

**Titre:** Gestion de l'incertitude causée par l'incohérence d'échelle spatiale  
à l'interface de l'inventaire et de l'analyse des impacts en ACV

**Auteur:** Guillaume Bourgault

**Date:** 2013

**Type:** Mémoire ou thèse / Dissertation or Thesis

**Référence:** Bourgault, G. (2013). Gestion de l'incertitude causée par l'incohérence d'échelle spatiale à l'interface de l'inventaire et de l'analyse des impacts en ACV [Thèse de doctorat, École Polytechnique de Montréal]. PolyPublie.  
Citation: <https://publications.polymtl.ca/1318/>

 **Document en libre accès dans PolyPublie**  
Open Access document in PolyPublie

**URL de PolyPublie:** <https://publications.polymtl.ca/1318/>  
PolyPublie URL:

**Directeurs de recherche:** Réjean Samson, & Pascal Lesage  
Advisors:

**Programme:** Génie chimique  
Program:

UNIVERSITÉ DE MONTRÉAL

GESTION DE L'INCERTITUDE CAUSÉE PAR L'INCOHÉRENCE  
D'ÉCHELLE SPATIALE À L'INTERFACE DE L'INVENTAIRE ET DE  
L'ANALYSE DES IMPACTS EN ACV

GUILLAUME BOURGAULT  
DÉPARTEMENT DE GÉNIE CHIMIQUE  
ÉCOLE POLYTECHNIQUE DE MONTRÉAL

THÈSE PRÉSENTÉE EN VUE DE L'OBTENTION  
DU DIPLÔME DE PHILOSOPHIAE DOCTOR  
(GÉNIE CHIMIQUE)  
DÉCEMBRE 2013

UNIVERSITÉ DE MONTRÉAL

ÉCOLE POLYTECHNIQUE DE MONTRÉAL

Cette thèse intitulée:

GESTION DE L'INCERTITUDE CAUSÉE PAR L'INCOHÉRENCE D'ÉCHELLE SPATIALE  
À L'INTERFACE DE L'INVENTAIRE ET DE L'ANALYSE DES IMPACTS EN ACV

présentée par : BOURGAULT Guillaume

en vue de l'obtention du diplôme de : Philosophiae doctor

a été dûment acceptée par le jury d'examen constitué de :

Mme DESCHÊNES Louise, Ph.D., présidente

M. SAMSON Réjean, Ph.D., membre et directeur de recherche

M. LESAGE Pascal, Ph.D., membre et codirecteur de recherche

M. VIRGILIO Nick, Ph.D., membre

M. JOLLIET Olivier, Ph.D., membre

## **DÉDICACE**

Je dédie cette thèse à Bernard Lapierre, professeur du cours d'éthique pour ingénieur à l'École Polytechnique de Montréal. M. Lapierre a posé la question suivante à la classe, indirectement responsable de mon retour aux études quatre ans plus tard : « Allez-vous faire partie du problème, ou allez-vous faire partie de la solution? »

## REMERCIEMENTS

Les personnes suivantes m'ont supporté à chacun leur manière. Merci à Christopher Mutel, Manuele Margni, Cécile Bulle, Anne-Marie Boulay, Pascal Lesage, Réjean Samson, Gontran Bage, Valérie Bécaert, Ariane Pinsonneault, Jean-François Ménard, Catherine Reid, Sandra Estella, Geneviève Martineau, Marylène Dussault, Dominique Maxime, Hugues Imbeault-Tétreault, Pablo Tirado-Seco, Sophie Fallaha, Andrew Henderson, Eric Savoy, Martin Lennox, Joseph Flanders. Merci aussi aux partenaires de la Chaire pour leur support financier.

## RÉSUMÉ

L'analyse du cycle de vie (ACV) se divise en deux phases de calcul. L'inventaire du cycle de vie (ICV) prédit la quantité de chaque polluant émis par les processus élémentaires (PÉ) (usines, transport, extraction de matières premières). Il est possible de trouver les coordonnées géographiques d'une minorité de PÉs. Les autres PÉs sont modélisés par une base de données générique d'ICV, dont la résolution spatiale est généralement nationale : le pays du PÉ est connu, mais pas sa position au sein du pays.

La seconde phase de calcul, l'évaluation des impacts de cycle de vie (ÉICV), modélise le sort, l'exposition et l'effet des polluants sur la santé humaine et la biodiversité, par l'intermédiaire de facteurs de caractérisation (FCs). Pour certaines catégories d'impacts dites « régionales », comme l'acidification, l'eutrophisation, les effets respiratoires, etc., les FCs varient fortement sur de courtes distances. Cette variabilité spatiale peut atteindre plusieurs ordres de grandeur sur le territoire d'un même pays. Dans l'absence d'informations spatiales plus précises sur l'inventaire, on doit caractériser les flux d'inventaire à l'aide de FCs agrégés. Cette situation cause une augmentation d'incertitude qui n'avait pas été quantifiée jusqu'à présent.

Cette thèse propose un ensemble d'outils nécessaires à la gestion de cette incertitude. L'algorithme d'agrégation permet la quantification de l'incertitude créée par cette opération et l'identification des pays où l'utilisation des FCs agrégés devrait être évitée. Les résultats préliminaires indiquent que pour la catégorie d'impact de l'utilisation de l'eau sur la santé humaine, la variabilité spatiale et l'incertitude couvrent la même gamme de valeur pour la majorité des pays. Le manque d'information spatiale ne sera donc pas une source majeure d'augmentation de l'incertitude, sauf pour certains cas, identifiés par un indicateur d'augmentation d'incertitude élevé. La méthodologie d'agrégation et de calcul de l'indicateur d'augmentation d'incertitude pourront être appliqués à d'autres catégories d'impact à mesure que l'incertitude sur les FCs natifs sera publiée.

La thèse introduit aussi la puissance statistique, outil tiré des tests d'hypothèse pour juger de la signifiante de différences entre deux variables aléatoires, et montre comment il pourrait être utilisé dans la pratique des analystes et le développement du modèle. La puissance statistique permet d'alimenter la discussion et l'interprétation de l'incertitude dans le cadre d'une étude. Cette quantité représente la probabilité que la conclusion soit la bonne, malgré la présence

d'incertitudes. Pour les développeurs de modèles, la puissance statistique permet de juger de la signifiante de la différence entre les différents FCs créés pour une même substance. Les résultats préliminaires indiquent que pour la catégorie d'impact de l'utilisation de l'eau sur la santé humaine, même pour les pays comportant plus d'une dizaine de FCs, il ne seraient pas justifié de créer plus de deux ou trois archétypes de FCs. La différence entre un plus grand nombre d'archétypes ne serait pas statistiquement significative.

La thèse argumente en faveur de la création de FCs par industrie. Les proxys d'émission par industrie permettent de localiser les émissions en fonction de la répartition géographique d'une industrie, qui pourrait être très différente de la répartition géographique de la population (l'hypothèse par défaut appliquée jusqu'à présent). Ces proxys d'émission produisent des FCs agrégés plus pertinents. De plus, chaque industrie se retrouve concentrée dans un petit nombre de cellules, ce qui réduit la variabilité spatiale, comparativement à une agrégation qui ne tient pas compte de la source de l'émission.

Finalement, un algorithme de calcul de scores d'impacts, focalisé sur les besoins des analystes, permet la spatialisation ponctuelle de PÉs. L'algorithme permet d'accéder au détail de la chaîne d'approvisionnement, information essentielle si un analyste veut remplacer un FC agrégé par un FC natif. L'accès à ce détail permet aussi d'identifier les points chauds d'incertitude et les meilleures opportunités de réduction de sa réduction.

## ABSTRACT

Life cycle assessment (LCA) is divided in two calculation phase. The first, life cycle inventory (LCI), predicts the quantity of pollutants emitted by each unit process (UP) (facility, transport, mines). Only a minority of those UP can be geolocalized. The others are modeled through a generic LCI database, whose spatial resolution is typically the national scale: the country of a UP is known, but not its position within the country.

The second phase of the calculation, life cycle impact assessment (LCIA), links the emissions to impacts on human health and biodiversity through characterization factors (CF) taking into account the fate, exposure and effect of pollutants. For the regional impact categories like acidification, eutrophication, respiratory effects, etc., the CFs can vary over several orders of magnitude within the same country. In absence of sufficient spatial information, elementary flows have to be characterized by aggregated CFs. This causes an increase in the uncertainty that had not been quantified until now.

This thesis proposes a suite of tools necessary for the management of this uncertainty. The aggregation algorithm allows the quantification of the uncertainty and helps identify the countries where the use of aggregated CFs should be avoided. Preliminary results indicate that for the human health impact of water consumption, spatial variability and uncertainty cover roughly the same range of value for a majority of countries. The lack of spatial information will not be a major source of uncertainty, except for certain cases, identified by the uncertainty increase indicator. The aggregation methodology and the calculation of the indicator will be applied to other impact categories as the uncertainty on native CFs is published.

The thesis introduces the statistical power, a tool borrowed from hypothesis tests, to evaluate the significance of the difference between two random variables, and shows how it should be used in the development of the model and by the practitioners of LCA. Statistical power can be used in the discussion and interpretation phase of an LCA. This quantity represents the probability that the conclusion is the right one, despite the presence of uncertainty. This number can be communicated to the decision-maker and checked against his risk aversion. For the model developer, the statistical power allows to judge the significance of the difference between different native CFs for the same substance. Preliminary results for the human health impact of water consumption indicate that even for countries with more than a dozen CFs, it is not justified



to create more than two or three archetypes of CFs. The difference between a larger number of archetypes would not be statistically significant.

The thesis argues in favor of the creation of industry-specific CFs to reduce even more the spatial variability. The industry-specific emission proxies allow may predict a very different spatial repartition of emissions, compared with population proxies (the default assumption) or industry-generic emission proxies. Those proxies produce more relevant CFs. Also, each industry is concentrated in a smaller number of cells, which reduces spatial variability, compared to an industry generic aggregation.

Finally, an algorithm of impact score calculation, with a focus on the needs of the practitioners, allows for the punctual spatialization of UP. The algorithm reveals all the relevant details of the structure of the supply chain. This information is essential for analysts who want to identify uncertainty hot spots and maximize the reduction of uncertainty.

## TABLE DES MATIÈRES

DÉDICACE.....	III
REMERCIEMENTS.....	IV
RÉSUMÉ.....	V
ABSTRACT.....	VII
TABLE DES MATIÈRES.....	IX
LISTE DES TABLEAUX.....	XII
LISTE DES FIGURES.....	XIII
LISTE DES SIGLES ET ABRÉVIATIONS.....	XV
LISTE DES ANNEXES.....	XVI
INTRODUCTION.....	1
1.1 Introduction.....	1
1.2 Plan de la thèse.....	2
1.2.1 Revue de littérature.....	2
1.2.2 Résumé des articles.....	2
1.2.3 Intégration des outils dans la gestion de l'incertitude.....	3
CHAPITRE 2 REVUE DE LITTÉRATURE.....	5
2.1 Le schéma général de l'ACV.....	5
2.1.1 Objectifs et champ de l'étude.....	5
2.1.2 Inventaire.....	5
2.1.3 Évaluation des impacts.....	6
2.1.4 Interprétation des résultats.....	7
2.2 Incertitude.....	8
2.2.1 Pourquoi traiter d'incertitudes.....	8
2.2.2 Variabilité versus incertitude.....	10
2.2.3 Sources d'incertitudes.....	10
2.2.4 Représentation de l'incertitude sur les données d'entrée.....	12
2.2.5 Propagation de l'incertitude.....	12
2.2.6 Contribution à l'incertitude.....	16
2.2.7 Application de l'analyse d'incertitudes chez les praticiens.....	17
2.3 Inventaire.....	19
2.3.1 Base de données d'ICV et résolution spatiale.....	19

2.4	Analyse d'impacts .....	20
2.4.1	L'état de l'art dans les modèles ÉICV .....	20
2.4.2	Incertitude sur les facteurs de caractérisation .....	23
2.4.3	La spatialisation des facteurs de caractérisation .....	24
2.5	L'interface ICV/ÉICV .....	31
2.5.1	La dichotomie ICV/ÉICV .....	31
2.5.2	Les contributions à l'intégration ICV/ÉICV .....	32
CHAPITRE 3	OBJECTIFS DE RECHERCHE .....	36
3.1	Besoins de recherche .....	36
3.2	Objectifs généraux et spécifiques .....	37
CHAPITRE 4	RÉSUMÉ DES ARTICLES .....	38
4.1	Article 1 : « Quantification of the increase in characterization factor uncertainty due to spatial variability » .....	38
4.1.1	Aperçu de l'article .....	38
4.1.2	Introduction : la dichotomie de résolution spatiale ICV/ÉICV .....	38
4.1.3	La procédure d'agrégation .....	39
4.1.4	La quantification de l'augmentation de l'incertitude .....	40
4.1.5	Application : l'impact de la consommation d'eau sur la santé humaine .....	41
4.1.6	Utilisation envisagée des aCF et nCF .....	42
4.2	Article 2 : « Statistical hypothesis testing tool applied to LCA: Judging the significance of differences» .....	44
4.2.1	Aperçu de l'article .....	44
4.2.2	Introduction : la puissance de test .....	45
4.2.3	Application 1 : cas génériques de distributions lognormales .....	48
4.2.4	Application 2 : comparaison de scores d'impacts et sondage d'analystes .....	50
4.2.5	Application 3 : création d'archétypes de facteurs de caractérisation .....	54
4.2.6	Avantages et limitations de la SP .....	56
4.3	Article 3 : «Relevance and feasibility of sector-specific country-level characterization factors in LCA» .....	57
4.3.1	Aperçu de l'article .....	57
4.3.2	Introduction : le problème des jeux de pondération .....	57
4.3.3	Corrélation entre émissions et population .....	58
4.3.4	Les jeux de pondération et FCs par industrie .....	62
4.3.5	Faisabilité .....	63

