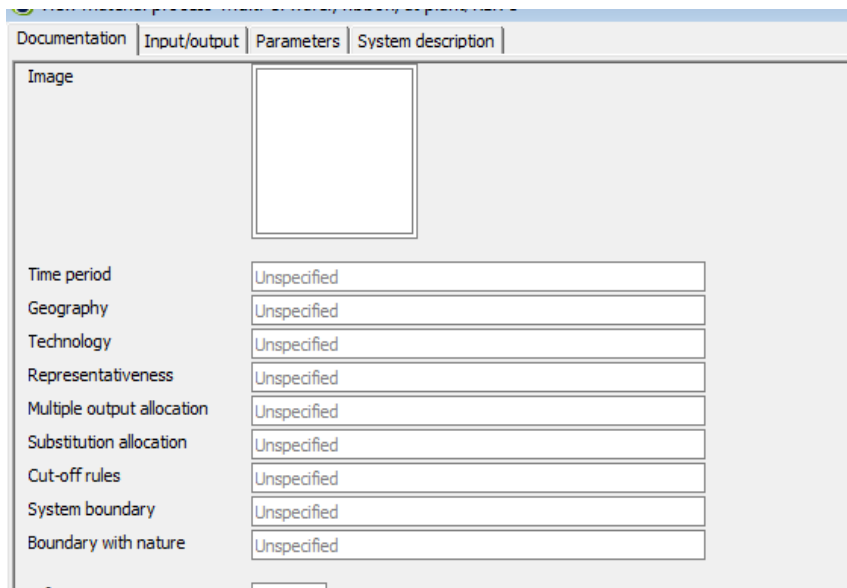
- La représentativité temporelle, géographique, technologique
- Les hypothèses (critères de cut off, frontières du système, allocations...)

Note : cette information n'est pas accessible systématiquement dans SimaPro en raison de différences entre les formats Ecospold.

Cela est par exemple encore disponible pour Buwal (1996-2000) et IDEMAT (dernière version : 2012).

Dans les versions précédentes de SimaPro (versions 7 et antérieures) ces champs étaient utilisés pour donner une vision de la qualité globale via des couleurs. Comme les « bibliothèques » de données plus récentes n'utilisent plus ces champs, cette fonctionnalité n'est plus utilisée.

**Figure 8 : Exemple de l'ancienne fonctionnalité de SimpaPro permettant la visualisation des informations des bases de données**



De plus les informations concernant les notes de qualité des données sont également visibles (voir exemple §VII.4.1.2 page 93).

## V.2.2. ANALYSE DES CONTRIBUTIONS

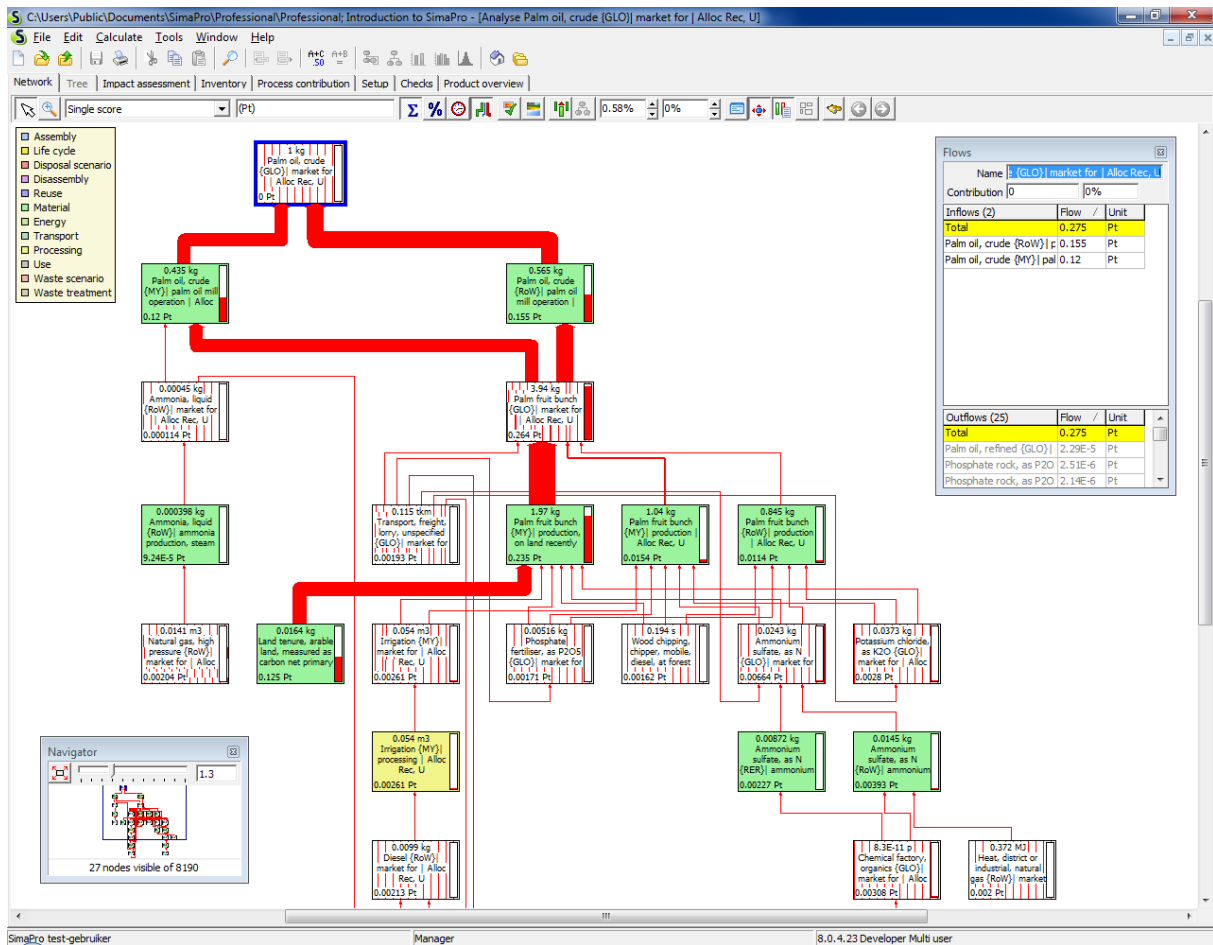
Les outils de modélisation permettent aisément de conduire des analyses de contributions. Ces analyses peuvent se présenter différemment.

Des exemples sont fournis ci-dessous :

Dans SimaPro, les fonctionnalités permettent de :

- Identifier, via le système de la largeur des flèches, les contributions pour toute la modélisation : plus la flèche est large, plus la contribution est importante
- Identifier les contributions des différents procédés via des barres sur les procédés.
- Sélectionner les procédés qui contribuent à plus de X %

Figure 9 : Diagramme avec différentes flèches et graduations pour montrer les contributions des procédés



Dans GaBi, l'approche « Weak Point analysis » permet de visualiser les différentes contributions par procédé et identifier ainsi où sont les enjeux clés : un lien est fait entre les procédés et les flux élémentaires

**Figure 10: Restitution de la « Weak point analysis » (GaBi) permettant de visualiser les contributions par procédé et par flux élémentaire**

Inputs/Outputs	GLO: Multi-functionality in End-of-Life Situations	DE: PE-LD credit	DE: PE-LD virgin	EU-27: Electricity credit	EU-27: Thermal energy credit
<b>Flows</b>	<b>0,000156</b>	<b>-3,87E-005</b>	<b>0,000285</b>	<b>-2,68E-005</b>	<b>-7,31E-005</b>
<b>Emissions to air</b>	<b>0,000156</b>	<b>-3,87E-005</b>	<b>0,000285</b>	<b>-2,68E-005</b>	<b>-7,31E-005</b>
<b>Organic emissions to air (group VOC)</b>	<b>0,000156</b>	<b>-3,87E-005</b>	<b>0,000285</b>	<b>-2,68E-005</b>	<b>-7,31E-005</b>
<b>Group NMVOC to air</b>	<b>0,000138</b>	<b>-3,31E-005</b>	<b>0,000243</b>	<b>-2,02E-005</b>	<b>-6,11E-005</b>
Butane (n-butane)	1,63E-005		3,13E-005		
Ethane	1,34E-005		3,29E-005		
NMVOC (unspecified)	7,89E-005	-1,72E-005	0,000126		-2,55E-005
Propane	2,93E-005		5,29E-005		-1,35E-005
Methane	1,82E-005		4,17E-005		

Dans l'exemple ci-dessus on observe que le LDPE vierge est le procédé le plus contributeur pour les NMVOC.

Il conviendra donc de s'assurer que les données relatives à ce procédé sont de bonne qualité par rapport au but et champ de l'étude.

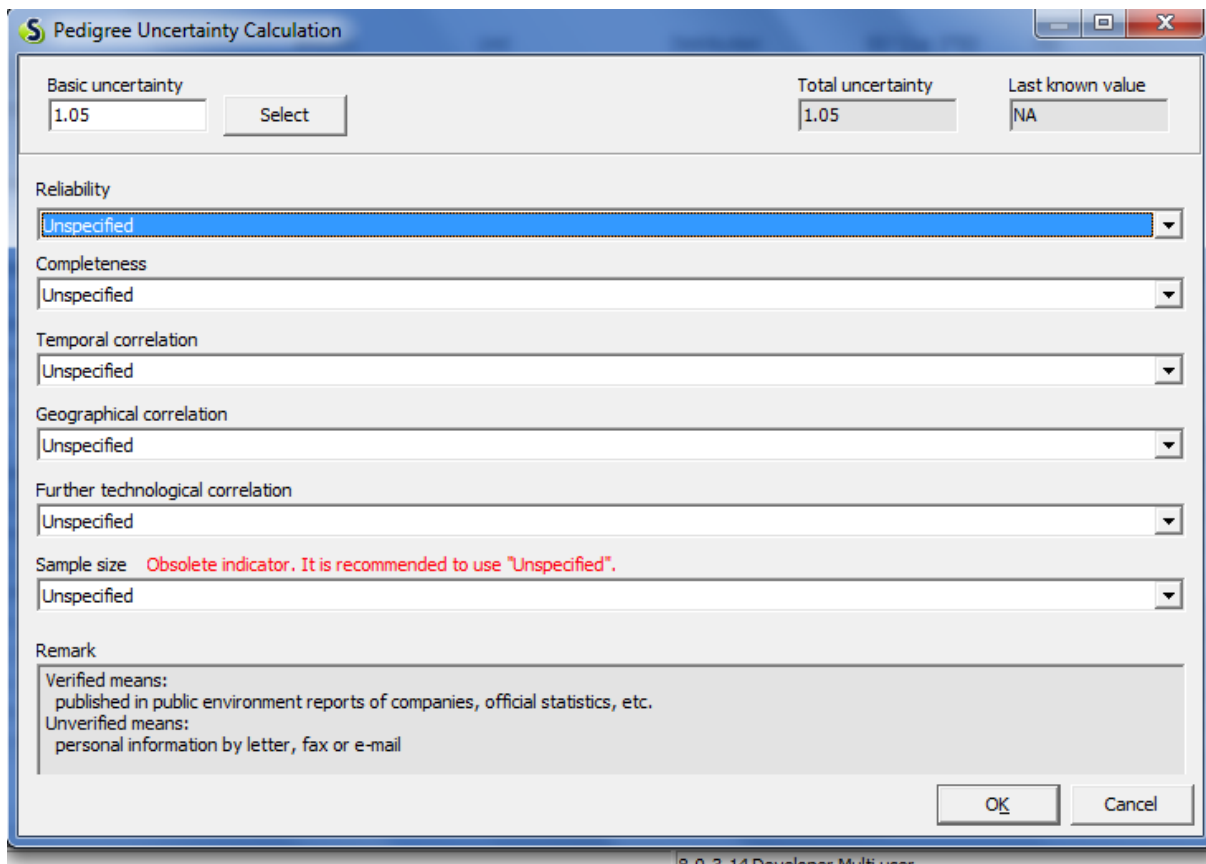
### V.2.3. QUALIFICATION DE LA QUALITÉ DES DONNÉES POUR SA PROPRE ÉTUDE

Les logiciels peuvent également permettre de qualifier soi-même la qualité des données que l'on utilise, en lien avec le but et champ de l'étude.

Par exemple, dans SimaPro, via l'outil Pedigree d'écoinvent, il est possible d'estimer l'incertitude.

L'utilisateur peut évaluer, via des menus déroulants, les différents critères qualité de sa donnée. L'incertitude est alors calculée (avec les règles d'écoinvent) pour chaque procédé, sur base de l'approche Pedigree matrice. Notons que ce travail est extrêmement fastidieux dans une logique de spécification initiale de la qualité des données.

**Figure 11 : Fonctionnalité SimaPro visant à renseigner la Pedigree matrix en qualifiant soi-même ses données pour une étude spécifique**

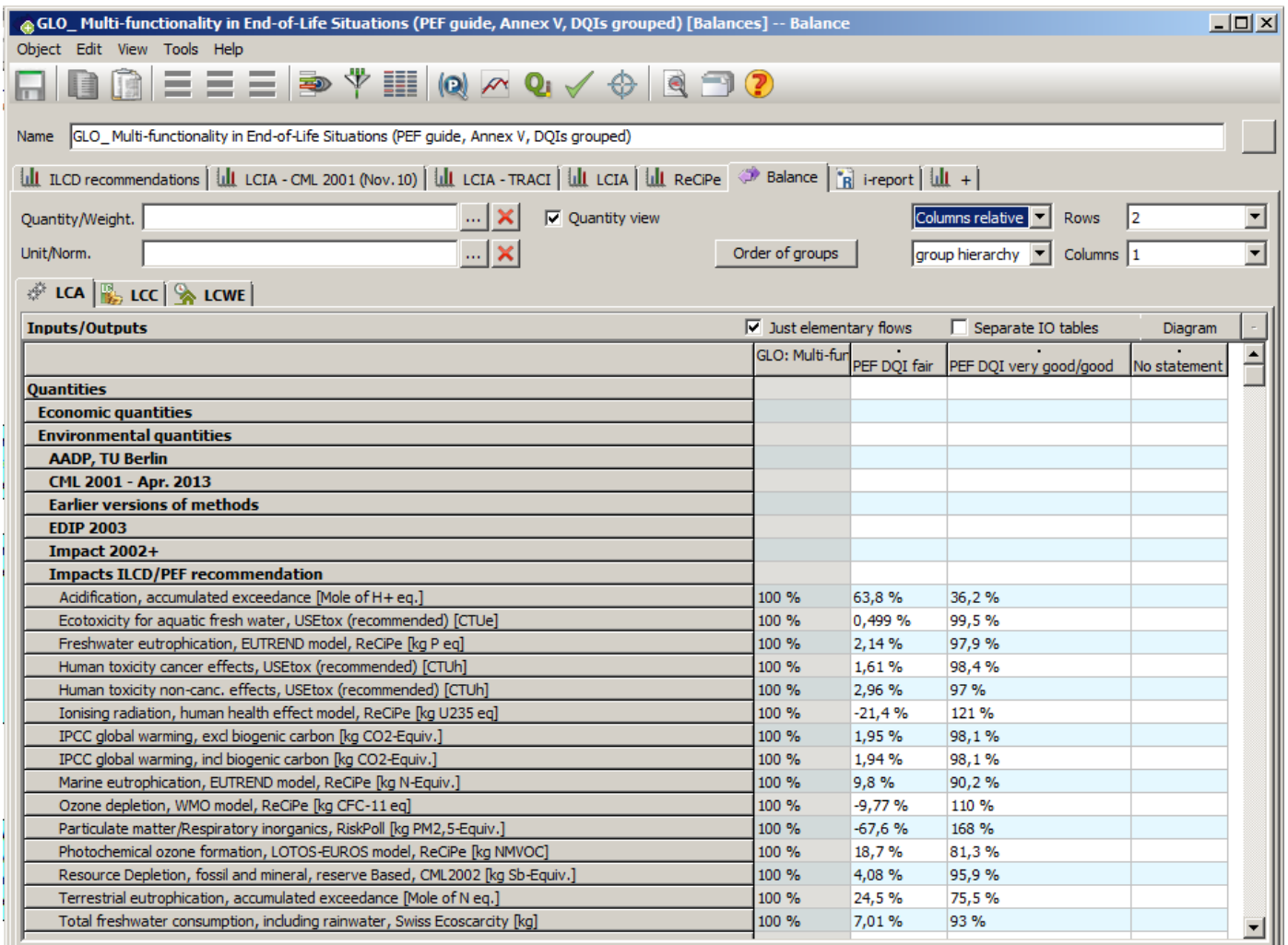


Des développements sont en cours pour également pouvoir intégrer le système de notation du PEF.

#### **V.2.4. VISUALISATION DES PARTS DE DONNÉES DE BONNE, MOYENNE ET MAUVAISE QUALITÉ**

GaBi a développé une fonctionnalité permettant de visualiser les proportions de jeux de données et leur note pour un modèle :

Figure 12 : Visualisation des pourcentages de données de bonne ou moins bonne qualité par indicateur de catégorie d'impact (GaBi)



Dans l'exemple ci-dessus on observe que pour le modèle analysé, les procédés regroupés et ayant une note « good/very good » représentent seulement 36.2 % des contributions à l'acidification. On ne respecte pas les objectifs du PEF (ce sont 70 % des contributions qui devraient avoir ce niveau de qualité).

### V.2.5. ANALYSE DE L'INCERTITUDE

Enfin, les différents logiciels permettent de réaliser des analyses d'incertitude :

- En spécifiant cette incertitude avec plusieurs probabilités de distribution (normale, log normale, triangulaire, uniforme...)
- En conduisant des analyses Monte Carlo dont le but est de déterminer, à partir de distributions sur un certain nombre de paramètres, la distribution finale des résultats.
- En développant des analyses de scénarios
- En développant une variation des paramètres
- En développant des analyses de sensibilité.

Cette approche est généralement très coûteuse en temps et est principalement utilisée de ce fait **par les scientifiques et plutôt en phase intermédiaire ou finale.**

## VI. Conclusions, Limites et Perspectives

### VI.1. La spécification initiale des besoins en qualité : identifier où mettre les efforts dans une démarche ACV

Les exigences de qualité seront très variables selon le type d'étude ACV conduite. On attend de l'ACV une réponse discriminante entre systèmes.

La définition initiale des besoins en qualité **permet de faire des choix et de déterminer là où l'on mettra et où l'on ne mettra pas les efforts pour obtenir une étude de qualité en limitant le travail**. Elle permet ainsi de réduire le temps de collecte de certaines données, voire de diminuer le nombre d'itérations.

### VI.2. La spécification initiale des besoins en qualité : un besoin d'expertise

Comme mentionné à plusieurs reprises par les praticiens, il s'agit beaucoup lors de cette phase initiale de disposer d'un **niveau d'expertise suffisant**<sup>6</sup> pour identifier les enjeux clés en matière de qualité. Cela passe par la recherche d'informations dans les ACV déjà publiées, la consultation des méta-données des procédés jugés clés, la sollicitation d'experts du secteur et, si nécessaire, une première itération grossière pour identifier les contributions principales et fixer des priorités de qualité sur ces données.

### VI.3. La recherche d'objectivité pousse à essayer d'obtenir un score de qualité, pas toujours pertinent

Le jugement d'expert inspirant parfois de la méfiance (l'expert en question était-il un « bon » expert ? est-il neutre ?), on observe une tendance à vouloir résumer ou analyser les aspects qualité par des formules mathématiques (scores globaux, lois d'incertitude).

Cela peut apparaître rassurant, mais cette note fournit intrinsèquement une évaluation de très mauvaise qualité, nettement moindre que celle de l'expert.

La raison en est que les incohérences (géographie, temps, technologie, nombre de sources) ont des influences combinées qui changent du tout au tout selon les cas.

Les outils quantitatifs sont très (trop) consommateurs de temps et trop peu spécifiques.

---

<sup>6</sup> Le commanditaire de l'étude n'est pas toujours un expert ; il peut donc être aidé par le praticien ACV, et des experts sectoriels pour cette étape de définition de la qualité initiale. Ce travail s'effectue souvent en collaboration.

C'est donc toujours à l'expert, et pas aux outils, que devrait revenir le rôle d'estimer l'incertitude relative sur base de ces métadonnées. Si des outils sont utilisés, ils peuvent aider l'expert mais ne pas fournir l'évaluation proprement dite.

Les experts consultés dans le cadre de cette étude ont mis en évidence que les approches utilisées par les outils risque d'atténuer systématiquement certains aspects de la qualité (exemple : la représentativité technologique).

## **VI.4. Une trop grande attention portée à l'incertitude par rapport à l'erreur de modélisation**

Certains guides déplorent un recours à la notion d'incertitude trop fréquent.

La précision de l'inventaire est une composante importante de la qualité, mais les aspects structurels et de modélisation des inventaires et des catégories d'impacts ont une influence en général plus grande sur les résultats. Les manques de données et hypothèses de modélisation peuvent induire des biais et affecter la précision des résultats, mais elles ne peuvent pas directement être abordées dans le calcul de l'incertitude. Les estimations d'incertitudes ne sont donc que des approximations. Elles tendent à sous-estimer la vraie incertitude mais en particulier elles ne captent pas complètement la précision des résultats obtenus.

L'une des erreurs est donc le trop grand recours aux calculs d'incertitude stochastique des données d'inventaire pour démontrer l'importance des différences entre les systèmes comparés.

L'incertitude globale et la précision permettent de juger également les autres composants de la qualité (hypothèses, méthodes, cohérence, faible représentativité des données...)

## **VI.5. Le manque d'interopérabilité entre les bases de données**

La pratique fréquente actuelle est de n'utiliser qu'une seule base de données pour garantir une certaine cohérence méthodologique (même modélisations de la fin de vie, des mix électriques...).

Cependant le manque de disponibilité de données dans certaines bases pourrait requérir de combiner des données issues de plusieurs bases de données, plus spécifiques.

Or il ne sert à rien à un utilisateur de disposer de données d'excellente qualité si l'utilisateur ne sait pas les importer dans son logiciel, ou si les flux utilisés sont incompatibles (pas de mapping existant).

Ce mapping est crucial car certains flux doivent parfois être « répartis » entre plusieurs flux d'une autre base de données (*ex : cas des flux de minerais « ore mix » devant être répartis entre différents types de minerais « cuivre » , « plomb »...*).

En effet, il peut exister des problèmes de nommages de flux (*ex : carbon dioxide dans une base versus CO2 dans une autre*), d'unité, nécessitant des conversions plus ou moins simples (*ex : le passage de g à kg est aisé, mais il s'agit parfois de passer de kg de fuel à des MJ et dans ce cas il convient de connaître la densité du fuel*).

Les enjeux de l'interopérabilité consistent donc en premier lieu à associer les bons flux élémentaires entre eux (mapping) et à vérifier qu'ils fonctionnent correctement avec les méthodes de caractérisation associées.

Cependant l'interopérabilité va plus loin et pose aussi la question de la nomenclature des informations, y compris pour les champs de données de contexte (sur les données de qualité notamment). Si les champs ne sont pas cohérents, un utilisateur peut parvenir à visualiser ces informations en lisant tout, cependant plus ces informations sont harmonisées, plus aisée sera l'analyse de l'utilisateur.

Enfin l'interopérabilité pose aussi la question des modalités de combinaisons de procédés les uns avec les autres, et pour cela il convient également de disposer d'une manière rigoureuse de pouvoir évaluer la qualité, mais surtout les modes de construction des données, leur cohérence méthodologique, les modalités de révision.... Ces enjeux sont donc fondamentaux pour disposer au final de données construites avec qualité.

## **VI.6. L'apport possible de la revue critique dans l'évaluation initiale de la qualité**

Actuellement les revues critiques sont généralement réalisées en fin d'étude, après plusieurs itérations.

La revue critique conduit donc une analyse de la qualité des données mais se base sur la qualité finale des données, non sur la spécification initiale des données.

En discutant avec certains praticiens ayant réalisé des revues critiques séquencées au fur et à mesure de l'étude, il apparaît qu'il est pertinent, lors de la revue notamment du but et champ de l'étude, de conduire une revue de la spécification initiale de la qualité des données.

Le réviseur pourrait ainsi être consulté sur la pertinence de l'approche proposée par l'auteur de l'étude.

## VII. Annexes

### VII.1. Annexe 1 : définition des principaux critères de qualité

#### VII.1.1. REPRÉSENTATIVITÉ

La représentativité indique **comment les données d’inventaire collectées représentent correctement le «vrai» inventaire** pour lequel on a collecté des données, sur le plan géographique, technologique et temporel.

Elle concerne donc la représentativité **de la construction du jeu de données**.

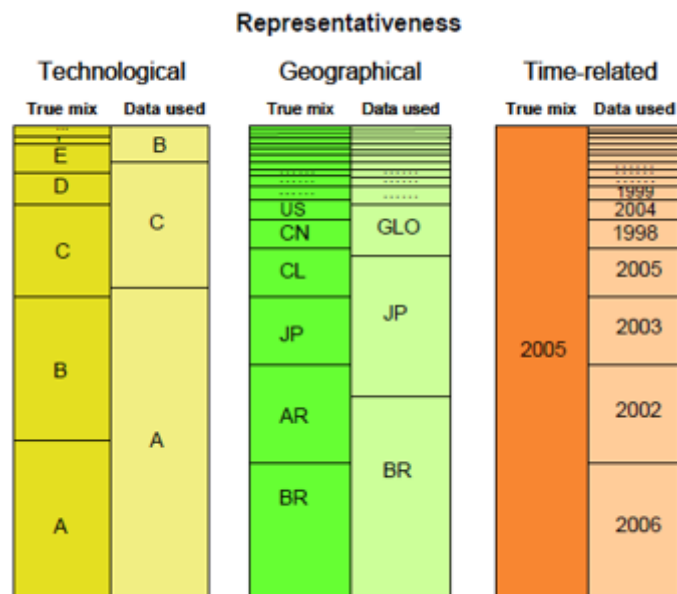
Le schéma ci-dessous représente cette notion :

*A gauche, chaque histogramme représente la donnée réelle, que l’on modélise. A droite, chaque histogramme représente la donnée modélisée.*

*On observe ainsi par exemple qu’alors que les impacts sont principalement causés par 1/4 de technologie A et 1/4 de technologie B, la modélisation retenue considère plutôt que la moitié des impacts sont causés par la technologie A et 1/8 par la technologie B.*

*De même pour la géographie, alors que la plus grande part des impacts est causée par des procédés basés au Brésil, en Argentine et au Japon, les données utilisées représentent principalement des données brésiliennes et japonaises :*

**Figure 13 : Illustration de la qualité**



**Source : ILCD Handbook, General Guide for Life Cycle Assessment – Detailed guidance, figure 28 p 325**

De façon plus large, lorsque l'étude ACV est construite sur la base d'un modèle ACV composé de plusieurs jeux de données tirés de bases de données, la représentativité peut également évoquer la justesse de la sélection d'un jeu de données dans le système modélisé pour représenter correctement la donnée

Elle concerne dans ce cas **la représentativité du jeu de données sélectionné**.

*Exemple : On veut modéliser un papier de bureau danois de 2005 et on utilise un papier journal danois de 2006 : sur le plan de la représentativité géographique, cela semble correct, de par l'année (car les technologies de l'industrie papetière évoluent peu) mais le choix est moins approprié concernant la représentativité technologique.*

Les différentes composantes de la représentativité et les éléments clés sur lesquels porter son attention sont présentées ci-dessous.

### VII.1.1.1. Représentativité géographique

La représentativité géographique fait appel à plusieurs notions : la zone géographique retenue pour la modélisation est-elle représentative de l'objet de l'étude ?

Plusieurs éléments sont à considérer :

- **Identifier ce que la zone géographique représente en termes de marché**

Il est important de distinguer dans l'objectif si on vise **un marché de production ou de consommation**.

*Exemple 1 : si on étudie la production de papier en France, cela ne signifie pas que toutes les pâtes à papier sont d'origine française, car il arrive que certaines papeteries françaises importent des pâtes scandinaves.*

*Exemple 2 : parler d'un acier européen quand on parle de la consommation d'acier en Europe n'a pas tellement de sens. En effet, l'acier consommé en Europe est produit un peu partout dans le monde.*

- Précautions à prendre en cas de **manque de données précises** pour la zone géographique recherchée

Il est parfois difficile de disposer des données précises issues d'un pays ou d'une zone géographique donnée. Des ajustements des données peuvent alors être nécessaires. Les précautions à prendre consistent à s'assurer, avant d'utiliser une donnée d'inventaire d'un autre pays, que les différences d'impacts sont faibles.

Une pratique régulièrement effectuée consiste à n'ajuster que les données liées au mix électrique pour les consommations. Cependant, ce qui va peut-être avoir davantage d'influence va être le sourcing de certains intrants, pour lesquels il conviendrait également d'ajuster les données. Si des correctifs ne sont pas effectués, l'incertitude peut être bien plus grande pour ce point.