4.3.6	Discussion .....	65
4.4	Article 4 : « Systematic disaggregation : a hybrid LCI computation algorithm enhancing interpretation phase in LCA » .....	67
4.4.1	Aperçu de l'article .....	67
4.4.2	L'approche matricielle : intérêts et limites .....	67
4.4.3	L'approche séquentielle : intérêts et limites .....	69
4.4.4	Combinaison des deux approches : la désagrégation systématique .....	69
CHAPITRE 5 DISCUSSION GÉNÉRALE .....		71
5.1	Objectif 1 : Quantification de l'incertitude .....	71
5.1.1	L'agrégation comme source d'incertitudes .....	71
5.1.2	Parallèle entre UII et PS .....	72
5.2	Objectif 2 : réduction de l'incertitude et parcimonie de complexification .....	73
5.2.1	Synergie entre PS et UII .....	73
5.2.2	Synergie entre UII et agrégation par industrie .....	74
5.2.3	Synergie entre PS et agrégation par industrie .....	74
5.3	Objectif 3 : intégration à la pratique de l'analyste .....	75
5.3.1	Spatialisation ponctuelle .....	75
5.4	Retour sur les besoins .....	77
5.4.1	Quantification de l'incertitude et de la compréhension de sa provenance .....	77
5.4.2	Interprétation de l'incertitude .....	77
5.4.3	L'incertitude comme guide de développement des modèles .....	78
5.4.4	La nature itérative de l'ACV .....	78
CHAPITRE 6 CONCLUSION ET RECOMMANDATIONS .....		79
6.1	L'incertitude comme guide d'amélioration du modèle et des données .....	79
6.2	La coordination de haut niveau dans la complexification de l'ACV .....	80
6.3	Développement d'outils pragmatiques d'application .....	81
6.4	Développements à venir .....	81
BIBLIOGRAPHIE .....		82
ANNEXES .....		88

## LISTE DES TABLEAUX

Tableau 1 : calcul d'un score d'impact à l'aide d'un inventaire fictif .....	7
Tableau 2 : Caractéristiques spatiales et temporelles des catégories d'impacts. Tableau tiré de (Owens 1997) .....	22
Tableau 3 : données sur les bassins versants ukrainiens pour l'eau de qualité S2a .....	42
Tableau 4 : détail de calcul du SP basé sur l'exemple de la Figure 17 .....	47
Tableau 5: matrices de la structure computationnelle adaptée.....	64
Tableau 6 : Variables et définitions dans la représentation matricielle en ACV, tiré de (Bourgault, Lesage et al. 2012) .....	68
Tableau 7 : Outils et acteurs impliqués dans l'atteinte des objectifs spécifiques .....	71

## LISTE DES FIGURES

Figure 1 : augmentation de l'erreur avec la sophistication d'un modèle .....	9
Figure 2 : représentation graphique d'une simulation Monte Carlo .....	14
Figure 4 : Schéma systématique d'analyse des incertitudes, tel que recommandé par (Clavreul, Guyonnet et al. 2012) .....	18
Figure 5 : Figure de points chauds (« heat map ») de la table des prix Eora MRIO 2009. Zoom sur la table japonaise entrée-sortie domestique. Tiré de (Lenzen, Kanemoto et al. 2012)....	20
Figure 6 : Compromis entre l'incertitude des paramètres et l'incertitude du choix du modèle. L'augmentation d'incertitude la plus marquée est causée par l'utilisation des produits de transformation. ....	24
Figure 7 : FC spatialisés endpoint pour a) le NO <sub>x</sub> , b) le NH <sub>3</sub> et c) le SO <sub>2</sub> pour l'impact d'acidification aquatique (tiré de (Roy 2012)) .....	27
Figure 8 : FC spatialisé pour l'eutrophisation (Jolliet, Helmes et al. 2010) .....	28
Figure 9 : FC spatialisés pour 7 substances ayant un effet sur la santé humaine au Canada, tiré de (Manneh, Margni et al. 2010). Le découpage du territoire en sous-bassins versants permet de capturer une plus grande variabilité spatiale que le découpage par écozones et provinces. ....	29
Figure 10 : efficacité d'exposition de l'acétaldéhyde en fonction des classes de densité de population, tiré de (Nigge 2001) .....	30
Figure 12 : score générique et spatialisé pour l'acidification, pour 1 kWh de « electricity, low voltage, at grid », tiré de (Mutel et Hellweg 2009) .....	33
Figure 13 : Changements dans la contribution des procédés introduits par la spatialisation des impacts de l'acidification .....	35
Figure 14: Illustration du calcul du aCF et du UII pour des situations simplifiées.....	41
Figure 15 : les étapes de l'agrégation pour les six bassins versants ukrainiens pour l'eau de qualité S2a .....	43
Figure 16: UII pour tous les pays, pour l'eau de qualité S2a .....	43
Figure 17 : Histogramme pour deux VAs (RV) .....	46
Figure 18 : $\mu_2/\mu_1$ VS GSD2 pour des distributions lognormales. Les points 1 à 4 réfèrent aux sous-graphiques de la Figure 19 .....	49
Figure 19 : quatre combinaisons de $\mu_2/\mu_1$ et GSD2 avec leur SP .....	49
Figure 20 : exemple de figure présentée aux analystes; SP = 0,85 et GSD2 = 3,5 .....	52
Figure 21 : résultats généraux du sondage .....	52
Figure 22 : Confort moyen en fonction de la SP et du GSD2 .....	52
Figure 23 : Changement du niveau de confort entre le contexte « gobelet » et « politique énergétique ». Un changement négatif indique un niveau de confort plus faible pour la politique énergétique. ....	53

Figure 24 : Résultats du Monte Carlo et valeurs de SP pour l'eau de qualité S3 en Inde.....	55
Figure 25 : survol international de la catégorie d'impacts pour la qualité d'eau S3.....	56
Figure 26 : Représentation de Bland-Altman pour les émission de phosphate en Angleterre.....	59
Figure 27 : Poids, différences de poids et contribution au FC agrégé pour les émissions de phosphate en Angleterre.....	60
Figure 28 : Poids, différences de poids et contribution au FC agrégé pour les émissions de phosphate en Autriche.....	61
Figure 29: ratio des FC agrégés par poids d'émission et de population, en fonction de la gamme de FC, pour le phosphate (chaque point est un pays).....	61
Figure 30: ratio des FC agrégés par poids d'émission et de population, en fonction de la gamme de FC, pour le NH <sub>3</sub> (chaque point est un pays).....	62
Figure 31 : FCs pondérés par population, émissions de phosphate génériques et par industrie en Italie, eutrophisation. La valeur finale des FCs par industrie dépend d'un plus petit nombre de contributeurs.....	63
Figure 32 : Gamme de FCs par industrie divisé par le FC pondéré par émissions de phosphate, eutrophisation.....	63
Figure 33 : SP en fonction de UII pour deux variables aléatoires.....	72
Figure 34 : Monte Carlo et PS pour les FC de l'eau de qualité S3 en Inde.....	74
Figure 35 : Identification du besoin de spatialisation ponctuelle.....	76
Figure 36 : Schéma d'application de la spatialisation ponctuelle.....	77

## LISTE DES SIGLES ET ABRÉVIATIONS

aCDF	aggregated cumulative density function
aCF	Aggregated characterization factor
ACV	Analyse du cycle de vie
aPDF	Aggregated probability density function
CDF	Cumulative density function
CF	Characterization factor
COV	Composés organiques volatiles
DALY	Disability adjusted life-year
E-PRTR	Europeah Pollutant Release and Transfer Registry
ÉICV	Évaluation des impacts du cycle de vie
FC	Facteur de caractérisation
GES	Gaz à effet de serre
GSD2	Geometric Standard Deviation Squared
ICV	Inventaire du cycle de vie
IO	Input-Output
LCA	Life cycle assessment
nCF	Native characterization factor
NEI	National Emission Inventory
NOEC	No effect concentration
nPDF	Native probability function
NPRI	National Pollutant Release Inventory
PDF	Probability density function
PS	Puissance statistique
TRI	Toxic Release Inventory



## **LISTE DES ANNEXES**

Annexe 1 : Article 1. « Quantification of the increase in characterization factor uncertainty due to spatial variability»

Annexe 2 : Article 2. « Statistical hypothesis testing tool applied to LCA: Judging the significance of differences»

Annexe 3 : Article 3. « Relevance and feasibility of sector-specific country-level characterization factors in LCA»

Annexe 4 : Article 4. « Systematic disaggregation: a hybrid LCI computation algorithm enhancing interpretation phase in LCA »

## INTRODUCTION

### 1.1 Introduction

La création d'un modèle informant une prise de décision demande de trouver un compromis entre simplicité et précision. Minimiser la quantité de données, de variables et d'équations se paie par une augmentation de l'incertitude.

Cette tension entre simplicité et précision est donc le lot de l'analyse du cycle de vie (ACV), qui vise à renseigner des décideurs au sujet des impacts environnementaux des diverses options envisagées. L'honnêteté intellectuelle demande que l'on communique les limites, les incertitudes et les risques d'erreurs associées à l'ACV, afin qu'elles soient pondérées dans la prise de décision.

Jusqu'à présent, la communauté développant l'ACV n'a pas porté à l'incertitude toute l'attention qu'elle méritait. Ignorer cet aspect du modèle nuit à sa crédibilité et à son acceptation auprès des décideurs. À leurs yeux, la capacité à discriminer les scénarios contemplés est la caractéristique essentielle de l'ACV. Un excès d'incertitude menace donc la raison d'être même du modèle.

Il est plus utile de regarder l'incertitude non pas comme un problème à régler, mais plutôt comme un guide attirant notre attention sur les opportunités d'amélioration. La gestion de l'incertitude en ACV ne se limite pas au calcul d'un intervalle de confiance sur les scores d'impacts. Le modèle assemble des données, modèles et hypothèses de plusieurs disciplines, chacune contribuant de manière plus ou moins importante à l'incertitude globale. Comprendre la provenance des incertitudes permettra de prendre les actions les plus pertinentes : nouvelles variables, équations plus complexes, données mieux adaptées ou plus précises.

La gestion de l'incertitude permet donc aux développeurs et aux praticiens de l'ACV de mieux effectuer leur travail. Étant donné la complexité du modèle et la diversité des situations auxquelles celui-ci doit s'adapter, la gestion de l'incertitude s'effectue avec un ensemble d'outils, s'attaquant chacun à un aspect du problème.

Les quatre articles présentés dans le cadre de cette thèse ont été conçus pour être combinés dans le cadre de la réduction de l'incertitude due à l'incohérence d'échelle spatiale à l'interface de

l'inventaire et de l'évaluation des impacts. Les solutions envisagées couvrent donc l'ensemble de la méthodologie.

## **1.2 Plan de la thèse**

### **1.2.1 Revue de littérature**

La revue de littérature débute par une présentation des informations minimales à la compréhension des deux phases du calcul de l'ACV : l'inventaire du cycle de vie (ICV) et l'évaluation des impacts du cycle de vie (ÉICV). La revue souligne ensuite les problèmes suivants:

- Le manque d'informations spatiales à l'inventaire
- L'incohérence des résolutions spatiales à l'interface de l'inventaire et l'analyse d'impact
- Le temps restreint et les outils limités à la disposition du praticien en ACV pour la gestion de l'incertitude

Un résumé étendu des quatre articles est inséré dans le corps du texte (Chapitre 4). Les articles, tels que publiés ou soumis aux différentes revues scientifiques, se trouvent en annexe.

### **1.2.2 Résumé des articles**

#### **1.2.2.1 Article 1 : « Quantification of the increase in characterization factor uncertainty due to spatial variability »**

L'article 1 vise à rétablir la cohérence d'échelles spatiales à l'interface de l'ICV et l'ÉICV. La solution retenue est la combinaison de facteurs de caractérisation natifs (nCFs) de fines résolutions spatiales en des facteurs de caractérisation agrégés (aCFs) à une échelle plus grossière, harmonisée à celle de l'inventaire. Cette agrégation s'effectue à l'aide d'une moyenne pondérée des fonctions de densité de probabilité (PDFs) des nCFs. Leur application respecte la structure computationnelle actuelle de l'ACV. L'article propose un moyen de quantifier l'augmentation de l'incertitude causée par cette agrégation et la compare à l'incertitude des modèles.

### **1.2.2.2 Article 2 : « Statistical hypothesis testing tool applied to LCA: Judging the significance of differences»**

L'article 2 introduit le concept de puissance statistique, tiré des tests statistiques hypothèses. La puissance du test consiste en une probabilité qu'une variable présentant une distribution soit plus élevée qu'une autre. Son application du côté de l'ÉICV permet d'éviter la multiplication d'archétypes de flux élémentaires associés à des facteurs de caractérisation qui ne sont pas significativement différents. Pour l'analyste, la puissance de test permet de prioriser et de justifier une meilleure cueillette de données pour satisfaire le niveau de confort du preneur de décision face à l'incertitude.

### **1.2.2.3 Article 3 : « Relevance and feasibility of sector-specific country-level characterization factors in LCA»**

La phase de calcul d'ICV contient de l'information sur l'identité des usines responsables des émissions. Jusqu'à présent, cette information n'a pas été propagée jusqu'à l'ÉICV pour augmenter la précision du choix des facteurs de caractérisation. Cet article démontre une grande variabilité entre les facteurs de caractérisation spatialisés en fonction des industries.

### **1.2.2.4 Article 4 : « Systematic disaggregation: a hybrid LCI computation algorithm enhancing interpretation phase in LCA »**

L'article 4 développe une méthode de calcul des impacts focalisée sur les besoins de l'analyste. Elle permet d'accéder aux détails de la répartition des impacts dans la chaîne d'approvisionnement, et le remplacement ciblé d'informations génériques par des informations plus représentatives. L'analyste est donc en mesure d'identifier rapidement les meilleures opportunités de diminutions de l'incertitude.

## **1.2.3 Intégration des outils dans la gestion de l'incertitude**

L'ACV a l'ambition de traiter tous les types de produits et services, et d'intégrer tous les types d'impacts environnementaux. Il n'est donc pas réaliste de chercher un unique outil-panacée pour la gestion de l'incertitude.

Les solutions développées au cours des quatre articles constituent une suite d'outils, devant être utilisée de manière séquentielle par divers acteurs. Le dernier chapitre de cette thèse se consacre à la coordination nécessaire à l'atteinte de l'objectif de réduction de l'incertitude.

## CHAPITRE 2 REVUE DE LITTÉRATURE

### 2.1 Le schéma général de l'ACV

L'ACV procède en quatre phases : objectifs et champ de l'étude, l'inventaire, l'évaluation des impacts et l'interprétation des résultats.

#### 2.1.1 Objectifs et champ de l'étude

À cette étape, une entreprise, un gouvernement, une agence de publicité, un groupe de consommateurs ou tout autre parti intéressé à prendre une décision de répondre à un certain nombre de questions dans le cadre de la définition des objectifs :

- Quelles raisons ont mené à la commande de l'ACV ?
- Quelle est l'application envisagée et quel est le public auquel s'adressent les résultats (analyse comparative, éco-conception) ?
- Quels types d'impacts veut-on couvrir ?
- Quelles sont les plus importantes hypothèses de l'étude (contexte géographique spécifique, scénario, etc.) ?

À l'aide d'un analyste, les commanditaires de l'ACV définissent les éléments suivants :

- Le système de produits à étudier
- La fonction du système
- Les flux de références
- Les frontières du système
- Les règles d'imputation
- Autres détails techniques nécessaires au travail de l'analyste.

#### 2.1.2 Inventaire

Les normes ISO 14040 et 14044 délimitent le calcul des scores d'impacts en deux étapes : l'inventaire et l'évaluation des impacts. Toutefois, l'étape d'inventaire se divise aussi en deux étapes : la mise à l'échelle et l'inventaire en tant que tel.

Le résultat de la mise à l'échelle consiste en une liste d'activités économiques, effectuées par des processus élémentaires, et la quantité fournie par chaque processus (Jolliet, Saadé et al. 2005): 2 kWh d'électricité, 4 tonnes d'acier, 1,5 tonne\*km de transport, etc. Ces quantités sont nommées flux économiques et représentent les activités requises pour soutenir directement et indirectement la production d'une unité fonctionnelle. Pour que cette étape produise un résultat réaliste, une grande variété de processus élémentaires doivent être modélisés, couvrant tous les secteurs de l'économie : transport, matériaux de construction, énergie, produits chimiques, processus agricoles, etc. La majorité de cette modélisation, dite d'arrière-plan, est effectuée par les fournisseurs de base de données (voir la section 2.3.1). Le reste de la collecte de données est accomplie par un analyste.

Le calcul de l'inventaire demande aussi d'attacher des consommations de ressources et des rejets de polluants (nommés flux élémentaires) à chaque flux économiques. L'inventaire est calculé en multipliant la mise à l'échelle par l'intensité de consommation de ressources et de rejets de chaque flux économique. Le résultat est une liste de polluants, identifiés par le nom de la substance consommée ou émise, ainsi qu'un compartiment et sous-compartiment (air, haute/basse densité de population, eau de surface/rivière/océan, sol industriel/agricole, etc.) Les intensités des émissions et consommations sont aussi fournies par les bases de données d'inventaire, mais sont aussi compilées à l'avant-plan par les analystes lorsque l'information est disponible.

### **2.1.3 Évaluation des impacts**

L'évaluation des impacts du cycle de vie (ÉICV) relie l'inventaire aux aires de protection (santé humaine, biodiversité des écosystèmes, épuisement des ressources, réchauffement climatique), via un mécanisme de cause à effet (Jolliet, Saadé et al. 2005). Cette quantification s'effectue par la multiplication du total de chaque émission par un facteur de caractérisation (FC), et la sommation des impacts par catégorie. Les unités d'un FC sont (unité de mesure du dommage)/(unité de mesure de l'émission). Les FC tiennent en compte le sort (la fraction du polluant atteignant physiquement le lieu de l'effet), l'exposition (la fraction présente dans l'environnement réellement ingérée par les organismes vivants) et l'effet (la quantification en années de vie perdue ou la diminution de par quantité ingérée) (Bennett, McKone et al. 2002).

Le Tableau 1 montre la structure du calcul d'un score d'impact. L'inventaire fictif représente le total des émissions de tout le cycle de vie d'un produit ou service à l'étude. Chaque flux d'inventaire est multiplié par un FC pour calculer sa contribution à l'impact total. Les contributions sont sommées pour obtenir le score d'impact total.

<b>substance émise</b>	<b>compartiment d'émission</b>	<b>sous-compartiment d'émission</b>	<b>Quantité émise (kg)</b>	<b>FC (DALY/kg)</b>	<b>Impact (DALY)</b>
Dioxine	Air	Faible population	3,00E-06	2,92E+04	8,75E-02
Arsenic	Sol	industriel	2,00E-05	3,10E-01	6,21E-06
Hexachloro-benzène	Air	Haute population	4,30E-04	2,04E-01	8,78E-05
Cadmium	Sol	agricole	4,78E-04	5,02E-02	2,40E-05
Arsenic	Eau	river	3,46E-07	3,17E-02	1,10E-08
Cadmium	Sol	Non-spécifié	1,34E-06	1,23E-02	1,65E-08
Hydrocarbones aromatiques polycycliques	Air	Non-spécifié	1,09E-03	9,91E-03	1,08E-05
Zinc	oll	agricole	0,1	8,59E-03	8,59E-04
				Somme	8,85E-02

Tableau 1 : calcul d'un score d'impact à l'aide d'un inventaire fictif

## 2.1.4 Interprétation des résultats

Finalement, les scores d'impacts sont observés afin de répondre aux objectifs de l'étude. L'analyste répond à plusieurs questions, par exemple, quels sont les principaux contributeurs (substances, usines, section de la chaîne d'approvisionnement, chaîne cause-à-effet) aux scores d'impacts? Les conclusions sont-elles sensibles aux différentes hypothèses? Quelles sont les opportunités de diminution des impacts?

La liste des questions dépend des objectifs et du champ de l'étude, et c'est la phase qui demande le plus de finesse, d'expérience et de jugement de la part de l'analyste.

La norme ISO14044 demande l'inclusion de questions liées à l'incertitude dans la phase d'interprétation. Leur objectif est de repérer la présence ou non de différences significatives, d'identifier des résultats d'ICV de vie négligeables et de guider le processus itératif de l'ACVI. Toutefois, la norme est laconique quant à la manière de répondre à ces questions.

De manière implicite, la norme répartit la responsabilité de la quantification de l'incertitude entre les développeurs des modèles d'ICV (base de données), d'ÉICV (facteurs de caractérisation) et



les analystes (données primaires, propagation des incertitudes et interprétation). Tous ces acteurs ont dû développer des moyens pour pallier le flou artistique créé par la norme autour de ces questions. La suite de la revue de littérature décrit leurs efforts pour créer le vocabulaire, comprendre et maîtriser l'incertitude.

## **2.2 Incertitude**

### **2.2.1 Pourquoi traiter d'incertitudes**

L'objectif d'une ACV est de fournir des informations sur les impacts environnementaux dans un contexte de prise de décision. Or, il est possible que les mauvaises décisions soient prises si les scores déterministes sont tenus en compte sans considérer l'incertitude qui leur est associée (Hammonds, Hoffman et al. 1994; Huijbregts 1998; Ciroth 2004). L'utilisation de scores déterministes permet le classement des options évaluées (impact de option A > impact de option B). Une analyse d'incertitude révèle la probabilité que ce classement soit erroné. Elle constitue une indication de la stabilité de la décision à prendre face à d'éventuels changements dans les données d'entrée (Ciroth 2004).

Discuter des incertitudes est le meilleur moyen de se prémunir contre un excès de confiance envers la signification des résultats (Geisler, Hellweg et al. 2005). La présentation d'estimés ponctuels constitue une surestimation de leur précision aux yeux de quiconque s'étant déjà adonné à la modélisation, ce qui pourrait entraîner la remise en question de la crédibilité de l'ACV (Björklund 2002).

La transparence dans le traitement de l'incertitude révèle la qualité des données et des modèles sur lesquels se base l'ACV. Il s'agit donc d'un puissant moteur incitant à leur amélioration (Björklund 2002; Ciroth 2004). De plus, une analyse d'incertitudes révélant les principaux contributeurs à l'incertitude est souhaitable même si l'analyste n'a pas de moyens pour les réduire (van Zelm et Huijbregts 2013).

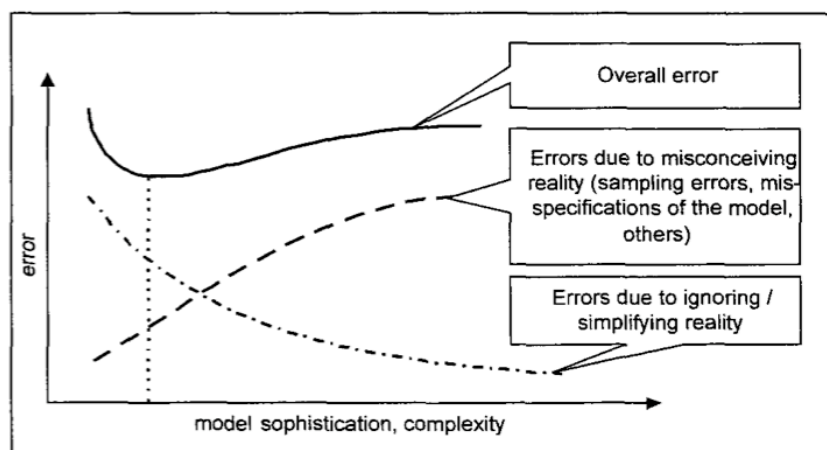
L'incertitude sur les données d'entrée se propage parfois de manière imprévisible vers le résultat final du modèle : de faibles variations s'amplifient, de larges variations s'avèrent insignifiantes (Ciroth, Fleischer et al. 2004). Il n'est donc pas surprenant que la USEPA (United States Environmental Protection Agency) ait souligné la capacité limitée de l'ACV à supporter des

décisions environnementales lorsqu'elle ne fournit que des valeurs déterministes (Lloyd et Ries 2007).

Les analyses d'incertitude sont plus efficaces lorsqu'effectuées de manière itérative. Quand le décideur ne tolère pas l'incertitude révélée, le modèle est revisité et amélioré, autant de fois que nécessaire. Chaque itération devient une occasion d'approfondir la connaissance et la compréhension du modèle (Hammonds, Hoffman et al. 1994).

Les modèles les plus sophistiqués ne sont pas toujours les plus précis. Il existe une tension entre, d'un côté, les erreurs dues à des simplifications, et de l'autre, les erreurs dues à des représentations inadéquates (Ciroth 2004). Tel qu'illustré à la Figure 1, l'erreur totale devrait être minimisée, et ce minimum ne se produit pas nécessairement du côté du modèle le plus sophistiqué.

Une analyse des incertitudes peut révéler le faible *pouvoir discriminant* de l'ACV, dans une comparaison de produits/services. Dans ce cas, aucune conclusion ne saurait être tirée. En l'absence d'une analyse des incertitudes, le *statu quo* et l'absence de conclusion ne sont même pas considérés.



**Fig. 2:** Too high sophistication in modelling raises the overall error, in a model (based on SRU: Umweltgutachten 1974, Stuttgart 1974, p 208, modified)

Figure 1 : augmentation de l'erreur avec la sophistication d'un modèle

Étant donné le potentiel de l'incertitude à paralyser l'application de l'ACV, la réduction de l'incertitude devrait être une priorité. Or, dans le contexte de temps et ressources limités d'une ACV ou du développement d'un modèle, le coût de l'acquisition d'informations supplémentaires

ou plus adéquates doit être mesuré en relation avec la diminution d'incertitude obtenue et désirée (Björklund 2002).