*Exemple : dans le cas d'un produit comme de la viande, les procédés de transformation peuvent nécessiter des ajustements en termes de mix électriques, cependant, les types d'élevage vont avoir sensiblement plus d'influence sur le bilan environnemental (élevage extensif versus élevage intensif, type d'alimentation...).*

- Complétude des données ayant **des effets locaux**

Enfin, la représentativité géographique doit aussi prendre en considération **la spatialisation de certains facteurs de caractérisation** (ex : utiliser de l'eau ou des ressources minérales à différents endroits du globe n'a pas la même incidence environnementale). Cela doit être pris en compte par les méthodes d'évaluation environnementale. Il est donc important, le cas échéant, d'identifier les **flux élémentaires** qui pourraient être concernés et de les renseigner spécifiquement.

### VII.1.1.2. Représentativité temporelle

L'analyse de la représentativité temporelle porte sur l'effet éventuel que peut avoir une différence entre l'année de la donnée et l'année à modéliser : dans quelle mesure les données utilisées représentent correctement l'objet modélisé. Selon les types de cas, des données anciennes peuvent être représentatives ou non de la situation à modéliser.

L'analyse de la représentativité temporelle est souvent difficile, pour plusieurs raisons :

- La construction d'une donnée ACV peut utiliser **plusieurs sources de données**, d'années différentes
- Il convient de s'assurer de disposer **d'une distinction claire entre l'année de publication de la donnée, l'année de modélisation d'une donnée et l'année de collecte de la donnée**
- En général, la date indiquée pour la collecte de données prend en compte la date de la dernière feuille de données. Ainsi seules les dernières données modifiées sont enregistrées comme récentes, mais pas les collectes de données antérieures qui peuvent dans certains cas avoir un impact bien plus élevé.
- Dans les bases de données, **une date d'expiration de la donnée devrait être proposée** ; les secteurs d'activité n'évoluant pas tous aussi rapidement, certaines données peuvent encore être bonnes au bout de 10 ans tandis que d'autres seront périmées.

Remarque : Certains référentiels stipulent qu'il convient de fournir des données récentes et indiquent un niveau d'exigence (ex : la norme 15804 sur les déclarations environnementales pour les produits de construction exige des données de moins de 5 ans). Toutefois, cette notion d'âge de la donnée reste floue.

*Exemple : Pour les industries produisant des ordinateurs, les évolutions technologiques sont rapides, tandis que pour le secteur de la production de matériaux issus de l'industrie lourde (aluminium, acier, verre), les technologies évoluent plus lentement. Si l'on souhaite modéliser la situation actuelle pour ces deux industries, dans le premier cas, il*

*conviendra de disposer des données collectées les plus récentes (2013 voire 2014) tandis que dans le second cas, une donnée d'il y a 10 ans pourrait suffire.*

Concernant la représentativité temporelle, dans certains cas, **il convient de prendre en compte les variations saisonnières ou jour/ nuit.**

Selon les systèmes modélisés, l'ILCD préconise les choix suivants :

- Pour des produits de vie courte : on peut se contenter de données d'une durée de validité courte
- Pour des produits de vie longue, des solutions d'éco-conception doivent être considérées sur de plus grandes échelles de temps.
- Les ACV prospectives doivent être valides sur de grands pas de temps.
  - Dans des cas de prospective court terme : les données les plus récentes suffisent (< 5 ans)
  - Dans des cas de prospective long terme (> 5 ans) : des données prospectives, ou à défaut, les données des BAT, ou à défaut, les données les plus récentes si elles sont considérées comme n'influençant pas les résultats.

### VII.1.1.3. Représentativité technologique

La représentativité technologie est capitale et très liée aux questions temporelles et géographiques.

La représentativité technologique s'apprécie au regard de plusieurs questions :

- La technologie que l'on représente est-elle actuelle, passée ou future ?
- Doit-elle représenter une technologie bien spécifique<sup>7</sup> ou une technologie moyenne du marché<sup>8</sup> (un mix de technologies) ?

L'ILCD Handbook mais aussi d'autres documents (ex : PE dans son Data quality Guidelines) soulignent l'importance de la représentativité technologique. En effet, l'évaluation de la représentativité d'une donnée dans une base de données est parfois limitée par le praticien à l'identification d'un procédé dont l'intitulé semble convenir, sans approfondir la question technologique.

Par ailleurs, le document Gabi Database & Modelling Principles 2013 souligne que certaines différences entre technologies peuvent générer des différences d'impacts bien

---

<sup>7</sup> Dans ce cas on parle de « **foreground data** » : les données doivent être collectées directement chez le fournisseur.

<sup>8</sup> Dans ce cas, on parle de « **background data** » : on peut utiliser les données moyennes du marché.

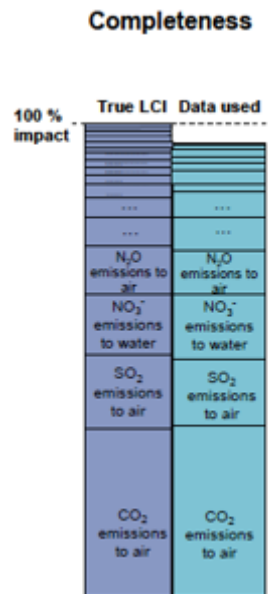
plus importantes que des différences de certains aspects géographiques, comme par exemple les mix électriques des pays. Il n'est donc pas toujours suffisant d'ajuster les données d'un jeu de données avec le mix électrique des différents pays.

### VII.1.2. COMPLÉTUDE

La complétude n'est pas toujours définie très clairement par les textes de référence. La norme ISO 14044 indique qu'il s'agit du pourcentage de flux élémentaires qui sont pris en compte par rapport à ceux qui **devraient être pris en compte** pour modéliser le cas réel.

Le schéma ci-dessous explicite comment la notion de complétude peut être appréhendée :

**Figure 14 : Illustration de la complétude**



**Source : ILCD Handbook, General Guide for Life Cycle Assessment – Detailed guidance, figure 28 p 325**

En théorie, tous les flux élémentaires qui sont attribuables au système sont à prendre en compte dans les limites du système. Cependant **tous ne sont pas pertinents quantitativement en terme de contribution aux impacts.**

Pour les moins pertinents, des données de plus faible qualité (estimations) peuvent être utilisées, ce qui limite l'effort de collecte ou l'obtention de données de grande qualité pour ces parties.

Parmi ces flux et procédés, **les non-pertinents peuvent être complètement «coupés»** (et l'effort qui serait nécessaire pour collecter ces données peut être utilisé pour se focaliser sur l'obtention de données de meilleure qualité sur les procédés et flux les plus pertinents).

La complétude de 100 % est toujours une approximation car si on connaissait l'impact total exactement, il n'y aurait pas besoin de le couper ; mais l'inventaire total n'est jamais connu, c'est toujours une approximation. Il est important que la complétude soit grande (le cut-off soit petit) car le manque de complétude sous-estime les impacts et limite la possibilité de mener des comparaisons. De plus, si le cut-off est trop grand, l'incertitude globale augmente.

**Notons qu'à moins qu'il ne soit défini initialement par des exigences intrinsèques (ex : un référentiel spécifiques, une norme...), le cut-off ne peut être approché que grossièrement dans la phase initiale de définition du champ.**

### VII.1.2.1. Cohérence méthodologique

Pour assurer la qualité des résultats, **toutes les hypothèses doivent être faites de façon cohérente pour les différentes parties du système.** Cela est en particulier fondamental pour les comparaisons.

**Ces méthodes doivent de plus être menées de façon uniforme en accord avec le but et champ de l'étude.**

Ces éléments de cohérence peuvent porter sur des aspects plus ou moins triviaux :

- Utilisation de la même nomenclature de flux dans tout le système, des mêmes méthodes de caractérisation des impacts.
- Approche conséquentielle
- Prise en compte des bénéfices du recyclage...

Lorsque les données s'appuient sur des inventaires de base de données, il convient cependant de vérifier l'importance de l'impact des jeux de données dans le résultat global. Ainsi on peut s'accommoder, le cas échéant de certaines cohérences méthodologique si elles n'ont pas d'influence sur les résultats.

*Exemple : dans une ACV de production de limonade en bouteille PEHD, l'allocation des bénéfices du recyclage du PEHD est une allocation au système qui produit la matière recyclée. Pour le lavage des machines, des bouteilles en PEHD de détergent sont également utilisées, puis collectées et recyclées. Comme l'impact de la production de ces bouteilles de détergent est faible par litre de limonade, on peut retenir un autre type d'allocation sur le recyclage de la bouteille en PEHD.*

En résumé, lors de la définition du champ de l'étude, des efforts doivent être faits pour assurer la cohérence méthodologique pour tous les procédés contributeurs. Cette cohérence doit être vérifiée lors de l'évaluation dans la phase d'interprétation.

### VII.1.3. INCERTITUDE

Les 3 principales sources d'incertitudes sont :

- L'incertitude stochastique (liée au hasard, statistique)

Elle peut être identifiée par une mesure de la moyenne, de la variation autour de la moyenne et d'une information sur le type de distribution.

- L'incertitude de choix

Cette approche est de forme discrète (une option ou une autre). Les choix concernent les principes de modélisation, les méthodes d'inventaires (normalisation et pondération), les décisions de cut off, les choix des inventaires, les choix des catégories d'impacts, les autres hypothèses. Il n'y a pas de distribution statistique associée, en général on raisonne en scénario. Les principaux choix ont le potentiel d'influencer l'incertitude des résultats finaux. L'approche pour les identifier est de calculer les différents scénarios possibles.

- Le manque de connaissance

Cela n'est pas chiffré sur le plan quantitatif mais plutôt via une revue critique.

#### VII.1.3.1. Interactions entre critères : complétude et incertitude

Les différents critères de qualité interfèrent. Ainsi, on peut obtenir une complétude de 90 % avec des données de haute qualité ou des données estimées. Mais si on utilise des données estimées, la référence à 100 % des impacts est moins précise. Donc quand on définit la complétude, cela doit être regardé avec le niveau de qualité minimum requis pour le but et champ de l'étude et avec le niveau de qualité nécessaire pour les autres composantes de la qualité.

Par exemple, si dans le cadre d'une comparaison, tous les indicateurs de catégories d'impacts sont en faveur du même système (plus de 60 à 90 % d'écart pour chaque catégorie) et si les données disponibles peuvent être très précises car les principaux procédés sont des procédés spécifiques, et les données sont mesurées annuellement, le niveau de complétude peut alors être bas (ex : 80 %) car cela permet encore de démontrer que les 2 alternatives sont significativement différentes.

## VII.2. Annexe 2 : ce que disent les textes de références sur la qualité initiale des données et la qualité

Nous avons étudié différents textes de référence :

- Les normes ISO 14040 -44
- L'ILCD Handbook
- La Product Environmental Footprint Methodology.

Nous présentons ici succinctement les éléments citant ces besoins de qualité initiale.

### VII.2.1. LES NORMES ISO 14040-44

#### VII.2.1.1. Importance de la spécification initiale

Comme indiqué ci-dessus (cf. § **Erreur ! Source du renvoi introuvable.** p.**Erreur ! Signet non défini.** et II.3 p.15), les normes ISO 14040 et 14044 insistent sur l'importance de la qualité initiale pour répondre au but et champ de l'étude.

#### VII.2.1.2. Critères de qualité

L'ISO liste 10 critères à considérer pour la qualité des données.

- Certains sont à décrire : les « couvertures » temporelles, géographiques et technologiques ou sources de données.
- D'autres sont à caractériser de façon qualitative : la représentativité, la cohérence, la reproductibilité.
- D'autres enfin sont à estimer de façon quantitative : la variabilité des données, la complétude et l'incertitude.

<b>ISO 14044 : 2006</b> <b>(§ 4.2.3.6.2)</b>	<p><i>The data quality requirements should address the following:</i></p> <ul style="list-style-type: none"><li><i>a) <b>time-related coverage:</b> age of data and the minimum length of time over which data should be collected;</i></li><li><i>b) <b>geographical coverage:</b> geographical area from which data for unit processes should be collected to satisfy the goal of the study;</i></li><li><i>c) <b>technology coverage:</b> specific technology or technology mix;</i></li><li><i>d) <b>precision:</b> measure of the variability of the data values for each data expressed (e.g. variance);</i></li><li><i>e) <b>completeness:</b> percentage of flow that is measured or estimated;</i></li><li><i>f) <b>representativeness:</b> qualitative assessment of the degree to which the data set reflects the true population of interest (i.e. geographical coverage, time period and technology coverage);</i></li><li><i>g) <b>consistency:</b> qualitative assessment of whether the study methodology is applied uniformly to the various</i></li></ul>
---	---

	<p>components of the analysis;</p> <p>h) <b>reproducibility</b>: qualitative assessment of the extent to which information about the methodology and data values would allow an independent practitioner to reproduce the results reported in the study;</p> <p>i) <b>sources</b> of the data;</p> <p>j) <b>uncertainty</b> of the information (e.g. data, models and assumptions).</p>
--	---

La norme ISO considère les **sources de données comme faisant partie de la qualité intrinsèque des données**. Cela démontre donc l'importance de la recherche et de la sélection, dès l'étape initiale, **des « meilleures » sources d'information**. Une expertise est généralement nécessaire.

Par ailleurs, **la complétude est généralement un critère qui est considéré comme traité plus tard dans l'étude ACV et non initialement, lors de l'étape de vérification**, comme en témoigne le point suivant :

<p><b>ISO 14044 : 2006</b> <b>(§ 4.5.3.2.)</b></p>	<p><i>The objective of the completeness check is to ensure that all relevant information and data needed for the interpretation are available and complete. If any relevant information is missing or incomplete, the necessity of such information for satisfying the goal and scope of the LCA shall be considered. This finding and its justification shall be recorded.</i></p>
<p><b>ISO 14044 : 2006</b> <b>(§ annex B, B.3.2.)</b></p>	<p><i>The completeness check attempts to ensure that the full required information and data from all phases have been used and are available for interpretation. In addition, data gaps are identified and the need to complete the data acquisition is evaluated. The identification element is a valuable basis for these considerations. (...) Nevertheless, completeness can only be an empirical value, ensuring that no major known aspects have been forgotten.</i></p>

Comme précisé, disposer d'une donnée chiffrée de cette complétude est très empirique. Il semble que la complétude fait essentiellement appel à un jugement d'expert.