### **2.2.2 Variabilité versus incertitude**

Certains auteurs font une distinction entre variabilité et incertitude (Huijbregts 1998; Hertwich, McKone et al. 2000). Un paramètre sera considéré comme « variable » lorsque sa valeur changera selon les circonstances. On parle alors d'une « incertitude de type A », de paramètre aléatoire, d'une incertitude irréductible ou stochastique. Cette variabilité dépend de facteurs aléatoires ou hétérogènes, tels que les caractéristiques d'un sol, les conditions météorologiques, la densité de population ou la fluctuation de l'efficacité de la combustion dans un moteur. Cette situation se présente lorsqu'un modèle n'admet qu'une seule valeur pour un paramètre qui, en réalité, en adopte plusieurs selon les circonstances. Dans cette situation, la variabilité se transforme en incertitude dans un compromis entre la simplicité d'un modèle et la précision de son résultat.

L'incertitude de type B, aussi appelée « vraie incertitude », réductible, épistémique, se présente lorsqu'une valeur ne peut être connue avec précision (Hertwich, McKone et al. 2000). Ce manque de précision résulte d'erreurs d'estimation, de moyens déficients de mesure, d'un manque de données ou de modélisations inadéquates.

La distinction est importante. La manière de gérer la variabilité et l'incertitude est différente, ainsi que la manière dont on peut espérer les diminuer, les intégrer dans une analyse quantitative ou qualitative, et prendre des décisions en leur présence. Par exemple, si l'incertitude est principalement due à une variabilité spatiale, cette connaissance imposera une grande prudence dans l'extrapolation des décisions vers un contexte géographique différent de celui de l'étude d'origine ou des sources de données.

### **2.2.3 Sources d'incertitudes**

Plusieurs auteurs ont pris la peine de classer les sources d'incertitudes en modélisation. Par exemple, de l'incertitude peut être introduite par le biais de l'exclusion de variables, l'utilisation de variables putatives, la présence de conditions extrêmes dans la nature, non tenu en compte dans le modèle, l'utilisation d'un modèle inadéquat (Hertwich, McKone et al. 2000).

L'incertitude peut apparaître en présence d'ambiguïté ou de controverse au sujet de la quantification d'objectifs sociaux. Cette source d'incertitude est toutefois rarement discutée ouvertement, car elle est perçue comme relevant des sciences humaines (Hertwich, McKone et al. 2000). C'est certainement le cas en ACV lorsque vient le moment de définir les catégories d'impacts et de les agréger, de les comparer entre eux et d'interpréter les résultats. Il est nécessaire d'abandonner l'idée de l'ACV comme outil neutre représentant la réalité telle qu'elle est, et de la remplacer par l'idée de l'ACV comme modèle représentant adéquatement les aspects pertinents de la réalité dans un contexte de prise de décision. Dans cette optique, il est plus facile d'accepter qu'elle puisse contenir des incertitudes.

Sans toutefois avoir conscience des rouages qui l'habitent, chaque être humain est équipé de mécanismes mentaux lui permettant de filtrer et de relativiser l'importance d'éléments contradictoires ou incertains dans le but d'aller de l'avant (Damasio 1997; Edwards 1999). Si l'on considère la complexité du système socio-économique, la phase d'inventaire au sein de l'ACV ne peut être considérée comme une méthode d'attribution des flux de matière et d'énergie sans ambiguïté, objective, exclusivement basée sur des données et des principes scientifiques (Werner et Scholz 2002).

La construction de modèles et leur amélioration requièrent l'identification des sources d'incertitudes d'une manière plus systématique. (Huijbregts 1998) propose une classification en six catégories : paramètre, modèle, choix, temporel, spatial et source/objet. L'incertitude des paramètres fait référence aux inexactitudes empiriques comme des erreurs de mesures, l'incomplétude, la désuétude ou le manque de données. L'incertitude du modèle est la conséquence de simplifications, comme les courbes de dose/réponse linéaires. L'incertitude due aux choix apparaît lorsque l'on doit sélectionner des options de modélisation, comme une méthode d'imputation ou de traitement du recyclage. Huijbregts souligne l'importance de l'incertitude due à la variabilité spatiale. Selon lui, l'inventaire et l'analyse d'impacts devraient être adaptés afin d'éviter la somme arbitraire d'émissions ou d'impacts ayant lieu dans des conditions modifiant considérablement leur intensité. Une remarque similaire s'applique à la variabilité temporelle : la température, la direction des vents, l'efficacité d'une centrale thermique sont des phénomènes cycliques modélisés par des valeurs uniques en ACV. Finalement, la variabilité entre sources et objets représente les écarts inévitables entre, par exemple, la

production d'un même produit avec deux technologies différentes, ou la variation de sensibilité d'une population envers un polluant.

De par leur nature même, les résultats d'une ACV ne se prêtent pas à la validation (Ciroth 2006), ce qui explique en partie l'omniprésence des incertitudes en ACV. Les indicateurs révèlent des impacts potentiels, dispersés sur tout le globe. Seules des validations empiriques d'éléments isolés sont possibles, mais pas les résultats de l'ensemble du modèle.

## **2.2.4 Représentation de l'incertitude sur les données d'entrée**

L'approche préconisée jusqu'à présent en ACV pour quantifier l'incertitude des paramètres consiste à quantifier l'incertitude sur les données d'entrée, afin de la propager jusqu'aux scores d'impacts. L'incertitude sur les données d'entrée est représentée par des fonctions de densité de probabilité (PDF), telles que l'uniforme, la normale, la triangulaire et la lognormale (Heijungs et Frischknecht 2005).

En cas de doute face à la représentativité d'une distribution face à la variable qu'elle décrit, il est possible de tester l'effet des différents choix sur le résultat final à l'aide d'une étude de sensibilité. Toutefois, en général, tant que la moyenne et la variance sont conservées, la forme de la distribution aura peu d'influence sur la moyenne, la variance et l'intervalle de confiance à 90% du résultat final (Hammonds, Hoffman et al. 1994).

## **2.2.5 Propagation de l'incertitude**

### **2.2.5.1 Simulation Monte Carlo**

La simulation Monte Carlo consiste à répéter le calcul demandé par le modèle un grand nombre de fois, en effectuant un tirage aléatoire sur tous les paramètres d'entrée à chaque itération (Kalos et Whitlock 2008). Cette méthode requiert des fonctions de densité de probabilité (PDF) pour les paramètres d'entrée. À la fin d'une itération, le résultat pour chaque indicateur d'impacts est gardé en mémoire. Lorsque le calcul a été effectué un grand nombre de fois (en ACV, typiquement entre 1 000 et 10 000 fois) (Huijbregts, Gilijamse et al. 2003; Peters 2007), la distribution du résultat est compilée en histogramme. La moyenne et l'écart type du résultat peuvent être calculés, ainsi qu'un intervalle de confiance au choix (90%, 95%, 99%). Il est aussi possible d'appliquer différents tests statistiques pour ajuster une PDF à l'échantillon. En ACV, la

distribution du résultat final aura une forme ressemblant à une normale, étant donné le théorème central limite et le grand nombre d'impacts sommés le long du calcul (Kalos et Whitlock 2008) (voir section 4.4.4).

L'échantillonnage stratifié diminue la probabilité de manquer une valeur extrême. Cette méthode consiste à diviser le domaine de chaque paramètre en intervalles de probabilité identique. Chaque intervalle sera donc tiré un nombre égal de fois, ce qui permet une convergence plus rapide (McKay et Conover 1979).

Dans le cas de comparaison de deux cycles de vie, il est pratique courante de calculer le ratio entre les deux scénarios dont le résultat a été obtenu à l'aide du même échantillonnage. L'hypothèse sous-tendue par cette procédure est la corrélation de l'incertitude présente dans les deux cycles de vie. Effectuer des tirages différents pour chaque cycle de vie et comparer le résultat final surestimerait la probabilité d'inversion du classement. La Figure 2 présente la distribution du ratio des scores d'impacts pour différents indicateurs pour deux différents pesticides, sous la forme « box and whiskers » (Geisler, Hellweg et al. 2005). Il n'y a toutefois aucune littérature qui justifie l'une ou l'autre des pratiques et il n'est pas clair non plus si le choix de la pratique a une influence significative sur le pouvoir discriminant de l'ACV (Bourgault, Lesage et al. 2012).

La simulation de Monte Carlo ne repose sur aucune simplification, hypothèse ou autre limitation, comparativement à la méthode analytique. Lorsque le nombre d'itérations est suffisamment élevé, le résultat de la simulation représente avec fidélité le comportement du modèle (Kalos et Whitlock 2008).

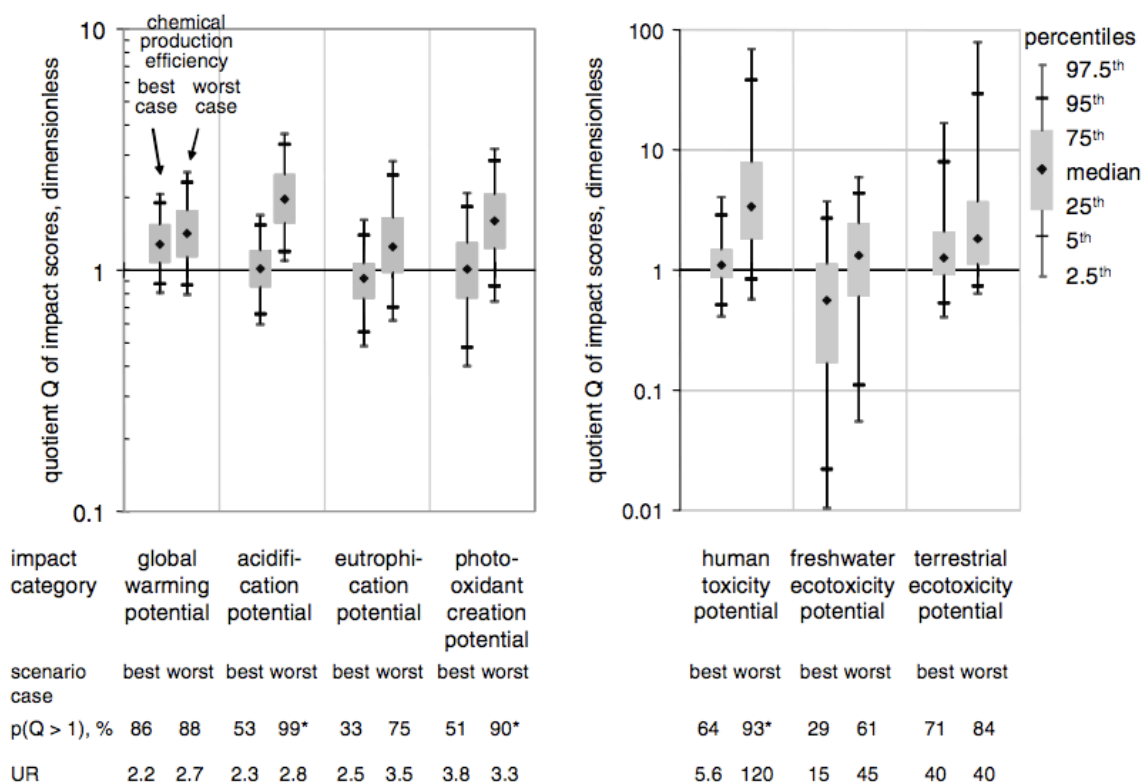


Figure 2 : représentation graphique d'une simulation Monte Carlo

### 2.2.5.2 Méthode analytique basée sur la série de Taylor

La série de Taylor consiste à l'approximation d'une fonction à l'aide d'une fonction polynomiale. On peut tronquer la série selon le niveau de complexité souhaité, mais cette simplification cause un rétrécissement du domaine de validité de l'approximation, ainsi qu'un risque d'erreur. En supposant une fonction lisse et de petites variations de paramètres, il est possible de ne garder que le premier terme. On peut utiliser cette simplification pour approximer la variance de  $y$ . Il faut toutefois supposer que les variables  $x_i$  sont indépendantes les unes des autres :

$$\text{Var}(y) \approx \sum_{i=1}^n \text{Var}(x_i) \left[ \frac{\partial y}{\partial x_i} \right]_{\bar{x}}^2 \quad \text{Équation 1}$$

Cette fonction exprime la variance du résultat final en fonction de la variance des paramètres d'entrée et d'une dérivée partielle, appelée coefficient de sensibilité. La dérivation de ces coefficients a été partiellement effectuée dans (Heijungs et Suh 2002) et complétée dans (Heijungs 2010). Les formules peuvent apparaître complexes à première vue. Leur interprétation est certes difficile, mais leur application est mécanique.

Le résultat de cette approche est l'expression de la variance d'un score d'impact en fonction de la valeur déterministe et de la variance des flux économiques, élémentaires et facteurs de caractérisation. La caractéristique la plus utile de cette équation, dans le cadre de ce projet, est l'expression de la variance par une somme, permettant d'isoler la contribution de chaque variance au total.

Il faut souligner les réserves de cette méthode. Elle suppose de faibles variations ainsi qu'une absence de corrélation entre les variables indépendantes. De plus, dans la nouvelle base de donnéesecoinvent v3, les coefficients des matrices ne sont plus directement des valeurs, mais le résultat d'un calcul à l'aide de paramètres, qui, eux aussi, sont incertains (Weidema, Bauer et al. 2013). Or, les formules permettant de calculer la sensibilité des coefficients ne permet pas d'intégrer cette paramétrisation.

On sait que la variance des paramètres peut atteindre des valeurs assez élevées, particulièrement dans le cas des facteurs de caractérisation. Si un contributeur est identifié comme important dans le cas d'une faible variance, il est raisonnable de supposer qu'il restera un contributeur important dans le cas de variances plus importantes. La valeur exacte de la contribution importe moins que la position dans le classement.