### VII.2.1.3. Méthodes d'évaluation des besoins en qualité initiale

Aucun critère concret n'est mentionné sur les modalités d'évaluation spécifiques de la qualité initiale.

En revanche, les éléments suivants sont détaillés car ils peuvent aider à définir en amont le besoin de qualité des données, et suggérer certaines réflexions / approches à avoir :

#### VII.2.1.3.1. Lors de la collecte des données

Lors de la collecte des données il est mentionné que pour les données qui pourraient avoir un impact sur les conclusions de l'étude, il est fondamental de collecter des informations sur la qualité des données.

Cela sous-entend donc qu'il convient d'avoir déjà une vision sur les étapes / données les plus influentes. Il n'est pas explicité comment cette évaluation est conduite.

<p><b>ISO 14044 : 2006</b> <b>(§ 4.3.2.1)</b></p>	<p><i>For those data that may be significant for the conclusions of the study, details about the relevant data collection process, the time when data have been collected, and further information about data quality indicators shall be referenced.</i></p> <p><i>If such data do not meet the data quality requirements, this shall be stated.</i></p>
---	---

	<p><i>To decrease the risk of misunderstandings (e.g. resulting in double counting (...))</i></p> <p><i>Since data collection may span several reporting locations and published references, measures should be taken to reach uniform and consistent understanding of the product systems to be modelled.</i></p>
--	--

### VII.2.1.3.2. Précisions sur les techniques additionnelles d'analyse de la qualité des données

Les techniques d'analyse de la qualité des données en cours d'étude sont citées et peuvent permettre, **lors d'une analyse bibliographique, d'identifier les parties de rapport plus spécifiques à analyser :**

Pour analyser la qualité des données, il existe différentes techniques :

- L'analyse de la gravité vise à identifier les principaux postes contributeurs.
- L'analyse d'incertitude vise à identifier l'impact des incertitudes sur les résultats.
- L'analyse de sensibilité vise à identifier l'impact des valeurs et choix méthodologiques sur les résultats.

<p><b>ISO 14044 : 2006</b> <b>(§ 4.4.4.2)</b></p>	<p><i>The specific techniques and their purposes are described below.</i></p> <p><i>a) Gravity analysis (e.g. Pareto analysis) is a statistical procedure that identifies those data having the greatest contribution to the indicator result. These items may then be investigated with increased priority to ensure that sound decisions are made.</i></p> <p><i>b) Uncertainty analysis is a procedure to determine how uncertainties in data and assumptions progress in the calculations and how they affect the reliability of the results of the LCIA.</i></p> <p><i>c) Sensitivity analysis is a procedure to determine how changes in data and methodological choices affect the results of the LCIA.</i></p>
---	--

Ces trois approches sont donc à analyser de façon attentive dans le cas de la définition des besoins en qualité.

### VII.2.1.4. Cas des ACV comparatives

Un point d'attention particulier vise les analyses comparatives.

Il est mentionné que les systèmes doivent être comparés avec les mêmes considérations méthodologiques et notamment concernant la qualité des données.

**S'il existe des incohérences entre systèmes, celles-ci doivent être regardées par rapport à l'objectif et au champ de l'étude.** Si certaines de ces incohérences ont un impact significatif sur les résultats, il convient d'en tenir compte avant d'établir des recommandations.

<p><b>ISO 14044 : 2006</b> <b>(§ 4.2.3.6.2)</b></p>	<p><i>Where a study is intended to be used in comparative assertions intended to be disclosed to the public, the data quality requirements stated in a) to j) above shall be addressed.</i></p>
---	---

<p><b>ISO 14044 : 2006</b> <b>(§ 4.2.3.7)</b></p>	<p>Systems shall be compared using the same functional unit and equivalent methodological considerations, such as performance, system boundary, <b>data quality</b>, allocation procedures, decision rules on evaluating inputs, and outputs and impact assessment. Any differences between systems regarding these parameters shall be identified and reported. If the study is intended to be used for a comparative assertion intended to be disclosed to the public, interested parties shall conduct this evaluation as a critical review.</p>
<p><b>ISO 14044 : 2006</b> <b>(§ annex B, B.3.4.)</b></p>	<p>The consistency check attempts to determine whether the assumptions, methods, models and data are consistent either along a product's life cycle or between several options. Inconsistencies are, for example:</p> <p>a) differences in data sources; e.g. Option A is based on literature, whereas Option B is based on primary data;</p> <p>b) differences in data accuracy; e.g. for Option A a very detailed process tree and process description is available, whereas Option B is described as a cumulated black-box system;</p> <p>c) differences in technology coverage; e.g. data for Option A are based on experimental process (e.g. new catalyst with higher process efficiency on a pilot plant level), whereas data for Option B are based on existing large-scale technology;</p> <p>d) differences with time-related coverage; e.g. data for Option A describe a recently developed technology, whereas Option B is described by a technology mix, including both recently built and old plants;</p> <p>e) differences in data age; e.g. data for Option A are 5-year-old primary data, whereas data for Option B are recently collected;</p> <p>f) differences in geographical coverage; e.g. data for Option A describe a representative European technology mix, whereas Option B describes one European Union member country with a high-level environmental protection policy, or one single plant.</p> <p>Some of these inconsistencies may be accommodated in line with the definition of the goal and scope. In all other cases, significant differences exist and their validity and influence need to be considered before drawing conclusions and making recommendations.</p>

Le tableau B13 de l'ISO précise et illustre cet élément :

Table B.13 — Result of a consistency check

Check	Option A		Option B		Compare A and B?	Action
	Literature	OK	Primary	OK		
Data source	Literature	OK	Primary	OK	Consistent	No action
Data accuracy	Good	OK	Weak	Goal and scope not met	Not consistent	Revisit B
Data age	2 years	OK	3 years	OK	Consistent	No action
Technology coverage	State-of-the-art	OK	Pilot plant	OK	Not consistent	Study target - no action
Time-related coverage	Recent	OK	Actual	OK	Consistent	No action
Geographical coverage	Europe	OK	USA	OK	Consistent	No action

Ainsi le tableau montre bien que seules les données pour lesquelles la précision est insatisfaisante au vu des objectifs de l'étude sont à revoir. Ainsi, pour l'objectif de l'étude, les données technologiques reflètent, pour la situation B une unité pilote, tandis qu'elle reflète pour A une situation de marché, mais cela n'est pas gênant étant donné le but de l'étude.

## VII.2.2. L'ILCD HANDBOOK

### VII.2.2.1. Importance de la spécification initiale

La spécification initiale des besoins de qualité est citée dès la définition du but et du champ de l'étude :

<b>§ 6.1. page 51</b>	<i>Next and main part of the scope definition is to derive the requirements on methodology, quality, reporting, and review in accordance with the goal of the study, i.e. based on the reasons for the study, the decision-context, the intended applications, and the addressees of the results.</i>
-----------------------	---

### VII.2.2.2. Critères de qualité

#### VII.2.2.2.1. Critères

Les critères de qualité de l'ILCD Handbook sont :

- La précision (représentativité, adéquation méthodologique et cohérence)
- L'incertitude intrinsèque
- La complétude de l'inventaire

<b>§ 6.9.2. page 136</b>	<i>Data quality is composed of accuracy (i.e. representativeness and methodological appropriateness and consistency), precision / uncertainty and completeness of the inventory.</i>
--------------------------	--

Un second aspect qui est cité pour la qualité des données et qui est complémentaire porte sur des aspects qui sont clés, notamment pour la sélection/ l'identification de données de qualité :

- La documentation et notamment son caractère approprié en terme de :
  - Couverture
  - Forme
  - Format de données et de rapport
- Leur revue : l'évaluation vise à savoir :
  - Si le type de revue appliquée est approprié
  - Si la revue est appliquée correctement en terme
    - De champ,
    - De méthode
    - Et de documentation
- Les efforts de cohérences (principalement sur la nomenclature et la méthodologie)

Ces efforts peuvent porter, notamment au niveau d'une base de données sur :

- Nomenclature et liste de flux élémentaires
- Les autres conventions (ex : nommage des fichiers...)
- Format informatique
- ...

<b>§ 12.1. Annexe A page 323</b>	<i>Data quality is composed of accuracy (i.e. representativeness and methodological appropriateness and consistency), precision / uncertainty and completeness of the inventory.</i>
--------------------------------------	--

#### VII.2.2.2.2. Qualité globale

L'ILCD insiste sur l'importance de l'ensemble des critères de qualité dans la vision d'une donnée ACV :

<b>§ 6.9.2. page 136</b>	<i>LCA results can be called valid ("of high overall data quality") if they are both accurate, precise, AND complete. The weakest of the criteria generally weakens the overall quality of the specific case. This is reflected by the ILCD data quality indicators. In LCA one can hence use the term "validity" to refer to the overall quality of the data (and the results of LCA studies).</i>
--------------------------	---

La qualité globale d'un jeu de données doit explicitement préciser les différents éléments de qualité des données décrits.

Un positionnement sur la qualité globale, non revue / vérifiée par un expert indépendant, ne doit pas être pris en considération pour évaluer la qualité des données.

**La qualité globale d'un jeu de donnée suit généralement la logique du maillon faible de la chaîne. C'est la base de l'approche considérée par l'ILCD, par opposition à d'autres approches comme celles du PEF, se basant sur une valeur moyenne.**

Si tous les critères de qualité sont excellents mais que l'un d'entre eux est très mauvais (par exemple la représentativité technologique – cela revient à utiliser par exemple les données d'éolien alors qu'on s'attend à considérer de l'électricité à base de charbon), alors cela conduit à une donnée de basse qualité pour le jeu de données. Une moyenne arithmétique ne permettrait pas d'identifier ces effets.

Le détail de la façon dont la base de données ELCD peut calculer cette qualité globale et la renseigner est mentionné au chapitre VII.3.1.1 p. 74.

#### VII.2.2.2.3. Recommandations et alertes sur la prise en compte des critères de qualité

L'ILCD estime que la recherche de **la qualité doit en premier lieu viser le caractère approprié de la modélisation au but et champ de l'étude.**

<p><b>§ 12.2. page 327</b></p>	<p><i>Procedurally, one can effectively work towards high quality data, but first precisely identifying the technological, geographical and time-related appropriateness, i.e. what the data set should represent. Next, completeness of the related inventory in coverage all to-be-included and relevant impact categories is aimed at. In quantifying the flows, paying attention to low variance of the values completes the approach.</i></p>
--------------------------------	--

L’ILCD mentionne par ailleurs **l’importance trop grande apportée aux erreurs stochastiques dans le calcul d’incertitude globale**, et que de ce fait on néglige parfois trop la précision liée à l’objet modélisé, ce qui peut entraîner des biais fondamentaux.

<p><b>§ 12.2. page 328</b></p>	<p><i>It is an increasingly found error to only consider the (known or estimated) quantitative stochastic inventory data uncertainty only and directly use this to demonstrate significance of differences in compared systems. The overall precision and accuracy however needs to also judge the other, structural components, assumptions, method appropriateness and consistency, limited representativeness of data, and the like. If only a partial analysis is done, it shall be clearly stated that the other part is lacking. In addition, it shall be clearly stated how accuracy and precision have been determined if one or both have been quantified. It is argued that in practice lack of accuracy is the more relevant problem than stochastic data uncertainty, why especially a lack of addressing the former while reporting the later can be understood as an attempt of misleading the target audience. This is even more so as lack of accuracy typically introduces a bias into the results. The judgement of the overall data quality can ultimately only be done by expert judgement. Uncertainty calculations and qualitative or quantified accuracy assessment can substantially help but provide supporting, quantitative information only.</i></p>
--------------------------------	--

### VII.2.2.3. Méthodes d’évaluation des besoins en qualité initiale

L’ILCD Handbook fournit quelques recommandations et réflexions qu’il est capital de prendre en compte pour l’évaluation des besoins initiaux en qualité.

#### VII.2.2.3.1. L’évaluation de la qualité, l’importance du jugement d’expert

L’ILCD Handbook insiste sur le fait que la **qualité globale consiste davantage en un jugement d’expert** qu’en une information uniquement quantitative, qui elle, consiste davantage en un support à l’évaluation.

<p><b>§ 12.2. page 328</b></p>	<p><i>The judgement of the overall data quality can ultimately only be done by expert judgement. Uncertainty calculations and qualitative or quantified accuracy assessment can substantially help but provide supporting, quantitative information only.</i></p>
--------------------------------	---

Dès lors, l’enjeu, dans le cadre de la spécification initiale de la qualité visera à mobiliser ces expertises plus qu’à disposer d’une valeur très précise d’incertitude.

#### VII.2.2.3.2. Analyse des flux et procédés les plus contributeurs pour déterminer l’exigence de qualité (p136)

Il convient de définir d’abord les exigences de qualité des jeux de données. Ensuite seulement ces exigences sont transcrites au niveau des flux élémentaires. Ce point est généralement fait **après la première itération, et non au début de l’ACV.**

Cette traduction requiert de connaître les facteurs de caractérisation des flux.

Il convient **donc de regarder les procédés contributeurs par flux élémentaires, de façon à identifier les enjeux sur lesquels il faut des données de bonne qualité.**

Cela demande donc une analyse initiale du cycle de vie, en pratique, c'est souvent un nombre très limité de flux et de procédés qui contribuent de façon significative aux impacts globaux.

#### VII.2.2.3.3. Utiliser des procédés existants, grossiers et le jugement d'expert (Chap 4 p 25)

Pour obtenir **la précision avec un minimum d'effort, il est recommandé de collecter les données d'une façon itérative.**

La première itération peut utiliser des données génériques ou moyennes pour les données de marché (background data) et même en partie pour des données spécifiques (foreground data).

Elles peuvent être combinées avec la notion de **jugement d'expert** pour identifier les procédés clés et les flux élémentaires clés. Ainsi la collecte sera focalisée sur les parties les plus pertinentes du système.

#### VII.2.2.3.4. S'appuyer sur des expériences antérieures (§7.3 résumé page 189)

En effet, ces expériences sont utiles :

- Soit parce qu'elles ont identifié là où sont les principales contributions,
- Soit parce qu'elles ont carrément défini un niveau d'exigence sur la qualité des données.

Par exemple, les PCR peuvent parfois fixer précisément des exigences en matière de complétude ou d'incertitude, sur base d'ACV réalisées de façon antérieure.