Les travaux de Hugues Imbeault-Tétreault sur une étude de cas ont révélé une bonne adéquation entre les résultats obtenus avec une simulation Monte Carlo et l'approximation analytique, et ce, malgré l'inclusion de  $GSD^2$  de 100 pour les facteurs de caractérisation en santé humaine et qualité des écosystèmes (Imbeault-Tétreault, Jolliet et al. 2013). Cette étude supposait toutefois la corrélation inter-systèmes. Les systèmes étaient effectivement très corrélés, représentant une flotte de véhicule avec et sans système d'optimisation des déplacements.



## 2.2.6 Contribution à l'incertitude

L'identification des sources d'incertitudes ayant le plus d'impact sur l'incertitude du résultat est une tâche importante, révélant la qualité des données et les opportunités d'amélioration, la sensibilité du résultat à certains paramètres ainsi que la possibilité d'augmentation de certitude de la conclusion de l'étude. Par exemple, dans son étude de la variabilité du bilan carbone des serveurs, Weber recommande une stratégie de communication aux utilisateurs, car l'incertitude sur l'impact dépend en grande partie de leur comportement (Weber 2012).

Un petit nombre d'articles traitent des méthodes existantes ou de leur application. Telle que décrite à la section sur la méthode basée sur la série de Taylor (2.2.5.2), la contribution à l'incertitude peut être évaluée à l'aide d'une série de Taylor de premier ordre.

Dans la méthode VUFAF (varier un facteur à la fois) ou par scénario, un paramètre ou un ensemble de paramètres sont variés et le changement sur le résultat du modèle est compilé (Huijbregts, Norris et al. 2001; Björklund 2002). Les avantages de ces méthodes (simplicité, caractère intuitif, modèles calibrés pour des valeurs médianes) en font des choix privilégiés en modélisation, mais présentent aussi plusieurs désavantages, comme les hypothèses de linéarité, d'absence d'interactions entre les paramètres, et la petite fraction explorée du domaine (Saltelli et Annoni 2010).

En présence d'un Monte Carlo, la corrélation de Spearman peut être appliquée entre chaque paramètre échantillonné et les différents scores d'impacts, ce qui a déjà été effectué à quelques reprises en ACV (Huijbregts, Gilijamse et al. 2003; Venkatesh, Jaramillo et al. 2011; Mattila, Leskinen et al. 2012; Weber 2012). Les coefficients les plus élevés indiquent un paramètre plus sensible dans le domaine exploré.

(Mutel, de Baan et al. 2013) proposent l'application à l'ACV d'une méthode en deux étapes : la première étape de triage est effectuée à l'aide de la méthode des effets élémentaires (MOEE : Method Of Elementary Effects); la seconde, appliquée sur les paramètres retenus à la première étape, est le test de contribution à la variance, une méthode explorant le domaine des paramètres dans leur globalité, basé sur la corrélation de Spearman.

## 2.2.7 Application de l'analyse d'incertitudes chez les praticiens

Lloyd et Ries ont effectué en 2007 une revue des pratiques de l'analyse de l'incertitude dans les publications scientifiques (Lloyd et Ries 2007). Vingt-quatre études ont été évaluées en regard de leur traitement de l'incertitude : quels types d'incertitudes ont été considérés; quelle méthode de propagation a été utilisée; comment l'incertitude a été interprétée et présentée. Toutes les études retenues ont rapporté l'incertitude sur les paramètres, 38% sur les scénarios et 33% sur les modèles. La majorité des études (67%) ont utilisé une méthode stochastique (Monte Carlo, Latin Hypercube, bootstrapping), et 29% ont utilisé une analyse par scénario. La grande diversité de méthodes d'analyse et de discussions (inspections visuelles, tests statistiques) démontre le manque de lignes directrices dans les standards ISO (Draucker, Kaufman et al. 2011).

Cette revue montre une nette augmentation du traitement de l'incertitude par rapport à une revue aux objectifs similaires, publiée en 2002 (Ross, Evans et al. 2002). Parmi les 30 études examinées, dix-neuf se déclaraient conformes aux normes ISO. Seulement quatre mentionnaient explicitement les problèmes reliés à l'incertitude, et dix les mentionnaient implicitement. Deux études traitaient le problème de manière qualitative, et une seule de manière quantitative.

L'incertitude a reçu plus de presse depuis 2007. Toutefois, les pratiques n'ont pas significativement évolué. Tel que mentionné par (Cooper, Noon et al. 2012), les analyses tendent à être modulaires et focalisées sur la quantification de l'incertitude sur les données d'avant-plan, spécifiques à l'étude (Ibáñez-Forés, Bovea et al. 2011; Rööös, Sundberg et al. 2011; Venkatesh, Jaramillo et al. 2011; Sills, Paramita et al. 2012; Weber 2012). L'analyse s'étend parfois à l'incertitude fournie par la base de données d'inventaire. Plusieurs études prennent en compte la corrélation entre systèmes de manière cavalière, à l'aide du pairage de chaque itération de Monte Carlo, sans discuter si cette méthode était appropriée (Rööös, Sundberg et al. 2011; Mattila, Leskinen et al. 2012), ou ne mentionnent pas le problème. D'autres concluent que l'empiètement des distributions empêche toute conclusion, sans toutefois quantifier cet empiètement à l'aide de tests statistiques (Venkatesh, Jaramillo et al. 2011). Un niveau de confiance de 90 ou 95% est parfois imposé, sans justifier si ce niveau est approprié au contexte de prise de décision (Mattila, Leskinen et al. 2012; Rack, Valdivia et al. 2013).

Deux études sur les biocarburants ont comparé l'incertitude des paramètres à celles des différentes méthodes de production et les méthodes d'imputation des coproduits (Spatari et

MacLean 2010; Stratton, Wong et al. 2011). Toutes les deux ont conclu que les choix menaient à une plus grande variabilité sur le résultat que l'incertitude sur les paramètres.

(Clavreul, Guyonnet et al. 2012) recommandent un schéma systématique d'analyse des incertitudes, tel que présenté à la Figure 3.

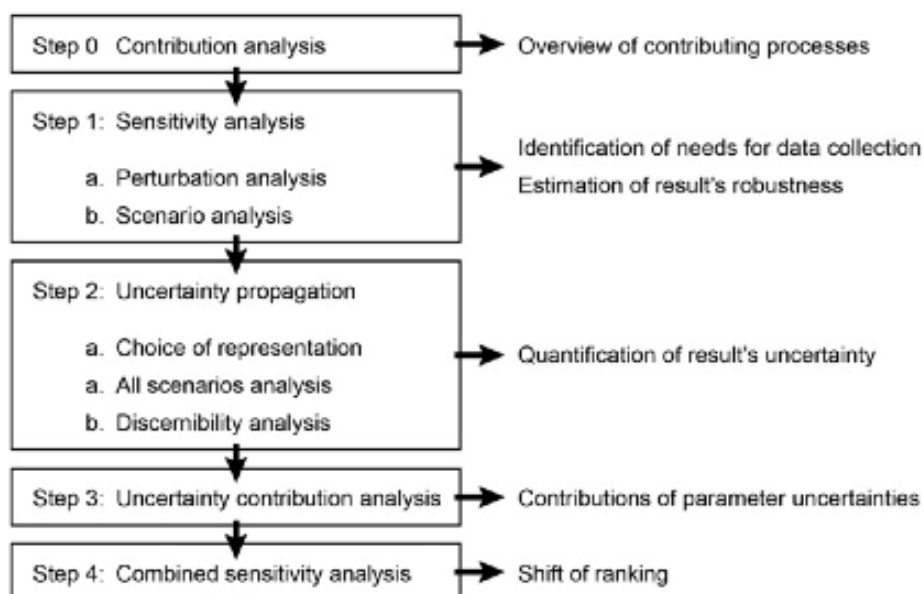


Figure 3 : Schéma systématique d'analyse des incertitudes, tel que recommandé par (Clavreul, Guyonnet et al. 2012)

Bien que plusieurs récents articles ont introduit des nouvelles approches de traitement de l'incertitude, leur intégration dans les bonnes pratiques des analystes n'est pas encore complétée (Baitz, Albrecht et al. 2013). Les analystes dépendent des logiciels commerciaux, qui ne sont pas nécessairement à jour par rapport aux derniers développements. De plus, même si les nouvelles méthodes d'analyse de l'incertitude sont accompagnées d'exemples d'application, une phase de test est requise pour s'assurer de l'acceptation et de la faisabilité des nouvelles pratiques. Finalement, étant donné les exigences des publications scientifiques, on peut supposer que les pratiques rapportées ici constituent les meilleurs cas d'application parmi toutes les ACV réalisées pour des clients commerciaux ou gouvernementaux.

## 2.3 Inventaire

### 2.3.1 Base de données d'ICV et résolution spatiale

Les bases de données d'inventaire sont constituées des matrices technologique et d'intervention, représentant respectivement les échanges technosphère-technosphère et technosphère-biosphère (Heijungs et Suh 2002). Des métadonnées sont aussi fournies afin d'assurer l'utilisation correcte des processus élémentaires et d'expliquer les hypothèses de modélisation, la qualité des données, l'incertitude, la source de l'échantillon, etc. Ces processus élémentaires sont utilisés dans la modélisation d'arrière-plan, et aussi pour compléter la cueillette de données d'avant-plan, lorsqu'elles ne sont pas disponibles.

Du côté de l'ACV basée sur les processus (« process LCA »), la base de données la plus utilisée est ecoinvent (Weidema, Bauer et al. 2013), compilée par un consortium suisse de groupes de recherche. La version 3, publiée à l'été 2013, modélise la production de près de 2 500 produits et services dans tous les secteurs importants de l'économie, et près de 950 polluants et ressources. Les processus élémentaires sont reliés entre eux par l'entremise de marchés, qui répartissent une demande vers les différents pays producteurs. La résolution spatiale est donc majoritairement nationale. Des processus élémentaires représentatifs de régions plus petites sont aussi en développement. Ils sont utilisés par les marchés lorsque suffisamment différents de la moyenne nationale.

Du côté de l'ACV entrée-sortie (« Input/Output LCA »), la base de données Eora constitue le plus ambitieux effort de modélisation couvrant l'économie mondiale (Lenzen, Kanemoto et al. 2012; Lenzen, Moran et al. 2013). La matrice technologique d'Eora couvre 187 pays. Chaque coefficient représente combien de dollars un secteur économique d'un pays dépense dans un secteur économique, local ou étranger (voir Figure 4).

Une mise à l'échelle basée sur cette matrice technologique permet donc de retracer la localisation des émissions à une résolution nationale. Le satellite environnemental (la matrice intervention) n'offre toutefois pas autant de détails que celui d'ecoinvent, répertoriant les principaux GES, NO<sub>x</sub>, COV, SO<sub>2</sub> et NH<sub>3</sub>.

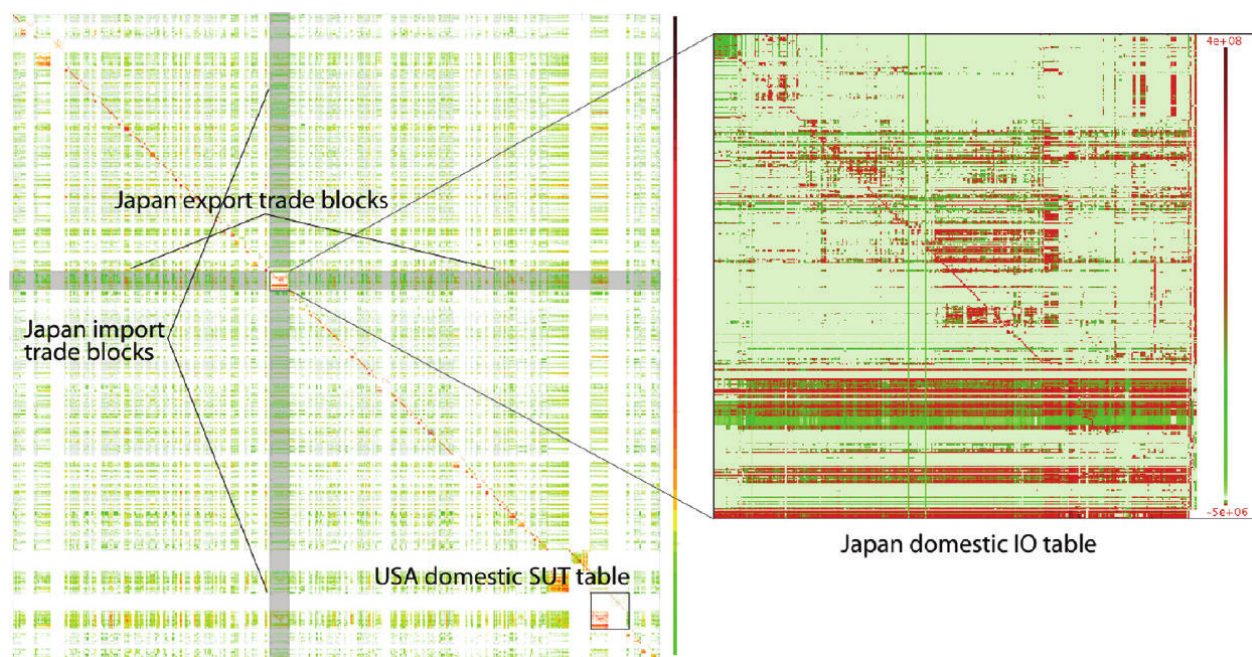


Figure 4 : Figure de points chauds (« heat map ») de la table des prix Eora MRIO 2009. Zoom sur la table japonaise entrée-sortie domestique. Tiré de (Lenzen, Kanemoto et al. 2012).