*Exemple d'EPD fixant un niveau d'exigence : At least 95 % completeness of overall environmental impact and maximum variance on inventory data lower than 10 % for Climate change and Primary energy, 25 % for Acidification, Eutrophication and Summer Smog impact potentials."*

## VII.2.3. LA PEF METHODOLOGY (2014)<sup>9</sup>

### VII.2.3.1. Importance de la spécification initiale

La spécification initiale de la qualité est mentionnée indirectement, au travers d'une étape que la méthodologie PEF recommande vivement, à savoir le « **screening** ». Cette étape permet de **cibler les activités de collecte des données et les priorités en matière de qualité** des données en vue de l'établissement du profil d'utilisation des ressources et d'émissions (Resource Use and emission Profile<sup>10</sup>)

### VII.2.3.2. Critères de qualité

#### VII.2.3.2.1. Critères

Les critères de qualité du PEF reprennent les critères de l'ILCD : (cf §VII.2.2.2.1 p63).

**Tableau 3** : Critères de qualité défini par la Méthodologie PEF

Data quality criteria, documentation, nomenclature and review	
Data quality criteria	<ul style="list-style-type: none"> <li>— Technological representativeness <sup>(1)</sup></li> <li>— Geographical representativeness <sup>(2)</sup></li> <li>— Time-related representativeness <sup>(3)</sup></li> <li>— Completeness</li> <li>— Parameter uncertainty <sup>(4)</sup></li> <li>— Methodological Appropriateness and Consistency <sup>(5)</sup> (the requirements as defined in Table 7 shall apply until end of year 2015. From 2016, full compliance with the PEF methodology will be required)</li> </ul>
Documentation	<ul style="list-style-type: none"> <li>— Compliant with ILCD format</li> </ul>
Nomenclature	<ul style="list-style-type: none"> <li>— Compliant with ILCD nomenclature (e.g. use of ILCD reference elementary flows for IT compatible inventories)</li> </ul>
Review	<ul style="list-style-type: none"> <li>— Review by "Qualified reviewer" (see chapter 8):</li> <li>— Separate review report</li> </ul>

**Source : Commission recommendation of 9 April 2013 on the use of common methods to measure and communicate the life cycle environmental performance of products and organisations, §5.6 tableau 3**

<sup>9</sup> Sur base des documents disponibles en 2014 soit : « Commission Recommendation of 9 April 2013 on the use of common methods to measure and communicate the life cycle environmental performance of products and organizations ». **Notons que ces approches pourraient évoluer à l'avenir.**

<sup>10</sup> Le Profil de Ressources et d'émissions comprend tous les matériaux et ressources énergétiques nécessaires et les émissions dans l'air, l'eau et le sol pour l'ensemble de la supply chain du produit utilisé.

Ces critères sont évalués différemment selon :

- La contribution des données aux impacts :  
Pour les données les plus contributrices l'évaluation doit être quantitative. Pour des données moins contributrices, on peut s'accommoder d'un jugement d'expert.
- Le type d'étude : les niveaux d'exigence seront différents pour une étude de « screening » et pour une étude répondant pleinement aux exigences de la PEF.

Un PEFCR (règles fixées pour estimer les impacts d'une catégorie de produits) fixera des exigences relatives à une catégorie de produit données, de façon plus figée.

Note : les PEFCR sont en cours d'élaboration

### VII.2.3.2.2. Qualité globale

Une formule est proposée pour calculer la qualité.

**Tableau 4 : Calcul de la qualité globale (méthode semi-quantitative) utilisé par la méthodologie PEF**

<p>Formula 1 <math display="block">DQR = \frac{TeR + GR + TiR + C + P + M}{6}</math></p> <p>— DQR: Note de qualité des données de l'ensemble de données</p> <p>— TeR: Représentativité technologique</p> <p>— GR: Représentativité géographique</p> <p>— TiR: Représentativité temporelle</p> <p>— C: Exhaustivité</p> <p>— P: Précision/incertitude</p> <p>— M: Pertinence et cohérence méthodologiques</p>
--

**Source : Commission recommendation of 9 April 2013 on the use of common methods to measure and communicate the life cycle environmental performance of products and organizations, §5.6 tableau 6**

Cette approche considère tous les critères sur un pied d'égalité sur une base de moyenne arithmétique des notes. Certaines autres méthodes d'évaluation de la qualité globale, et notamment l'ILCD fonctionnent autrement en considérant un poids supérieur pour le critère ayant la note la plus faible. (cf. Chap VII.2.2.2.2 p64).

Au regard de la note obtenue, on peut indiquer une note de qualité :

**Tableau 5 : Classification des notes de qualité et appréciation selon la méthodologie PEF**

Overall data quality level according to the achieved data quality rating	
Overall data quality rating (DQR)	Overall data quality level
≤ 1,6	"Excellent quality"
1,6 to 2,0	"Very good quality"
2,0 to 3,0	"Good quality"
3 to 4,0	"Fair quality"
> 4	"Poor quality"

**Source : Commission recommendation of 9 April 2013 on the use of common methods to measure and communicate the life cycle environmental performance of products and organizations, §5.6 tableau 6**

### VII.2.3.3. Exigences requises

#### VII.2.3.3.1. Approche qualitative ou semi quantitative selon l'impact des données

Le PEF fixe des exigences de qualité minimale des données en fonction de la contribution des données aux impacts totaux :

**Tableau 6 : Niveau d'exigence sur la qualité des données en fonction de la contribution des procédés pour une étude complète<sup>11</sup>**

	Exigence minimale de qualité	Type d'évaluation
Données couvrant au moins 70 % des contributions à chaque catégorie d'impact du PEF	Qualité globale des données «GOOD» (DQR ≤ 3,0)	Semi-quantitative
Données concernant 20 à 30% des contributions à chaque catégorie d'impact du PEF	Qualité globale des données «FAIR»	Avis qualitatif d'expert Aucune quantification requise
Données utilisées pour les approximations et pour combler les lacunes (pas plus de 10 % des contributions à chaque catégorie d'impact du PEF)	Meilleures données disponibles	Avis qualitatif d'expert

**Source : Commission recommendation of 9 April 2013 on the use of common methods to measure and communicate the life cycle environmental performance of products and organizations, §5.6 tableau 4**

<sup>11</sup> Dans le cas d'une étude de type « screening », le niveau d'exigence requis pour la qualité est de « FAIR » au moins pour les données couvrant au moins 90 % des contributions à chaque catégorie d'impact.

- Pour les données couvrant au moins 70 % des contributions à chaque catégorie d'impact, un niveau de qualité « au moins bon » est requis. Le mode d'évaluation de la qualité est alors semi-quantitatif.

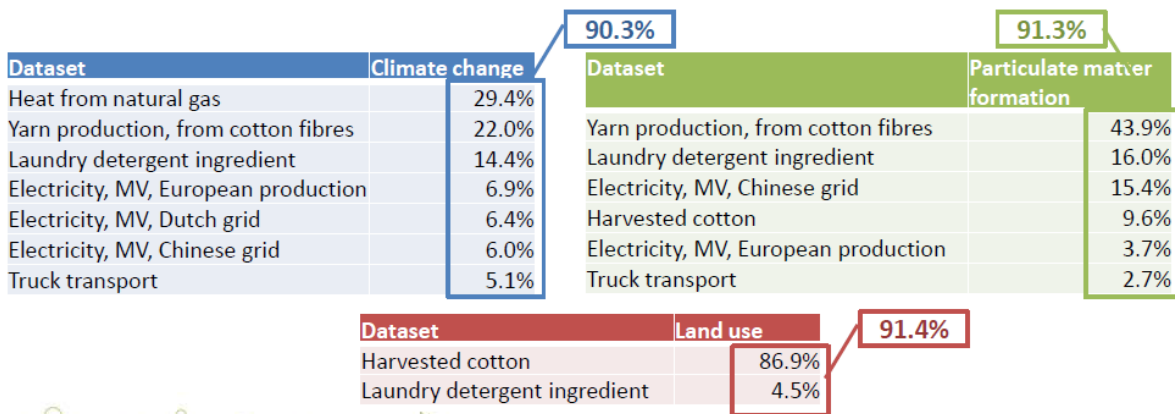
Note : Le seuil de 70 % a été défini par la Commission européenne comme un seuil permettant de concilier la robustesse des données et la faisabilité de l'évaluation de la qualité.

- Pour les données représentant 20 à 30 % des contributions, le niveau d'exigence est de disposer d'une qualité « FAIR ».
- Pour les données utilisées pour combler des lacunes, elles doivent être les meilleures possibles.

Remarque : **Cette approche peut manquer de se concentrer sur les procédés clés** : en effet, la formation proposée par le helpdesk de la CE sur l'évaluation de la qualité des données propose la démarche suivante :

- Identifier, pour chaque catégorie d'impact étudiée, les données contribuant à plus de 90 % (en les classant)

**Figure 15 : Exemple d'analyse de la qualité pour une étude d'ACV screening selon la méthode PEF– Etape d'identification des données couvrant au moins 90 % des contributions aux catégories d'impacts**



Source : Webinar – Data Quality in the context of the EF pilot phase of the European Commission DG ENV (18 July 2014)

Dans le schéma proposé ci-dessus, on a identifié, pour chaque catégorie d'impact (en bleu : le changement climatique, en vert : la formation de particules et en rouge : l'utilisation des terres), les procédés les plus contributeurs, en les classant du plus contributeur au moins contributeur. L'ensemble de ces procédés représente plus de 90 % des impacts pour chaque catégorie. On est dans une étude de type screening, il faut que l'ensemble de ces données aient un niveau de qualité « FAIR3 ».

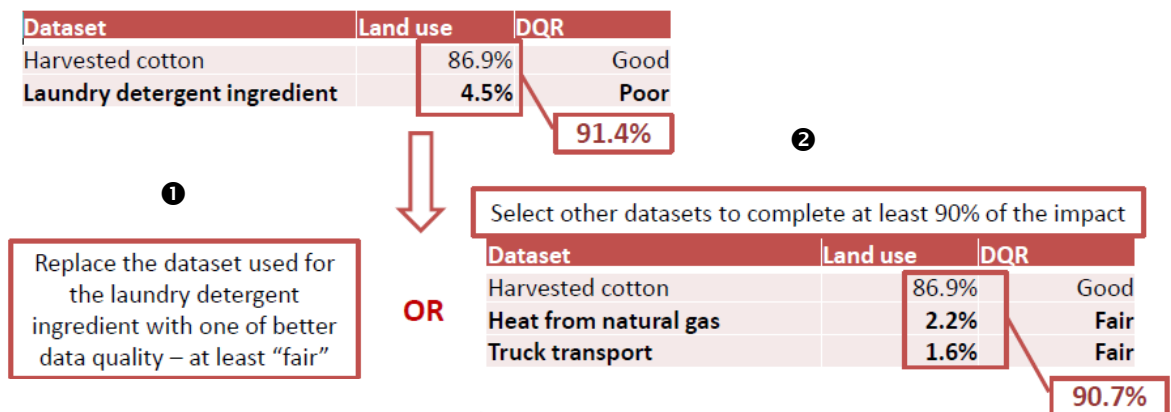
- Vérifier la qualité des données pour chacun de ces procédés.

Dans cet exemple, on est dans une étude de type screening, il faut que l'ensemble de ces données contribuant à plus de 90 % des catégories d'impacts aient un niveau de qualité au moins « FAIR ».

Si les procédés n'ont pas la qualité requise, 2 alternatives sont proposées :

- La première consiste à améliorer la qualité des données utilisées pour les procédés les plus contributeurs et contribuant à 90 % des impacts (solution ❶)
- La seconde consiste à sélectionner d'autres données de meilleure qualité, permettant également d'atteindre 90 % des impacts. Cela revient à rechercher des procédés moins contributeurs ayant le niveau de qualité requise s'il est plus facile d'accéder à ces données. (solution ❷)

**Figure 16 : Exemple du traitement de l'amélioration de la qualité pour une étude d'ACV screening selon la méthode PEF- Etape de recherche de données meilleures pour les mêmes procédés ou de procédés de meilleure qualité**



**Source : Webinar – Data Quality in the context of the EF pilot phase of the European Commission DG ENV (18 July 2014)**

Dans le cas présenté ci-dessus, le procédé « laundry detergent ingredient » a une note de qualité trop faible (POOR). Soit on affine la donnée pour améliorer son niveau de qualité (solution ❶) soit on regarde les procédés contributeurs suivants permettant d'atteindre 90 % des impacts (ici « heat from natural gas » et « truck transport ») et leur niveau de qualité (solution ❷). Le niveau de qualité répond alors aux exigences (niveau « FAIR »), donc cela convient.

Remarque : selon le niveau de contribution des procédés, l'approche 2 peut dans certains cas, manquer de pertinence car alors on aura de la très bonne qualité pour des procédés peu contributeurs. De plus la recherche d'une bonne qualité pour de nombreux procédés peu contributeurs peut s'avérer inutile car alors on peut imaginer qu'une compensation

statistique va se faire naturellement. Dans le cas présent, cela peut être pertinent car on ne remplace qu'un procédé par deux autres. L'un des critères pour opter vers la solution ② plutôt que la ① est l'accessibilité des données.

#### VII.2.3.3.2. Des exigences obligatoires ou recommandées selon le but de l'étude

- Pour les études PEF :
  - Dans le cadre d'une étude de « screening » l'objectif est que pour les données contribuant au moins à 90 % des impacts estimés pour chaque catégorie, la note de qualité soit « Fair ». Cette évaluation est alors qualitative.
  - Si l'étude est communiquée à l'extérieur (B2B ou B2C), les exigences recommandées ci-dessus sont obligatoires
  - Si le but de l'étude est interne, les exigences recommandées ci-dessus sont recommandées et non obligatoires.

Les évaluations concernant la représentativité doivent être soumises à une revue.

- Pour le développement de PEFCR

Les PEFCR doivent fournir des éléments de guidance sur l'évaluation de la qualité des données notamment pour les aspects de représentativité.

Note : certains PEFCR peuvent formuler des exigences plus strictes si cela est pertinent notamment pour :

- Les procédés gate to gate
- Les phases amont ou aval
- Les activités de la chaîne logistique qui sont clés pour la catégorie de produit
- Les catégories d'impacts qui sont clés pour la catégorie de produit

### VII.2.3.4. Synthèse

Le tableau ci-dessous présente les exigences en fonction des différents cas et portées d'études :

**Tableau 7 : Synthèse des exigences de qualité des données de la méthodologie PEF selon les types d'études**

Types d'étude	Portée	Exigences requises définies dans le PEF	Evaluation
Etude PEF	Communication externe	Obligatoires	Semi quantitative pour les données couvrant au moins 70 % des impacts les plus contributrices (GOOD) Qualitatives pour le reste
	Interne	Recommandées	
Etude screening PEF		Obligatoire (mais niveau d'exigence différent d'une étude PEF)	Qualitative (FAIR) pour les données couvrant au moins 90 % des impacts
Etude sur base d'un PEFCR		Exigences Etude PEF + critères additionnels possibles	Modalités d'évaluation Etude PEF + précision sur représentativité

**Source : Commission recommendation of 9 April 2013 on the use of common methods to measure and communicate the life cycle environmental performance of products and organisations**

## VII.3. Annexe 3 : Les informations des bases de données sur la qualité

Les bases de données sont de bons outils pour pouvoir apprécier, en première approche, la qualité de certaines données.

En effet, dès l'étape initiale, il convient d'observer dans quelle mesure les données potentiellement disponibles dans les bases de données pourront répondre ou non à des exigences de qualité. L'incertitude que l'on va déterminer ici est l'incertitude relative.

Il est donc nécessaire de savoir si les bases de données fournissent des informations sur la qualité de leurs données, comment elles les évaluent, et surtout, comment un praticien peut avoir accès facilement et rapidement à ces informations.

En effet, l'accessibilité à l'information est cruciale dans cette étape de spécification initiale.

### VII.3.1. DESCRIPTION DE LA QUALITÉ DANS LES BASES DE DONNÉES

#### VII.3.1.1. La base de données ELCD

L'ILCD propose une approche d'estimation de la qualité globale sur une base semi-quantitative, basée sur l'évaluation de 6 indicateurs de qualité des données :

- Représentativité technologique (TeR)
- Représentativité géographique (GR)
- Représentativité temporelle (TiR)
- Complétude (C)
- Précision (P)
- Cohérence méthodologique (M)

La formule de calcul de la « note » de qualité est la suivante :

$$\text{Formule 3} \quad DQR = \frac{TeR + GR + TiR + C + P + M + X_w * 4}{i + 4}$$

- *DQR* : Data Quality Rating of the LCI data set; see Table 7
- *TeR, GR, TiR, C, P, M* : see Table 5
- *X<sub>w</sub>* : weakest quality level obtained (i.e. highest numeric value) among the data quality indicators
- *i* : number of applicable (i.e. not equal "0") data quality indicators

Cette note se base sur l'hypothèse forte que la note globale de qualité doit être atténuée au regard du critère le plus faible, qui est de ce fait sur-pondéré dans la formule de calcul.

Cela permet par exemple à des données de très grande précision mais parfaitement inadaptées en termes de représentativité par rapport au but et champ de l'étude, d'avoir une note « dégradée ».

Des tableaux sont fournis pour estimer les notes de chaque critère :

**Tableau 8 : Grille de notation pour différents critères de l'ILCD Handbook**

Quality level	Quality rating	Definition	Completeness overall environmental impact	Precision / uncertainty overall env. impact (relative standard deviation in
<b>Very good</b>	1	"Meets the criterion to a very high degree, having or no relevant need for improvement. This is to be judged in view of the criterion's contribution to the data set's potential overall environmental impact and in comparison to a hypothetical ideal data quality."	≥95 %	≤7 %
<b>Good</b>	2	"Meets the criterion to a high degree, having little yet significant need for improvement. This is to be judged in view of the criterion's contribution to the data set's potential overall environmental impact and in comparison to a hypothetical ideal data quality."	[85 % to 95 %]	(7 % to 10 %]
<b>Fair</b>	3	"Meets the criterion to a still sufficient degree, while having the need for improvement. This is to be judged in view of the criterion's contribution to the data set's potential overall environmental impact and in comparison to a hypothetical ideal data quality. "	[75 % to 85 %]	(10 % to 15 %]
<b>Poor</b>	4	"Does not meet the criterion to a sufficient degree, having the need for relevant improvement. This is to be judged in view of the criterion's contribution to the data set's potential overall environmental impact and in comparison to a hypothetical ideal data quality."	[50 % to 75 %]	(15 % to 25 %]
<b>Very poor</b>	5	"Does not at all meet the criterion, having the need for very substantial improvement. This is to be judged in view of the criterion's contribution to the data set's potential overall environmental impact and in comparison to a hypothetical ideal data quality."	50 %	25 %
<b>Not evaluated / unknown</b>	5	"This criterion was not judged / reviewed or its quality could not be verified / is unknown."		
<b>Not applicable</b>	0	"This criterion is not applicable to this data set, e.g. its geographical representativeness cannot be evaluated as it is a location-unspecific technology unit process."		

L'évaluation des critères avec ces notes, est basée **sur un jugement d'expert**. Elle se situe au niveau du jeu de données directement.

La quantification de la précision est réalisée par une simulation Monte Carlo ou d'autres procédures, mais c'est le réviseur qui au final décide de la note.

Il existe deux niveaux d'exigences qui ne font pas systématiquement appel à cette exigence de notation :

- **Le ILCD DataNetwork – entry level** : ce niveau d'exigence n'impose pas la notation, mais essentiellement une revue sur les 6 critères qualité.

Concernant la qualité, il n'existe pas d'évaluation selon des niveaux de qualité, mais plutôt une appréciation de celle-ci. Au minimum, les représentativités technologiques, géographique et temporelles doivent être documentées. La qualité est portée par la documentation qui doit être supportée par le format informatique de l'ILCD. Un minimum de documentation doit être spécifié.

Une revue doit être effectuée : le réviseur doit répondre à une qualification spécifique (connaissance du secteur, du produit et expérience en ACV).

Le réviseur peut être interne ou externe, mais dans le cas d'une revue interne un rapport séparé est requis. Si la revue est externe, les informations peuvent être directement consignées dans le jeu de données.

- **Le full ILCD compliance** :

Pour assurer la qualité globale du jeu de données, l'ILCD considère 5 aspects comme clés : Method ; data quality, nomenclature, documentation et Review.

Pour les critères qualité, l'ILCD full compliance impose de quantifier chaque critère qualité sur la base du tableau précédent, puis de calculer le DQR selon la formule.

On obtient ainsi, en fonction de la note globale, 3 niveaux possibles :

- High quality data
- Basic quality data
- Data estimate

Le tableau ci-dessous présente ces seuils :

**Tableau 9 : Grille de notation globale de la qualité**

Overall data quality rating (DQR)	Overall data quality level
$\leq 1.6^{218}$	"High quality"
$>1.6 \text{ to } \leq 3$	"Basic quality"
$>3 \text{ to } \leq 4$	"Data estimate"

Les exigences en matière de documentation sont plus élevées

Les exigences en termes de revue critique imposent une revue externe et indépendante, menée par un réviseur enregistré. Il doit être indiqué si les éléments révisés l'ont été correctement au regard du champ, des méthodes et de la documentation.

Note : A ce jour, les jeux de données de la base de données ILCD suivent les recommandations de l'ISO. Cependant plusieurs jeux de données (environ 34 et plus encore en préparation) répondent à l'exigence ILCD entry level.

### VII.3.1.2. La base de données GaBi

#### VII.3.1.2.1. Qualité

La base de données prend en compte 7 critères :

- Représentativité temporelle : PE juge très bonne la qualité temporelle de vieilles données sur certains jeux dans la mesure où ils reflètent toujours la technologie actuelle
- Représentativité géographique : 4 critères doivent être renseignés :
  - La technologie est-elle représentative pour la région / le pays indiqué ?
  - Les matériaux utilisés sont-ils représentatifs pour la région / le pays considéré ?
  - Les énergies sont-elles représentatives pour la région/ le pays considéré ?
  - Le champ « mix and location type » représente-t-il celui cité dans la documentation ?
- Représentativité technologique : c'est un critère crucial pour PE qui est souvent oublié.
- Complétude : c'est sur base du jugement d'expert que l'on identifie les flux les plus impactants ; cependant avoir une note « very good » est très difficile car elle mentionne « tous les flux pris en compte » ; il y a moins de 20 jeux de données qui auraient ce critère, la plupart étant dans la catégorie good.
- Consistency : cohérence méthodologique
- **Un autre critère est pris en compte, non considéré par l'ILCD et le PEF : Fiabilité / origine** : il vise à étudier l'origine de la donnée et l'éventuelle vérification qui est faite par un expert. C'est un point important, qui peut avoir des répercussions sur la précision.
- Précision : PE ne souhaite pas indiquer de donnée **mathématiques en particulier sur la précision**. La raison est liée aux limitations intrinsèques des simulations Monte Carlo qui ne gèrent pas les lacunes, les questions de modélisation, les hypothèses et les corrélations entre flux. En effet, pour eux cela provient d'un jugement d'expert et cela doit apparaître comme tel. Cette donnée doit permettre au praticien de se poser la question suivante : puis-je faire confiance à ces chiffres ?

Ces critères sont évalués par un jugement d'expert sur la base d'une échelle de notation (allant de poor à very good). Ces critères ont été implémentés pour l'ensemble des jeux de données de la base, soit environ 7.200 jeux de données.

Le tableau ci-dessous récapitule ces données :

**Tableau 10 : Grille de notation pour différents critères dans GaBi**

	<b>Time representativeness</b>	<b>Geographical representativeness</b>	<b>Technological representativeness</b>	<b>Completeness</b>	<b>Precision</b>	<b>Consistency</b>
<b>Very good</b>	Completely representative - Check of representativeness or main data source not older than 3 years	Completely representative - all 4 criteria met	Completely representative - Technology mix or solely existing technology in the market regarding unit process and related main precursors (energy and materials)	All flows recorded: The entire process is covered by complete access to process data or the process was modelled in a very detailed form. Processes in which the cut-off rules were applied and checked can also be considered complete.	Measured /calculated AND verified	Defined methodology or standard, certified compliance
<b>Good</b>	Completely /partly representative	Completely /partly representative - 3 out of 4 criteria met	Completely / partly - Main technology in the market AND precursors from the main technology of the market	All relevant flows recorded: The relevant flows of the process are covered. When all flows cannot be recorded, this is the next option, which still enables good quality of results in terms of evaluation.	Measured / calculated / literature and plausibility checked by expert	GaBi modelling principles
<b>Fair</b>	Partly representative - Check of representativeness or main data source not older than 3 years, known changes but still partly representative	Partly representative - 2 out of 4 criteria met	Partly representative - one of the relevant technologies in the market and precursors from the main technology of the market OR main technology of the market and precursors from one of the relevant technologies in the market	Individual relevant flows recorded: Only particular flows are recorded. It must be clear that in this case some important flows can have been omitted, so only medium quality of data can be achieved. If possible, further research should be performed.	Measured / calculated / literature and plausibility not checked by expert OR Qualified estimate based on calculations plausibility checked by expert	ISO 14.040 with additional method/consistency requirements mainly met
<b>Poor</b>	Partly / not representative	Partly / not representative - 1 out of 4 criteria met	Partly / not - one of the existing technologies and precursors from one of the existing technologies in the market	Some relevant flows not recorded: If good quality is desired, this case should not occur. In the case that no data is available, reasons for using this kind of data should be documented.	Qualified estimate based on calculations, plausibility not checked by expert	ISO 14.040 with additional method/consistency requirements partly met
<b>Very</b>	Not representative - technology that is	Not representative - unit process and main	Not representative - one of the existing	no statement about	Very poor: Rough estimate with known	Methodology or consistency with

« SPECIFICATION INITIALE DES BESOINS DE QUALITE EN ACV »

---

<b>Poor</b>	known to be not representative	precursors representing another geography than the area stated and are known to be not representative	technologies that is known to be not representative"	completeness available"	deficits, not based on calculations	known deficits"
-------------	--------------------------------	---	--	-------------------------	-------------------------------------	-----------------

La qualité globale est calculée à l'échelle d'un jeu de données :

Trois approches existent :

- Moyenne arithmétique de leurs 6 critères,
- Moyenne arithmétique PEF (pas la même échelle de valeur)
- Approche ILCD

PE mentionne ainsi que la qualité des jeux de données dans la base peut se représenter ainsi :

- 100 % des jeux de données sont utilisables dans des projets ILCD/PEF
- 95 % des jeux de données atteignent une qualité globale « GOOD » et sont utilisables dans des projets PEF sans restrictions.
- 4 % des jeux de données atteignent une qualité globale « FAIR » et sont utilisables dans des projets PEF, si ces 4 % ne pèsent pas plus de 30 % des résultats finaux dans chaque catégorie d'impact.
- 1 % des jeux de données atteignent une qualité globale « POOR » et sont utilisables dans des projets PEF, si ces 1 % ne pèsent pas plus de 30 % des résultats finaux dans chaque catégorie d'impact.

Par ailleurs, PE indique que les jeux de données suivent les exigences requises par différents systèmes qualité:

- 100% suivent le GaBi Database and Modelling Principles 2013
- 100% suivent l'ISO 14040/14044
- une grande majorité des jeux de données suit l'"ILCD Data Network - entry-level". Le principal obstacle à la conformité est le processus de revue qui n'est pas correctement défini et est très coûteux.
- BP X 30: certains jeux de données GaBi data composent la base de données Impacts de l'ADEME et suivent les principes du BP X 30 -323
- PEF : comme le PEF n'est pas implémenté, il est trop tôt pour indiquer sa conformité aux exigences.

#### VII.3.1.2.2. Revue

Le système de revue mis en place au sein de la base de données est le suivant :

Toutes les entrées subissent dans un premier temps une revue interne indépendante, par un des experts de l'industrie / la technologie en question. Ce réviseur de contenu travaille en collaboration avec le gestionnaire du projet qui a produit le jeu de données. Il évalue la cohérence méthodologique.

Par ailleurs certaines parties de la base de données GaBi ont été soumises à des revues externes indépendantes qualifiées dans le cadre du ILCD data network entry level

puisque les données GaBi contribuent à l'alimentation de la base ELCD de la Commission européenne.

De plus, toutes les données d'énergie de la base GaBi dans la base de données ELCD ont été récemment analysées plus en détail dans le cadre d'une revue critique externe indépendante par CIEMAT et l'ENEA (deux grandes organisations énergie-recherche de l'Espagne et de l'Italie) et comparées aux autres sources de données.

Enfin, PE a entamé une revue plus systématique du processus de revue lui-même, incluant le processus d'assurance et la revue interne indépendante pour l'ensemble de la base de données avec l'organisation d'accréditation DEKRA. Ce processus de révision est en cours, les principales technologies avec un grand impact sur les résultats globaux de la base de données ont déjà été revues. Un examen du contenu de tous les ensembles de données est en cours de discussion.