D'autres bases de données et études, moins exhaustives, ont été publiées récemment, avec une résolution spatiale élevée. Par exemple, (Tessum, Marshall et al. 2012) ont produit un inventaire de  $\text{NO}_x$ , COV,  $\text{PM}_{10}$ ,  $\text{PM}_{2.5}$ ,  $\text{SO}_x$  et  $\text{NH}_3$  causé par l'utilisation d'essence, d'éthanol de grains de maïs et de feuilles et tiges de maïs, désagrégé spatialement sur une grille de 4 km, à résolution temporaire d'heure, pour le territoire des États-Unis. Une étude a quantifié l'impact de la culture du blé d'automne en France (Berthoud, Maupu et al. 2011), à l'aide du modèle de caractérisation USEtox. L'usage de pesticides a été répertorié pour les 6 679 lots pratiquant cette culture, produisant un inventaire spatialisé. Ces études sont caractéristiques des efforts de collection des données d'inventaire en ACV : *ad hoc*, coûteux et déconnectés des pratiques courantes des entreprises y prenant part (Rack, Valdivia et al. 2013).

## 2.4 Analyse d'impacts

### 2.4.1 L'état de l'art dans les modèles ÉICV

L'ÉICV est considérée comme un système d'indicateurs (Owens 1999). Cela signifie qu'elle ne mesure pas d'impacts réels, ne prédit pas d'impacts potentiels, n'estime pas les risques et ne fait

pas de liens directs entre les émissions et des impacts réels. Il a été nécessaire de ramener l'ÉICV au stade d'indicateurs suite à la constatation des nombreuses simplifications derrière le calcul des FC ne permettant pas de quantifier les effets environnementaux des émissions, ainsi que la reconnaissance des aspects subjectifs dans la sélection des catégories d'impacts, indicateurs, modèles, regroupements, pondération et autres procédures (Owens 1999).

Multiplier une quantité émise par un FC implique deux suppositions : toute émission a un impact (aucun effet de seuil), et l'impact augmente linéairement avec la quantité émise (Owens 1997). Ces simplifications se défendent par l'approche « less is better » (Potting et Hauschild 1997). L'ACV ne se préoccupe pas de savoir si une NOEC (no effect concentration) est atteinte, mais plutôt par la probabilité qu'elle le soit. Cette approche est en partie expliquée par la compression spatiale et temporelle nécessaire au calcul. De plus, un cycle de vie calcule les émissions pour un seul système de produits. Or, si une NOEC doit être atteinte ou dépassée, cela sera le fait du cycle de vie à l'étude et de tous les autres cycles de vie contribuant aux mêmes impacts. Les sources identifiées par le cycle de vie peuvent être considérées comme marginales (Potting et Hauschild 1997). En l'absence d'informations plus précises, il faut considérer que chaque kilogramme d'émission supplémentaire contribue également à la probabilité qu'une NOEC soit atteinte.

L'approximation linéaire est plus pertinente dans le cas des substances à effets globaux et des substances transportées sur de longues distances. Par contre, à mesure que les effets deviennent locaux et transitoires, la simplification n'approche plus adéquatement la réalité (Owens 1997). Le Tableau 2 résume la situation pour différentes catégories.

Tableau 2 : Caractéristiques spatiales et temporelles des catégories d'impacts. Tableau tiré de (Owens 1997)

<i>Impact categories</i>	<i>Spatial</i>	<i>Temporal</i>	<b>Dose response</b>	<i>Threshold</i>	<i>Accuracy relevance</i>
Global warming	Global	C/Dc	Linear (?)	No (?)	Good
Stratospheric Ozone	Global	C/Dc	Linear	No	Good
Acidification	Cn/Reg.	Years	Nonlinear	Yes	Fair
Eutrophication	Reg./Local	Years	-Linear	Yes	Fair-Poor
Photochemical smog	Reg./Local	Hours/Day	VNonlinear	No (?)	Poor
Ecotoxicity	Local	H - Years	Nonlinear	Yes	Poor
Habitat loss	Reg. (?)/Local	Dc/Years	Nonlinear	Yes	Very poor
Biodiversity	Reg. (?)/Local	Years	Nonlinear	Yes	Very Poor

**Cn: Continental; Reg.: Regional; C: Centuries; Dc: Decades; (?) Some uncertainty in answer; VNonlinear: Very Non-linear**

Un indicateur midpoint est placé dans la chaîne cause à effet au point où les mécanismes de dommage des différentes substances cessent d'être hétérogènes (Hauschild, Goedkoop et al. 2013). Par exemple, pour le réchauffement climatique, le midpoint est le forçage radiatif. Au-delà de ce point, la chaîne cause à effet vers l'aire de protection est identique, peu importe la substance émise.

Un processus exhaustif de revue des modèles ÉICV a été effectué en 2009 par le Joint Research Centre of the European Commission, afin d'identifier les meilleures pratiques (Hauschild, Goedkoop et al. 2013). La norme ISO 14044 est trop générale pour guider les analystes dans leur évaluation de la pertinence des modèles à utiliser. Une série de critères (35 à 50, selon les catégories d'impacts), a été développée et appliquée à 91 modèles représentant quatorze catégories d'impacts midpoint et endpoint. Ces critères couvraient tant la qualité scientifique que l'acceptation des parties prenantes : complétude du champ d'étude, pertinence environnementale, robustesse scientifique, documentation-transparence-reproductibilité, et applicabilité. Les modèles ont été catégorisés selon 4 niveaux de qualité : recommandé et satisfaisant, recommandé avec améliorations requises, recommandé mais à appliquer avec précaution, intérim.

Même après plusieurs groupes de travail, organisés dans le cadre de la UNEP-SETAC Life Cycle Initiative, un modèle faisant pleinement consensus n'a pas encore été développé, et aucun des modèles ne satisfait réellement l'exigence d'acceptation universelle demandée par la norme ISO.

Plusieurs modèles caractérisant les mêmes substances mènent à des résultats différents, une situation préoccupante qui entrave l'acceptation et l'application de l'ÉICV (Rack, Valdivia et al. 2013).

Les chaînes d'approvisionnement étant la plupart du temps globale, il est nécessaire de développer des modèles applicables à des émissions en provenance de toutes les régions du globe. Du côté des incertitudes, seule le modèle midpoint de potentiel de réchauffement climatique du GIEC fournit des informations quantitatives jugées satisfaisantes. Pour la majorité des modèles, l'information est au mieux incomplète, au pire, inexistante.

Au niveau endpoint, seulement les modèles de particules fines, de la formation d'ozone photochimique et l'effet cancérigène en santé humaine n'ont pas reçu le classement « intérim ». Les autres ont été jugées valides comme base pour les développements à venir, mais non recommandées par le comité.

#### **2.4.2 Incertitude sur les facteurs de caractérisation**

La quantification de l'incertitude n'est pas pratique commune dans le développement des modèles d'ÉICV (van Zelm et Huijbregts 2013). (Björklund 2002) explique en partie cette absence d'analyse par la difficulté à quantifier l'incertitude sur les paramètres des modèles. Malgré cette difficulté, cet aspect est considéré explicitement comme critère d'évaluation des modèles (Hauschild, Goedkoop et al. 2013).

(van Zelm et Huijbregts 2013) ont exploré le concept de point optimal de complexité vis-à-vis de l'incertitude dans le modèle d'impacts des pesticides sur l'écotoxicité aquatique. L'incertitude de huit variations du modèle ont été comparées, combinant toutes les possibilités de deux options pour trois choix de modélisation : facteur d'effet linéaire ou non linéaire, présence ou absence de facteurs de dommage et inclusion ou non de produits de transformation. Les auteurs ont divisé l'incertitude en deux type : incertitude du modèle et incertitude des paramètres. L'incertitude du modèle est calculée comme la médiane d'un modèle divisée par celle du modèle le plus complexe (non linéaire, facteur de dommage, produits de transformation). L'incertitude des paramètres a été calculée comme le ratio du 95<sup>ème</sup> percentile au 5<sup>ème</sup> percentile. Après avoir attribué une distribution aux paramètres pour lesquels de l'information était disponible, un Monte Carlo de 10 000 itérations a été effectué pour propager l'incertitude au résultat final. L'inclusion des



produits de transformation induit la plus grande incertitude, due principalement à l'incertitude sur la toxicité des produits de transformation (Figure 5).

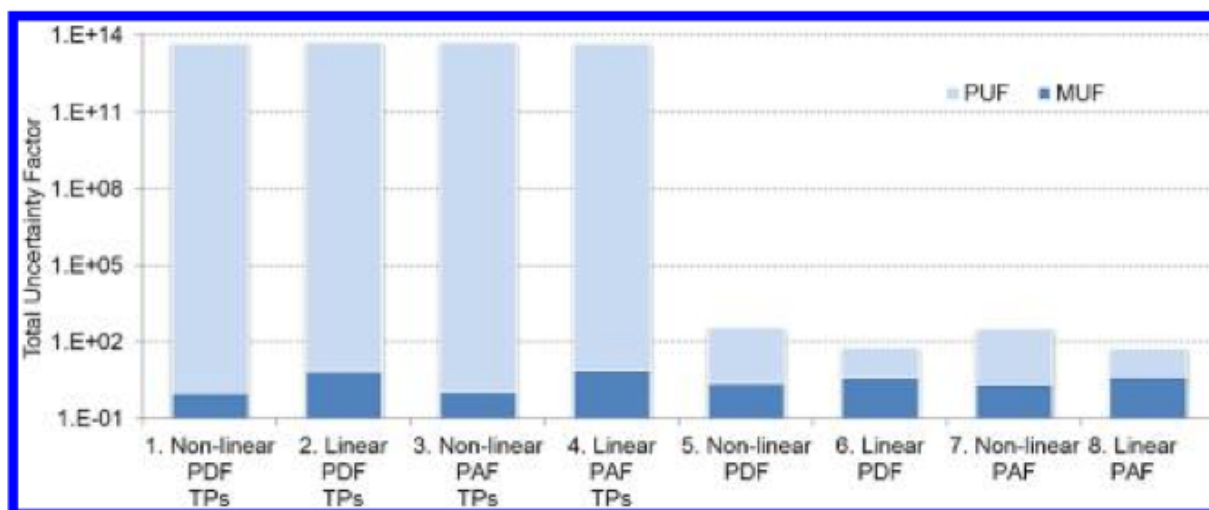


Figure 5 : Compromis entre l'incertitude des paramètres et l'incertitude du choix du modèle.

L'augmentation d'incertitude la plus marquée est causée par l'utilisation des produits de transformation.

### 2.4.3 La spatialisation des facteurs de caractérisation

Il y a maintenant consensus sur la pertinence de la spatialisation en ÉICV, qui se justifie par les constats suivants :

- les chaînes d'approvisionnement sont la plupart du temps globales (Sleeswijk 2010; Wegener Sleeswijk 2011; Baitz, Albrecht et al. 2013; Hauschild, Goedkoop et al. 2013; Rack, Valdivia et al. 2013);
- il est nécessaire d'utiliser le même modèle pour caractériser ces émissions, car les différents modèles reposent sur des hypothèses hétérogènes (Sleeswijk 2010; Hauschild, Goedkoop et al. 2013);
- pour les catégories d'impacts dites régionales, la variabilité spatiale peut atteindre plusieurs ordres de grandeur à l'intérieur du même pays.

Avant l'avènement de la spatialisation en ÉICV, les FCs dépendaient uniquement des caractéristiques physico-chimiques des émissions (temps de demi-vie, coefficient de partition, toxicité (Potting et Hauschild 2006). Or, le concept de fraction prise (sort x exposition) a fait

ressortir l'importance de paramètres géographiques (*landscape parameters* : direction des vents, pluies, etc.) et de paramètres d'exposition (*exposure parameters* : diète des populations, densité de population, densité de production agricole, bioaccumulation dans la chaîne alimentaire, etc.). Les premières études de spatialisation ont fait ressortir que les variations d'un FC sur un territoire étaient au moins aussi importantes que les variations de FC d'une substance à l'autre (Margni 2010). Il y a un consensus autour de l'amplification de ces effets pour les substances à courte demi-vie ou transportées sur de petites distances (Owens 1997; Hertwich, McKone et al. 2000).

#### **2.4.3.1 Facteurs de caractérisation localisés, spatialisés et génériques**

Les premiers efforts de spatialisation ont utilisé la terminologie localisé-spatialisé-générique (Potting et Hauschild 2006). L'article 1 propose une nouvelle terminologie plus appropriée; toutefois, la première terminologie ayant été reprise par presque tous les auteurs, les termes sont quand même définis ici.

Les facteurs *localisés* sont calculés en tenant compte des conditions que rencontreront les émissions ayant lieu dans un site de production précis et connu. Ces facteurs sont applicables lorsqu'une cueillette de données primaires révèle l'identité d'un fournisseur dans la chaîne d'approvisionnement. Cette application sera donc marginale dans le quotidien de l'ACV.