### VII.3.1.3. La base de données ecoinvent

La base de données ecoinvent distingue deux approches de la qualité :

- La qualité des points de données (chaque flux élémentaire et intermédiaire)
- La qualité du jeu de données

#### VII.3.1.3.1. Qualité au niveau des points de données

Ecoinvent fournit une information sur la qualité au niveau de chaque flux. Il n'existe pas à ce stade de données sur la qualité du jeu de données global, mais des projets sont prévus en ce sens.

Le calcul de la qualité de la donnée se base sur 5 caractéristiques qui sont :

- Corrélation temporelle
- Corrélation géographique
- Corrélation technologique
- Complétude
- Fiabilité.

Un jeu de 5 scores entre 1 et 5 sont fournis pour chaque point individuel (input et output exchange).

Le tableau qui fait foi est la matrice pedigree qui se base sur les mêmes principes que la matrice pedigree proposée initialement :

**Tableau 11 : Matrice pedigree d’ecoinvent**

Indicator score	1	2	3	4	5 (default)
<b>Reliability</b>	Verified <sup>5</sup> data based on measurements <sup>6</sup>	Verified data partly based on assumptions or non-verified data based on measurements	Non-verified data partly based on qualified estimates	Qualified estimate (e.g. by industrial expert)	Non-qualified estimate
<b>Completeness</b>	Representative data from all sites relevant for the market considered, over an adequate period to even out normal fluctuations	Representative data from >50% of the sites relevant for the market considered, over an adequate period to even out normal fluctuations	Representative data from only some sites (<<50%) relevant for the market considered or >50% of sites but from shorter periods	Representative data from only one site relevant for the market considered or some sites but from shorter periods	Representativeness unknown or data from a small number of sites <b>and</b> from shorter periods
<b>Temporal correlation</b>	Less than 3 years of difference to the time period of the dataset	Less than 6 years of difference to the time period of the dataset	Less than 10 years of difference to the time period of the dataset	Less than 15 years of difference to the time period of the dataset	Age of data unknown or more than 15 years of difference to the time period of the dataset
<b>Geographical correlation</b>	Data from area under study	Average data from larger area in which the area under study is included	Data from area with similar production conditions	Data from area with slightly similar production conditions	Data from unknown or distinctly different area (North America instead of Middle East, OECD-Europe instead of Russia)
<b>Further technological correlation</b>	Data from enterprises, processes and materials under study	Data from processes and materials under study (i.e. identical technology) but from different enterprises	Data from processes and materials under study but from different technology	Data on related processes or materials	Data on related processes on laboratory scale or from different technology

Par ailleurs, une autre information est fournie, il s’agit de l’incertitude de la mesure, appelée « basic uncertainty ». Cette incertitude est établie de façon différente selon la disponibilité des données :

- Si un échantillon de données est disponible la distribution de probabilité et l’écart type peuvent être définis directement ; cependant ce cas est rarement disponible.
- Sinon, l’incertitude peut être fournie par le fournisseur de données lui-même en donnant des informations sur l’écart-type.
- Enfin, quand aucun élément n’est disponible, Ecoinvent propose des distributions par défaut. La distribution proposée est toujours une distribution de probabilité log normale, car elle évite le problème de valeurs négatives que l’on peut rencontrer avec une distribution normale.

Le tableau ci-dessous présente les valeurs par défaut proposées :

**Tableau 12 : Valeurs par défaut pour les incertitudes de différents flux dans écoinvent (sur base du jugement d’expert)**

Table 10.3. Default basic uncertainty (variance  $\sigma_b^2$  of the logtransformed data, i.e. the underlying normal distribution) applied to intermediate and elementary exchanges when no sampled data are available; c: combustion emissions; p: process emissions; a: agricultural emissions

Input / output group	c	p	a	Input / output group	c	p	a
<b>demand of:</b>				<b>pollutants emitted to air:</b>			
thermal energy, electricity, semi-finished products, working material, waste treatment services	0.0006	0.0006	0.0006	CO <sub>2</sub>	0.0006	0.0006	
transport services (tkm)	0.12	0.12	0.12	SO <sub>2</sub>	0.0006		
Infrastructure	0.3	0.3	0.3	NMVOC total	0.04		
<b>resources:</b>				NO <sub>x</sub> , N <sub>2</sub> O	0.04		0.03
Primary energy carriers, metals, salts	0.0006	0.0006	0.0006	CH <sub>4</sub> , NH <sub>3</sub>	0.04		0.008
Land use, occupation	0.04	0.04	0.002	Individual hydrocarbons	0.04	0.12	
Land use, transformation	0.12	0.12	0.008	PM>10	0.04	0.04	
<b>pollutants emitted to water:</b>				PM10	0.12	0.12	
BOD, COD, DOC, TOC, inorganic compounds (NH <sub>4</sub> , PO <sub>4</sub> , NO <sub>3</sub> , Cl, Na etc.)		0.04		PM2.5	0.3	0.3	
Individual hydrocarbons, PAH		0.3		Polycyclic aromatic hydrocarbons (PAH)	0.3		
Heavy metals		0.65	0.09	CO, heavy metals	0.65		
Pesticides			0.04	Inorganic emissions, others		0.04	
NO <sub>3</sub> , PO <sub>4</sub>			0.04	Radionuclides (e.g., Radon-222)		0.3	
<b>pollutants emitted to soil:</b>							
Oil, hydrocarbon total		0.04					
Heavy metals		0.04	0.04				
Pesticides			0.033				

Ces données par défaut ont été établies sur base de jugement d’experts. Un groupe de travail au sein d’écoinvent a été constitué sur la question de l’incertitude et de l’établissement de ces valeurs.

Les corrélations éventuelles pour mesurer l’incertitude sont prises en compte via des formules mathématiques permettant de lier, par exemple des flux élémentaires entre eux ou des liens au niveau de leurs propriétés.

Un calcul d’incertitude globale est ensuite mené permettant de livrer une information se traduisant sous une forme de distribution.

Le tableau ci-dessous indique la formule de calcul :

**Figure 17: Formule de calcul de l’incertitude globale dans ecoinvent**

Since each normal distribution is assumed to be independent, i.e. their covariance is zero, the variance of the summed final distribution is then:

$$\sigma^2(X+Y) = \sigma^2(X) + \sigma^2(Y) + 2\text{cov}(X,Y)$$

$$\sigma^2 = \sum_{n=1}^6 \sigma_n^2$$

with:

$\sigma_1^2$ : basic uncertainty (variance measured or estimated according to Table 10.3)

$\sigma_2^2$ : uncertainty factor (variance) of reliability distribution

$\sigma_3^2$ : uncertainty factor (variance) of completeness distribution

$\sigma_4^2$ : uncertainty factor (variance) of temporal correlation distribution

$\sigma_5^2$ : uncertainty factor (variance) of geographical correlation distribution

$\sigma_6^2$ : uncertainty factor (variance) of other technological correlation distribution

Les variances des différents facteurs d’incertitude sont établis à partir d’un tableau de correspondance des notes de la matrice pedigree.

**Tableau 13 : Correspondance entre les incertitudes liées aux différents critères et les notes de la matrice pedigree dans ecoinvent**

Table 10.5. Uncertainty factors (variances of the underlying normal distributions) used to convert the data quality indicators of the pedigree matrix in Table 10.4 into additional uncertainty.

Indicator score	1	2	3	4	5
Reliability	0.000	0.0006	0.002	0.008	0.04
Completeness	0.000	0.0001	0.0006	0.002	0.008
Temporal correlation	0.000	0.0002	0.002	0.008	0.04
Geographical correlation	0.000	2.5e-5	0.0001	0.0006	0.002
Further technological correlation	0.000	0.0006	0.008	0.04	0.12

On obtient ainsi une incertitude totale, basée sur des distributions log normales.

**VII.3.1.3.2. Informations sur la qualité du jeu de données**

Concernant la documentation, chaque jeu de données est extensivement documenté. Ecoinvent fournit aussi des rapports par grandes thématiques (rapports sectoriels) reprenant la construction des jeux de données, les sources, les hypothèses de modélisation...

Par ailleurs, il existe une notion de validité temporelle du jeu de données au niveau du jeu de données.

Enfin, les informations concernant les représentativités géographiques et technologiques existent au niveau du jeu de donnée, tout comme l’approche méthodologique utilisée.

Il n’existe pas de critère « cut off » autorisé dans la nouvelle base de données, le manque de flux élémentaire n’est pas autorisé. Cela signifie que le jeu de données se veut complet. Tous les flux intermédiaires et élémentaires doivent être répertoriés sauf s’ils sont exclus du champ du jeu de données.

### VII.3.1.3.3. Revue des jeux de données

Concernant la revue des jeux de données, le processus est le suivant : un « editorial board » est constitué.

Il est composé de 50 éditeurs, experts dans différents domaines, chaque domaine recouvrant :

- Une activité économique (agriculture, industries extractives...), classifiée dans le schéma ISIC<sup>12</sup>

Cet éditeur s'appelle le « Main editor ». Il s'occupe de garantir la qualité du jeu de données entrant et sa cohérence dans le secteur spécifiques où l'activité doit s'intégrer.

- Une zone géographique
- Certains types d'émissions
- Certains champs de la base comme par exemple l'incertitude.

Les éditeurs qui sont concernés par les 3 tâches ci-dessus sont des éditeurs transversaux et s'intéressent à la cohérence transversale des jeux de données dans la base.

Chaque jeu de données produit passe donc au minimum par 3 éditeurs :

- L'éditeur activité économique
- 2 éditeurs transversaux (incertitude, émissions, documentation...)

Tous les commentaires sont consignés et documentés par ecoinvent.

On peut considérer ces revues comme indépendantes et externes (pour les fournisseurs extérieurs) ; pour des développements propres, la revue est interne, mais on s'assure qu'elle soit indépendante (pas de conflit d'intérêt du réviseur) et le même type de revue est assuré.

### VII.3.1.4. La base de données Impacts

La base Impacts est une base de données qui met à disposition des utilisateurs uniquement des résultats d'impacts (selon différentes méthodes de caractérisation, sélectionnée sur base de recommandation du JRC).

En ce sens, il est fondamental que les données fournies par les fournisseurs de données soient complètes pour tous les flux couverts par les différentes catégories d'impacts.

La base Impacts se base sur le niveau Data Entry level de l'ILCD.

Comme elle considère le même format que l'ILCD elle dispose des mêmes champs.

---

<sup>12</sup> ISIC : International Standard Industrial Classification of all Economic Activities de l'ONU.

A ce stade elle a sélectionné certains champs comme étant plus importants que d'autres pour décrire le jeu de données.

Pour l'instant, les rapports de revue (interne) des fournisseurs de données ne sont pas disponibles, alors qu'ils doivent être normalement rendus publics pour répondre à l'exigence de l'ILCD data entry level. L'ADEME travaille à respecter cette exigence, en établissant des rapports sectoriels rappelant les types de procédés, la représentativité et les règles de modélisation ou données transversales utilisées.

### VII.3.1.5. Synthèse

Le tableau ci-dessous récapitule les critères de qualité des données pris en compte et les modes d'évaluation de ces critères, en précisant :

- Le caractère descriptif
- Le caractère semi-quantitatif (grille de notation mise en place)
- Le caractère quantitatif (un chiffre pour le critère en question ou une valeur d'écart-type)

**Tableau 6 : Récapitulatif des critères de qualité couverts par les bases de données et les évaluations pratiquées**

Termes en FR	Termes en EN	ELCD Database		GaBi		Ecoinvent		Base Impacts	
		Présence du critère	Nature de l'évaluation	Présence du critère	Nature de l'évaluation	Présence du critère	Nature de l'évaluation	Présence du critère	Nature de l'évaluation
<b>Caractère approprié de la modélisation / Appropriateness</b>									
Représentativité temporelle	Time-related representativeness	X	Semi-quantitatif	x	Semi-quantitatif	x	Quantitatif & semi-quantitatif	x	descriptif
Représentativité géographique	Geography representativeness	X	Semi-quantitatif	x	Semi-quantitatif	x	Quantitatif & semi-quantitatif	x	descriptif
Représentativité technologique	Technology representativeness	X	Semi-quantitatif	x	Semi-quantitatif	x	Quantitatif & semi-quantitatif	x	descriptif
Complétude	Completeness	X	Quantitatif	x	Semi-quantitatif	x	Quantitatif & semi-quantitatif		
Fiabilité	Reliability					x	Quantitatif & semi-quantitatif		
Cohérence méthodologique	Methodological appropriateness and consistency	X	Semi-quantitatif	X	Semi-quantitatif	x	Semi-quantitatif/quantitatif		
<b>Précision intrinsèque / Precision</b>									
Précision	Precision / uncertainty	X	Quantitatif		Quantitatif		Quantitatif		
<b>Incertitude globale / overall uncertainty</b>									
	Overall uncertainty	X	Semi-quantitatif	x	Semi-quantitatif	x	Quantitatif		
<b>Autres critères</b>									
<b>Couverture temporelle</b>	Time-related coverage	X	descriptif	X	descriptif	X	descriptif	X	descriptif

« SPECIFICATION INITIALE DES BESOINS DE QUALITE EN ACV »

<b>Couverture géographique</b>	Geographical coverage	X	descriptif	X		X	descriptif	X	descriptif
<b>Couverture technologique</b>	Technology coverage	X	descriptif	X		X	descriptif	X	descriptif
<b>Reproductibilité</b>	Reproducibility	X		X		X			
<b>Source</b>	Sources of data	X	joint	X		X	Rapport annexé	X	Joint
<b>Documentation</b>	Documentation (source, format transparency to ensure reproducibility)	X	joint	XX		x	Rapport dans le jeu de données	X	Joint + rapport annexé
<b>Revue</b>	Review aspects	X	Cité rapport joint +	X		x	Nom des réviseurs dans le jeu de données Rapport annexé accessible sur demande	X	Cité + rapport annexé
<b>Nomenclature</b>	Nomenclature (structure and naming the elementary flows)	X	Non visible	X	Non visible	x	Non visible	X	Non visible

## **VII.4. Annexe 4 : Informations utiles dans les bases de données pour la spécification initiale de la qualité**

Nous détaillons ci-dessous, pour chaque base de données les « champs » pertinents à étudier lors de la consultation de la base de données, qui permettront de bien se focaliser sur les enjeux de qualité :

### **VII.4.1.1. Base de données ELCD**

Selon les types d'inventaires, il peut exister un rapport de revue succinct, qui va reprendre des éléments sur la qualité des données, de façon très simple et visuelle.