Les facteurs *spatialisés* prennent en compte les conditions d'une surface délimitée par des caractéristiques pratiques selon le média d'émission ou d'autres considérations: un sous-bassin versant, une écozone, une province ou un état, une cellule d'une grille aérienne. Toute émission ayant lieu à l'intérieur de cette surface se verra attribuer le FC attaché à cette surface.

Lorsqu'un modèle ne prend pas en compte la variabilité spatiale des paramètres d'entrée et ne produit qu'un seul FC par substance et par compartiment d'émission, on dit que les FC sont *génériques*.

#### **2.4.3.2 Acidification**

La première catégorie d'impacts à recevoir un traitement spatial détaillé a été l'acidification (Potting, Schöpp et al. 1998). EMEP (Cooperative Program for Monitoring and Evaluation of the long-range transmission of air pollutants in Europe) est un modèle de trajectoire éolien à deux dimensions. Il établit une matrice de transfert « région vers cellule ». Une émission ayant lieu dans une des 44 régions se trouve ainsi répartie sur l'ensemble des 612 cellules de la grille de

déposition. Le modèle RAINS (Regional Air Pollution Information and Simulation) se base sur la matrice de transfert d'EMEP en y ajoutant des informations sur la sensibilité des écosystèmes sur chacune des 612 cellules de la grille de déposition. Le résultat est un ensemble de FC intégrant le lieu d'émission de l'agent acidifiant à une résolution nationale, sa dispersion et une mesure de la sensibilité des écosystèmes où il sera éventuellement déposé. Chacune des substances acidifiantes possède donc 44 facteurs de caractérisation (Seppälä, Posch et al. 2006)

Ces travaux ont fait ressortir une variabilité de trois ordres de grandeur entre le plus faible et le plus élevé FC (Potting, Schöpp et al. 1998). La méthode contient des incertitudes (non quantifiées) mais il est entendu qu'elles sont compensées par l'introduction d'une nouvelle puissance de discrimination en ACV.

Une approche comparable a été retenue par TRACI. Une matrice « état-à-état » trace le déplacement des émissions entre les 48 états contigus des États-Unis. Toutefois, la modélisation s'arrête à la déposition, car aucune base de données sur la sensibilité des écosystèmes n'était disponible pour ce territoire (Norris 2002). Les FC ainsi basés sur cette chaîne cause-à-effet plus courte présentent une plus faible variabilité comparée à la méthode européenne. Le ratio entre le plus élevé et le plus faible FC pour l'acidification de la méthode TRACI est d'environ cinq.

La méthode canadienne LUCAS utilise aussi un modèle sort-effet pour l'acidification. Le facteur d'effet se base sur la charge critique comme indicateur de sensibilité, une donnée compilée à une résolution spatiale provinciale (Fréchette-Marleau, Bécaert et al. 2008). Ainsi, dix facteurs de caractérisation sont produits par substance acidifiante.

Finalement, les travaux de Pierre-Olivier Roy (Roy 2012) ont démontré la pertinence d'une résolution spatiale encore plus fine. En appliquant un modèle sort/sensibilité/effet pour l'acidification terrestre sur une grille 2,5 degrés par 2 degrés, les FCs calculés varient sur une plage atteignant six ordres de grandeur.

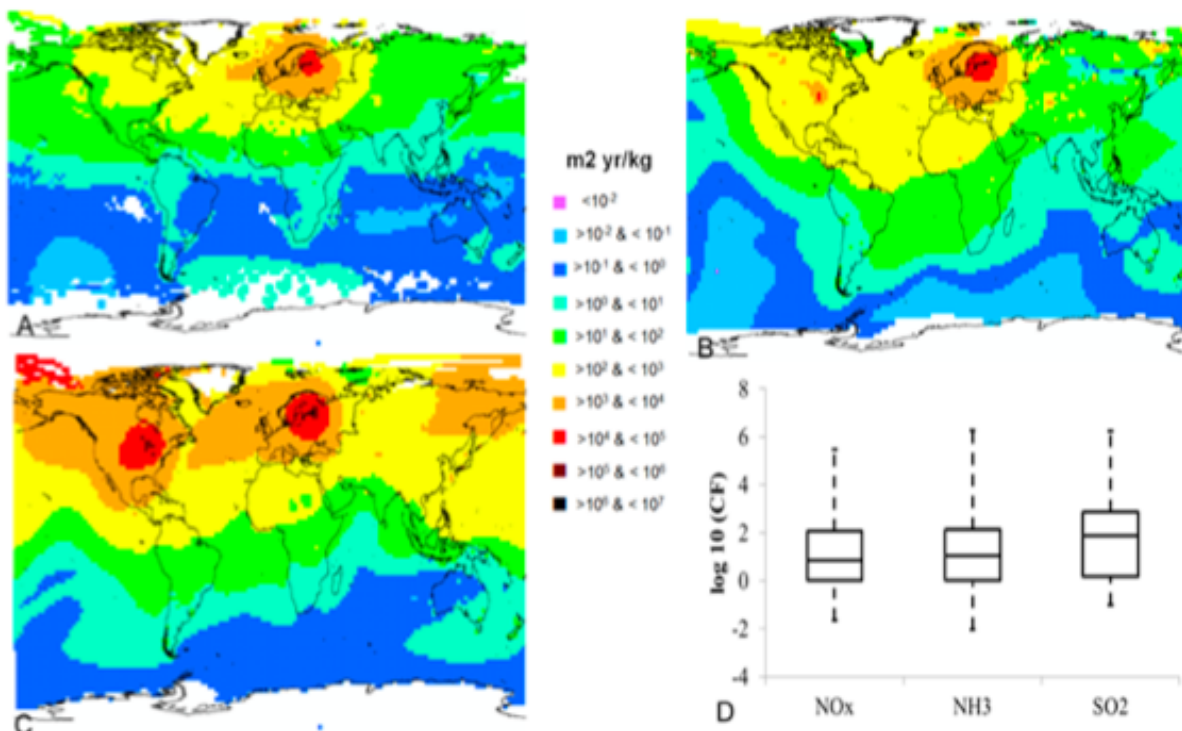


Figure 6 : FC spatialisés endpoint pour a) le NO<sub>x</sub>, b) le NH<sub>3</sub> et c) le SO<sub>2</sub> pour l'impact d'acidification aquatique (tiré de (Roy 2012))

### 2.4.3.3 Eutrophisation

Pour l'eutrophisation, TRACI distingue les zones P-limitantes et N-limitantes. La majorité des eaux côtières sont N-limitantes, et vice versa pour les eaux douces. Les FC développés pour cette catégorie d'impacts couplent cette information avec un modèle de transport aérien ou aquatique. Ainsi, le FC pour une émission aérienne de NH<sub>3</sub> pour un état est compris entre un facteur 0,4 et 1,4 de la moyenne nationale (Norris 2002).

Des FCs spatialisés pour des émissions eutrophisantes dans le sol et dans l'eau sur une grille de résolution 0,5° x 0,5° ont été dévoilés en novembre 2010 (Jolliet, Helmes et al. 2010). Le sort des substances eutrophisantes se base sur la rétention par sédimentation, l'irrigation, l'usage humain ainsi que l'advection par les petits et principaux cours d'eau. Le facteur d'effet est dépendant de la quantité d'eutrophisation ayant déjà eu cours. La variation de sort s'étend sur sept ordres de grandeur, et celle du facteur d'effet, sur deux ordres de grandeur. Les FCs

résultants couvrent donc une gamme de valeurs allant de  $3 \times 10^{-4}$  à  $1,5 \times 10^5$  PAF\*année\*m<sup>3</sup>/kg, tel que présenté à la Figure 7.

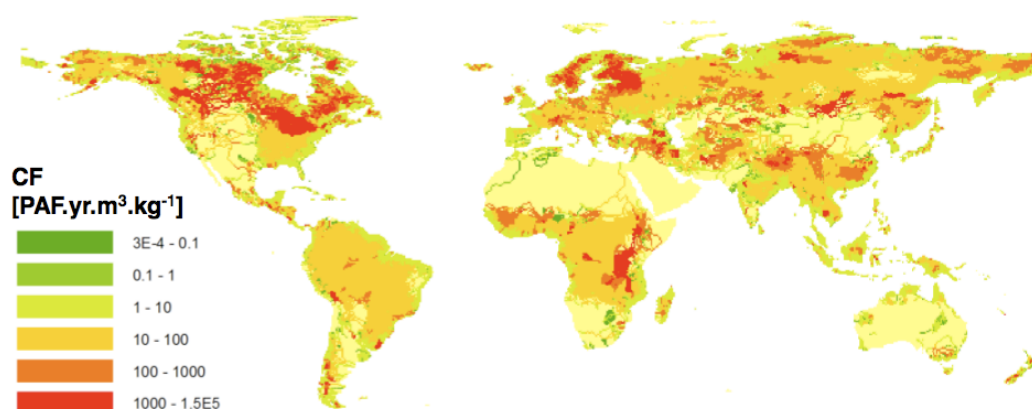
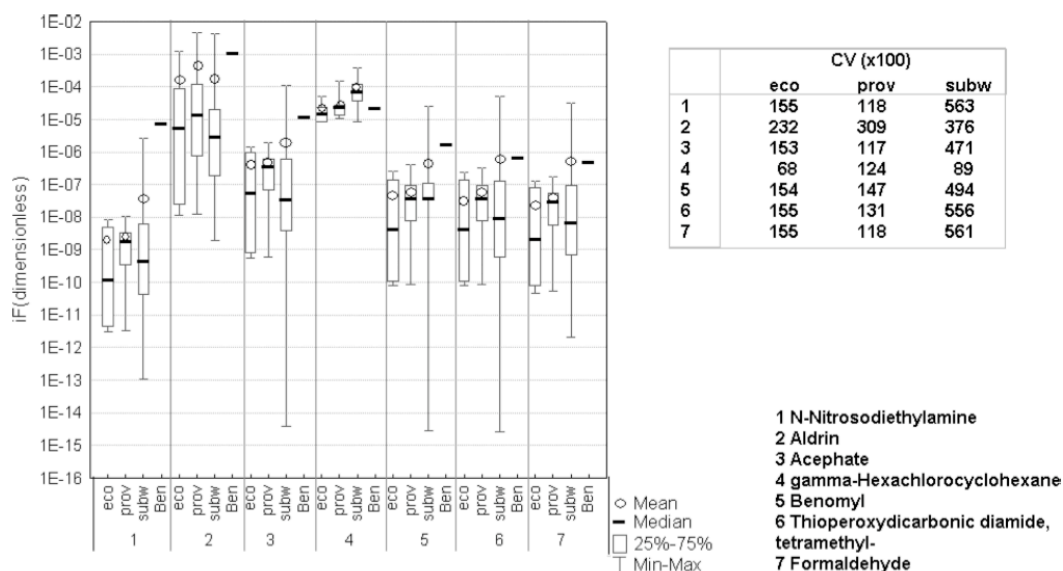


Figure 7 : FC spatialisé pour l'eutrophisation (Jolliet, Helmes et al. 2010)

#### 2.4.3.4 Toxicité humaine et écotoxicité

La variabilité des FC causée par l'hétérogénéité des paramètres de calcul sur un territoire donné a été quantifiée à plusieurs reprises dans le cadre de modèles multimédias (MacLeod, Bennett et al. 2004; Pennington, Margni et al. 2005; Rochat, Margni et al. 2006). D'un continent à l'autre, les FC peuvent varier d'un facteur cinq à dix (Rochat, Margni et al. 2006). Au sein de l'Europe, il a été calculé que la fraction prise pouvait être surestimée ou sous-estimée par trois ordres de grandeur (Pennington, Margni et al. 2005).

De récents travaux sur la variabilité spatiale de la fraction prise en Amérique du Nord ont aussi fait ressortir l'importance de cette variabilité (Humbert, Manneh et al. 2009; Manneh, Margni et al. 2010). La résolution spatiale considérée est l'intersection des sous-bassins versants et d'une grille aérienne de 2° de latitude par 2,5° de longitude.



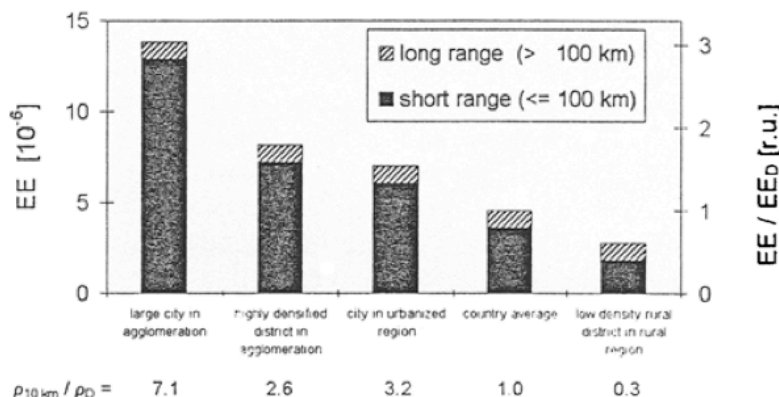
**FIGURE 3. Box plots and coefficients of variation (CV) of iFs calculated for emissions of organic substances (1–7) into water and for the following three spatial resolutions: ecozones (eco), provinces (prov), and subwatersheds (subw). iFs obtained by Bennett and colleagues (Ben) are also shown.**

Figure 8 : FC spatialisés pour 7 substances ayant un effet sur la santé humaine au Canada, tiré de (Manneh, Margni et al. 2010). Le découpage du territoire en sous-bassins versants permet de capturer une plus grande variabilité spatiale que le découpage par écozones et provinces.