Dans les inventaires consultables en ligne, différents champs des métadonnées sont particulièrement intéressants :

- Geographical representativeness,
- Time representativeness : la durée de validité de la donnée et la représentativité temporelle de la donnée sont considérées
- Technological representativeness

Figure 18 : Exemple de champs pertinents à consulter pour la base de données ELCD

Chlorine;production mix for PVC production, at plant (en)	
Table of Contents: <a href="#">Process information</a> - <a href="#">Modelling and validation</a> - <a href="#">Administrative information</a> - <a href="#">Inputs and Outputs</a>	
Process information	
Key Data Set Information	
Location	RER
Geographical representativeness description	The data set represents the region specific situation, focusing on the main technologies and the region specific characteristics.
Reference year	1993
Name	Base name: Treatment, standards, routes: Mix and location types Chlorine;production mix for PVC production, at plant
Use advice for data set	The LCI data set should be used for LCI/ LCA studies where chlorine is used as raw material / feedstock for the PVC production.
Technical purpose of product or process	This Chlorine is just for the use as feedstock / raw material for the production of polyvinylchloride PVC.
Classification	Class name / Hierarchy level Materials production / Inorganic chemicals
General comment on data set	Raw data for intermediate products are collected by several producers in Europe; the average is calculated by the consultant based on the production capacity of the companies. The LCI result is calculated with background data from the Boustead database.
Copyright? Yes    Owner of data set (contact data set) <a href="#">Plastics Europe</a>	
Quantitative reference	
Reference flow(s)	chlorine - 1.0 kg (Mass)
Time representativeness	
Data set valid until:	2005
Time representativeness description	The data corresponding to brine purification and brine electrolysis were collected in 1994. All the calculations were carried out in 2005.
Technological representativeness	
Technology description including background system	ELECTROLYSIS: Over 90% of all industrial chlorine is produced by electrolysis, a process in which an electric current is passed through a brine solution. In its simplest form and electrolytic cell is as shown in Figure 1. Two plates, or electrodes, are inserted into a brine solution and connected to a DC power supply. The electrode connected to the negative terminal is called the cathode and the electrode connected to the positive terminal is called the anode. When the current passes, chlorine gas is liberated at the anode, hydrogen gas is liberated at the cathode and the electrolyte is gradually converted from sodium chloride to sodium hydroxide. The net result of the reaction can be written as: $\text{NaCl(aq)} + \text{H}_2\text{O(l)} = \text{NaOH(aq)} + \frac{1}{2}\text{H}_2(\text{g}) + \frac{1}{2}\text{Cl}_2(\text{g})$ All of the products of this reaction are marketable and so should be as pure as possible - hence one reason for the initial purification of the brine. The background system is addressed as follows: Electricity: Electricity used, is modelled according to the individual country-specific situation. Modelling the electricity production system covers the production and delivery of primary fuels, the steam generator, the electricity generator and the distributor. Electricity is regarded to be derived from the public electricity supply and from on-site electricity generation of plants. Input materials used for producing process steam are fossil fuels, like coal, oil and gas. Process Steam: Process steam/thermal energy used, is modelled according to the individual country-specific situation. Process steam is regarded to be derived as by-product from the public electricity supply and from on-site steam generation of plants. Input materials used for producing process steam are fossil fuels, like coal, oil and gas. Another source for process steam are waste incineration plants. Transport: All relevant and known transport processes used, are included. Overseas transport to and from major ports as well as all relevant and known pipeline and / or tanker transport of crude oil imports are included. Energy carriers: Crude oil is modelled according to the production and supply of crude oil to Western Europe. The crude oil supply is regarded regional. Therefore, interchanges between countries within that region do not count as imports or exports. Oil processing is reckoned as standard worldwide with overall energy production efficiency of the order of 90%. The IEA data on world oil production and processing data have been used for the use of energy. The total inputs have been partitioned between the products using energy as the partitioning parameter. Modelling covers the extraction, the transport and the delivery including emissions arising at the extraction site and reported losses during subsequent distribution of the oil. The transport for indigenous crude oil production is based on the characteristics of North Sea production since this represents 98% of indigenous production. For North Sea oil some 75% is delivered by pipeline with the remainder delivered by tanker. Coal modelling covers the production, the transport and the delivery.

Note : les champs Completeness existe également, le champ Precision n'est pas toujours renseigné.

2 champs souvent oubliés, mais utiles pour la sélection du bon procédé (en rouge dans le procédé présenté ci-dessus) :

- **Use advice for data set**
- **Technical purpose of product or process**

Ces 2 champs peuvent aider, à **vérifier la pertinence de retenir un inventaire** et fournissent des éléments complémentaires sur le caractère approprié de la donnée relativement au but et champ de l'étude menée

Enfin, dans le cas où l'on recherche de façon fine **l'évaluation des différents critères**, il convient d'analyser la partie relative à la validation :

Cette partie peut se présenter, selon qu'il existe une évaluation semi quantitative ou qualitative ainsi :

**Tableau 14 : Présentation des méthodes d'évaluation de la qualité dans la base de données ELCD**

Validation			
Review	<i>Independent external review</i>		
	<i>Scope of review</i>	<i>Method(s) of review</i>	
	<i>Documentation</i>	<i>Expert judgement</i>	
Data quality indicators	<b>Name</b>	<b>Value</b>	
	<i>Geographical representativeness</i>	<i>Good</i>	
	<i>Methodological appropriateness and consistency</i>	<i>Good</i>	
	<i>Time representativeness</i>	<i>Good</i>	
	<i>Completeness</i>	<i>Not evaluated / unknown</i>	
	<i>Overall quality</i>	<i>Good</i>	
	<i>Precision</i>	<i>Not evaluated / unknown</i>	
	<i>Technological representativeness</i>	<i>Good</i>	
Review	<i>Dependent internal review</i>		
	<i>Scope of review</i>	<i>Method(s) of review</i>	
	<i>LCI results or Partly terminated system</i>	<i>Expert judgement</i>	
		<i>Cross-check with other data set</i>	
	<i>LCIA results</i>	<i>Expert judgement</i>	
	<i>Documentation</i>	<i>Expert judgement</i>	

Source : <http://eplca.jrc.ec.europa.eu/ELCD3/resource/processes/945256f7-e041-46c8-b3bb-0d447d4b398b>

**VII.4.1.2. La base de données GaBi**

On peut visualiser, pour un inventaire donné les éléments sur la qualité des indicateurs, **en plus** des autres champs typiques tels que la représentativité ...

Les résultats sont donnés selon les différents systèmes de notations GaBi, PEF et ILCD.

**Figure 19 : Exemple de visualisation du niveau de qualité pour un inventaire donné dans GaBi**

The screenshot shows the 'Validation' window in GaBi. It includes a dropdown for 'Type of review' (set to 'Dependent internal review') and a 'Delete review' button. Below is a table for 'Scope of review' with columns for 'Scope of review' and 'Method(s) of review'. A red box highlights the 'Quality indicators' table, which lists various indicators and their values. Below this is a 'Review details' text area, a list of 'Reviewer name and institution' (PE INTERNATIONAL, LBP-GaBi, IBP-GaBi), and another red box highlighting 'Other review details' which provides a summary of quality scores for GaBi, ILCD, and PEF.

Scope of review	Method(s) of review
Raw data	Validation of data sources, Sample tests on calculations, Cross-check with other source, Expert judgement
Unit process(es), single operation	Energy balance, Element balance, Cross-check with other source, Cross-check with other data set, Expert judgement
Unit process(es), black box	Energy balance, Element balance, Cross-check with other source, Cross-check with other data set, Expert judgement
LCIA results or Partly terminated system	Energy balance, Element balance, Cross-check with other source, Cross-check with other data set, Expert judgement
LCIA results	Cross-check with other source, Cross-check with other data set, Expert judgement, Compliance with ISO 14040 to 14044
Documentation	Expert judgement, Compliance with ISO 14040 to 14044
Life cycle inventory methods	Compliance with ISO 14040 to 14044
LCIA results calculation	Expert judgement, Compliance with ISO 14040 to 14044
Goal and scope definition	Expert judgement, Compliance with ISO 14040 to 14044

Quality indicators	Value
Technological representativeness	Good
Time representativeness	Very good
Geographical representativeness	Good
Completeness	Good
Precision	Good
Methodological appropriateness and consistency	Good
Overall quality	Good

**Other review details:**  
 Overall quality according to different validation schemes  
 GaBi = 1,8 interpreted into "good overall quality" in the GaBi quality validation scheme  
 ILCD = 1,9 interpreted into "basic overall quality" in the ILCD quality validation scheme  
 PEF = 1,8 interpreted into "very good overall quality" in the PEF quality validation scheme

**VII.4.1.3. La base de données ecoinvent**

On peut visualiser les données relatives à l’incertitude par flux élémentaire, ainsi que les notes individuelles de chaque critère de la matrice pedigree **en plus** des autres champs typiques tels que la représentativité.

**Figure 20 : Exemple de qualité des données disponible au niveau d’un point de données (flux élémentaire)**

Unit Process Exchanges				
Name	Amount	Unit	Uncertainty SD	
<b>Reference Products</b>				
+ partide board, for outdoor use	1	m3		
<b>By-product/Waste</b>				
+ wood chips, dry, measured as dry mass	0	kg		
<b>To Environment</b>				
air				
- Formaldehyde	0.0029403	kg	Lognormal	1.9993
<b>Uncertainty</b>	Lognormal (Geometric mean=0.0029403, Variance of log-transformed data=0.12, Arithmetic mean of log-transformed data=-5,81, Standard deviation=1.9993, CI/2wP, half range of confidence interval=2.2294, Variance of data with pedigree=0.1607) Pedigree matrix: 1 3 5 3 1, Comment:			
Compartment	air	Subcompartment	unspecified	
Classification		Formula		
CAS-Number	000050-00-0	Synonym		
Activity Link		Variable Name		
Mathematical Relation		Comment	(1,3,2,3,1,1);	
Tags		Source		
Page Numbers		Annual Production Volume		
Production Volume Variable Name		Production Volume		
Production Volume Comment		Mathematical Relation		
		Specific Allocation Property		
+ NMVOC, non-methane volatile organic compounds, unspecified origin	0.1627	kg	Lognormal	1.9993
- Particulates, < 2.5 um	0.0038224	kg	Lognormal	3.0999
<b>Uncertainty</b>	Lognormal (Geometric mean=0.0038224, Variance of log-transformed data=0.32, Arithmetic mean of log-transformed data=-5,55, Standard deviation=3.0999, CI/2wP, half range of confidence interval=3.3794, Variance of data with pedigree=0.3707) Pedigree matrix: 3 3 5 3 3, Comment:			
Compartment	air	Subcompartment	unspecified	
Classification		Formula		
CAS-Number		Synonym		
Activity Link		Variable Name		
Mathematical Relation		Comment	(3,3,1,3,3,5);	
Tags		Source		
Page Numbers		Annual Production Volume		
Production Volume Variable Name		Production Volume		
Production Volume Comment		Mathematical Relation		
		Specific Allocation Property		

Exemple d’informations utiles dans le jeu de données

<b>Activity</b>	
Activity Name	EUR-flat pallet production
Type	UnitProcess
Special Type	OrdinaryTransformingActivity
Inheritance Depth	LinkedDataset
General Comment	<p>Module includes only materials (no production process) and should only be used for the packing and transportation of products.</p> <p>[This dataset was already contained in the ecoinvent database version 2. It was not individually updated during the transfer to ecoinvent version 3. Life Cycle Impact Assessment results may still have changed, as they are affected by changes in the supply chain, i.e. in other datasets. This dataset was generated following the ecoinvent quality guidelines for version 2. It may have been subject to central changes described in the ecoinvent version 3 change report (<a href="http://www.ecoinvent.org/database/ecoinvent-version-3/reports-of-changes/">http://www.ecoinvent.org/database/ecoinvent-version-3/reports-of-changes/</a>), and the results of the central updates were reviewed extensively. The changes added e.g. consistent water flows and other information throughout the database. The documentation of this dataset can be found in the ecoinvent reports of version 2, which are still available via the ecoinvent website. The change report linked above covers all central changes that were made during the conversion process.]</p>
Included Activities Start	
Included Activities End	Includes only the materials and not the process of construction. The examined system is from gate to gate and as in most cases the paletts have a long life-span the waste treatment is not included. It must be included in the packing module
Synonym	Palett, stillage
Tags	ConvertedDataset
Energy Values	Undefined
Master Allocation Property	
Allocation Comment	
Dataset Icon	
<b>Classifications</b>	
System : Value	EcoSpold01Categories: wooden materials/processing
System : Value	ISIC rev.4 ecoinvent: 1623:Manufacture of wooden containers
<b>Geography</b>	
Shortname	RER
Longname	Europe
Comment	
<b>Technology</b>	
Technology Level	Current
Comment	Standard composition of materials
<b>Time Period</b>	
Start Of Period	2000
End Of Period	2014
Data Valid For Entire Period	Yes
Comment	
<b>Macro Economic Scenario</b>	
Value	Business-as-Usual
Comment	
<b>Modelling And Administrative</b>	
<b>Representativeness</b>	
System Model	Allocation, ecoinvent default
Percent	
Sampling Procedure	internet search
Extrapolations	This dataset has been extrapolated from year 2002 to the year of the calculation (2014). The uncertainty has been adjusted accordingly.

## VII.5. Annexe 5 : Liste des personnes interviewées sur la qualité initiale des données

- Des commanditaires d'études ACV :
  - Sophie Bonnier (Eco-Emballages),
  - Cécile Desabbayes (Eco-mobilier)
  - Les membres de SCORELCA qui ont notamment contribué lors d'une réunion de travail à l'affinage de la méthode sur base de l'étude de cas pratiques (Anne Prieur - GDF Suez, Denis Le Boulch - EDF, Barbara Forrière - Renault)
- Des industriels praticiens de l'ACV :
  - Gert Van Hoof de Procter and Gamble
- Des consultants en ACV
  - Jad Zoghaïb (Solinnen)
  - An Vercalster et Katrien Boonen (VITO)
- Des fournisseurs de bases de données ACV
  - Marco Recchioni (JRC)
  - Olivier Réthoré (ADEME)
  - Emilia Moreno Ruiz (ecoinvent)
  - Thilo KUPFER (PE International)
- Des fournisseurs d'outils ACV
  - Tommie Pensioen (pour Pre consultant)
  - Thilo Kupfer (pour GaBi)

Note : RDC Environnement et plus spécifiquement Bernard DE CAEVEL a également apporté ses éléments et son expertise.