En découplant le territoire avec une telle précision, des variations importantes apparaissent entre les différents FC. Pour une même substance, à la résolution des sous-bassins versants, le coefficient de variation se situe entre 8 900 et 56 300 (Figure 8).

Avant l'apparition du concept de fraction prise, Nigge a proposé une classification en archétypes génériques basée sur la densité de population pour les émissions dans l'air (Nigge 2001). La densité de population est effectivement l'un des facteurs principaux déterminant le FC pour la toxicité humaine. La Figure 9 démontre l'influence des différentes classes de densité de population sur l'efficacité d'exposition (équivalent à la fraction prise par inhalation).

L'utilisation de ces classes génériques implique d'attacher à l'inventaire des informations sur la densité de population. La version 2 de ecoinvent contient de telles informations. Sur les 39 555 flux élémentaires émis dans l'air, 6 325 ont reçu la mention « unspecified ». Les autres, soit 84% des flux élémentaires émis dans l'air, ont été classifiés comme « high population density » ou « low population density ».



**Fig. 4:** Spatial variation of exposure efficiencies EE from traffic emissions of acetaldehyde in Germany.  $EE_D$  country average,  $\rho_{10\text{ km}}$  population density within the first 10 km,  $\rho_D$  country average population density

Figure 9 : efficacité d'exposition de l'acétaldéhyde en fonction des classes de densité de population, tiré de (Nigge 2001)

Encore plus récemment, en combinant USEtox et des modèles pour 760 archétypes de sols, il a été démontré que les FC d'écotoxicité terrestre du cuivre et du nickel variaient de 3,5 et 3 ordres de grandeur, respectivement (Owsianiak, Rosenbaum et al. 2013). Pour le cuivre, la variation est principalement liée au pH du sol ainsi qu'à son contenu en carbone organique; pour le nickel, la principale variable affectant le FC est la concentration en  $Mg^{2+}$  en phase aqueuse.

#### 2.4.3.5 Utilisation de l'eau

L'article 1 contenu dans cette thèse utilise les travaux de Anne-Marie Boulay, liant l'utilisation de l'eau à un impact sur la santé humaine (Boulay, Margni et al. 2009). La chaîne cause-à-effet inclut des informations concernant la rareté de l'eau (surface, sous terrain et pluie), la capacité d'adaptation au manque d'eau, la distribution des utilisateurs ainsi que les scénarios de compensation pour pallier le manque d'eau. La plupart des paramètres entrant dans le calcul du FC ont une forte variation spatiale qui n'est pas corrélée avec les frontières politiques. La méthode contient donc 808 cellules de formes et surfaces hétéroclites. Les FCs couvrent jusqu'à sept ordres de grandeur pour la même qualité d'eau.

## 2.5 L'interface ICV/ÉICV

### 2.5.1 La dichotomie ICV/ÉICV

La dichotomie entre la phase d'inventaire et d'évaluation des impacts a été notée il y a longtemps. « These LCA components [inventory and impact assessment] evolved somewhat separately » (Owens 1997). Signe plus récent de cette évolution indépendante, la série de deux articles publiés en 2008 « A survey of unresolved problems in life cycle assessment » (Reap, Roman et al. 2008; Reap, Roman et al. 2008) consacre le premier article aux problèmes inhérents à la définition des objectifs et de l'inventaire, et le second à l'évaluation des impacts et à l'interprétation. Sous-entendu : les problèmes en ACV se situent uniquement dans une phase ou une autre de la méthodologie.

Le problème auquel s'attaque ce projet a déjà été clairement formulé : la compression de la dimension spatiale dans la phase d'inventaire introduit une incertitude dans le calcul de l'impact (Ross, Evans et al. 2002). Cette incertitude est le résultat direct du manque de coordination entre les deux phases centrales de l'ACV.

Observant que l'état actuel de l'inventaire ne permettait pas une application de leurs travaux, (Manneh, Margni et al. 2010) notent : « When calculating human health impacts in LCA, efforts should be focused not only on spatially differentiated intake fractions but also on the emissions of substances. Future efforts may, therefore, address the development of regionalized life cycle inventories. This would imply ensuring consistency between the spatial resolution of inventory data and the resolution of spatially differentiated intake fractions and characterization factors, when necessary. » .

En utilisant la base de données Input-Output américaine, il a été calculé que la modélisation de seulement quinze à vingt contributeurs suffisait à pour capturer 40% des impacts d'un cycle de vie (Norris 2002). Ce nombre monte à 4 000 pour atteindre 60% des impacts. Ce nombre pourrait être encore plus élevé en ACV basée sur les procédés (process LCA), car un contributeur en ACV Input-Output représente potentiellement plusieurs procédés unitaires. Il ne faut donc pas s'attendre à ce qu'un inventaire spatialisé soit produit pour une majorité des impacts par les praticiens en ACV.



Ross et Evans soulignent l'impossibilité, inhérente à l'état actuel de la méthodologie, d'obtenir des informations spatiales plus précises pour l'inventaire (Ross et Evans 2002). Un raffinement méthodologique serait donc nécessaire, intégrant de nouvelles sources d'informations.

Le manque de coordination entre les deux étapes du calcul a encore été remarqué récemment : « the connection between LCI and LCIA sometimes seems to be lost » (Heijungs 2012). Heijungs focalise son commentaire sur l'augmentation irréaliste de la taille des bases de données d'ICV à construire si les FC à haute résolution spatiale devaient être appliqués tels que développés. Dans leur commentaire sur la tension entre les derniers développements scientifiques de la méthode et les meilleures pratiques courantes en ACV, (Baitz, Albrecht et al. 2013) notent : « Regionalization is also a matter of interaction between LCI and LCIA; regionalization in LCI demands for suitable impact assessment evaluation and regionalization in LCIA methods demands a respective basis of regionalized LCI data. »

Dans la liste des moyens pour augmenter la fiabilité des ACV, (Björklund 2002) propose la validation des données. C'est précisément cette validation qu'il n'est pas possible de mettre en application dans le problème du manque d'informations spatiales en ICV.

La dichotomie ICV/ÉICV a aussi été baptisée « data asymmetry » : « Data asymmetries in comparative (i.e. most) LCA studies using more demanding methods may shift problems from LCIA to the LCI databases ». (Klöppfer 2013). L'acceptation des nouveaux modèles spatialisés par les praticiens de l'ACV s'en trouve compromise, car un aspect essentiel de leur travail (l'accessibilité des données) n'a pas été pris en compte dans leur développement. L'applicabilité fait toutefois partie des critères d'évaluation des modèles d'ÉICV : « Characterization factors [are] provided for the important elementary flows for this impact category in a form that is straightforward to apply » (Hauschild, Goedkoop et al. 2013). En d'autres mots, les nouvelles pratiques à implanter doivent s'insérer dans les routines et pratiques, afin d'assurer leur applicabilité (Baitz, Albrecht et al. 2013).

## **2.5.2 Les contributions à l'intégration ICV/ÉICV**

L'équivalence est un concept-clé dans l'intégration ICV/l'ÉICV (Owens 1997). Prenons un exemple simple. Une pomme plus une banane ne donne pas deux pommes-bananes. Toutefois, en regard du guide alimentaire canadien, pommes et bananes sont équivalentes en terme de

portions de fruits. Lorsqu'une telle généralisation est possible, il est possible d'agrèger des quantités autrement disparates. En ACV, le concept d'équivalence intervient à plusieurs étapes du calcul. Une première agrégation est effectuée dans l'étape de mise à l'échelle : les demandes en produits intermédiaires (flux économiques) sont additionnées les unes aux autres. Une seconde agrégation est effectuée en sommant les émissions d'une même substance en provenance de différents procédés unitaires. Finalement, les effets de plusieurs substances sur plusieurs voies d'exposition sont additionnés en ÉICV. Chacune de ces agrégations suppose le concept d'équivalence pour toutes les étapes du calcul en aval de l'agrégation, malgré l'absence de cette équivalence.

C'est pour éviter cette agrégation qu'une adaptation du calcul matriciel a été proposée par (Mutel et Hellweg 2009). Ce logiciel sépare la contribution à l'inventaire de chaque procédé élémentaire, permettant d'appliquer, par exemple, les FC spatialisés au niveau national pour l'acidification. La Figure 10 démontre l'écart qui peut survenir entre le résultat générique et spatialisé.

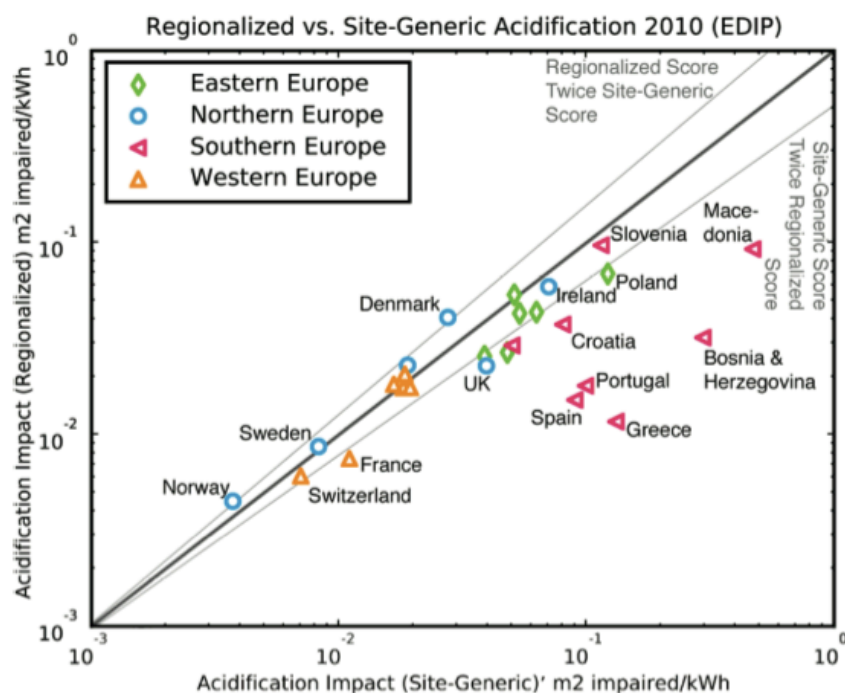


Figure 10 : score générique et spatialisé pour l'acidification, pour 1 kWh de « electricity, low voltage, at grid », tiré de (Mutel et Hellweg 2009)

Un élément de la base de données ecoinvent démontre une pensée englobant l'inventaire et l'ÉICV. Des facteurs de caractérisation existent pour les différents compartiments d'émissions. La base de données donne donc la possibilité aux fournisseurs de données de spécifier le compartiment dans lesquels les flux élémentaires sont émis par le procédé unitaire.

Ciroth affirme que l'intangibilité des résultats de l'ACV empêche une validation sur les résultats (Ciroth 2006). La validation doit alors être faite sur les méthodes de calcul et les sources d'informations. De plus, avant d'implémenter de nouveaux éléments méthodologiques, il faut répondre aux questions suivantes : quelles seront les améliorations ? Dans quelles circonstances ce nouvel élément sera pertinent? Ces deux remarques nous permettent d'énoncer les caractéristiques essentielles des moyens à prendre dans l'intégration inventaire-ÉICV :

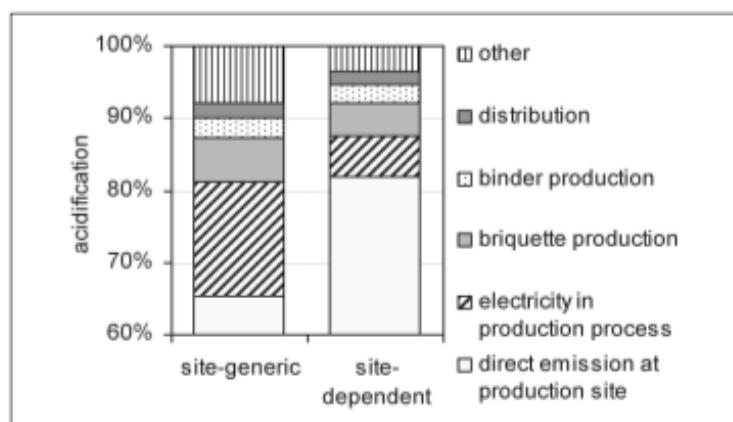
- Source de données de qualité supérieure;
- Méthodes de calcul transparentes et simples;
- Effets mesurables sur la qualité du résultat final;

Jusqu'à présent, les résultats de l'ACV n'ont pas été intégrés avec d'autres sources d'informations comme les surveillances environnementales (type Toxic Release Inventory) ou les analyses de risques (Owens 1997). Ces sources pourraient informer l'ACV sur l'occurrence, la magnitude et la fréquence des effets environnementaux. Ces sources de données sont collectées de manière réglementaire, exhaustive et vérifiée, ce qui en fait des données de haute qualité.

La spatialisation de l'inventaire a été appliquée avec succès sur un petit nombre de procédés unitaires. Cette spatialisation est possible pour les procédés unitaires inclus dans la cueillette de données primaires. En effet, il n'y a aucune impossibilité intrinsèque à l'extension de la cueillette de données primaires. Seul des considérations d'ordre pratique (manque de temps, confidentialité) empêchent son expansion. Ainsi, Ross et Evans ont comparé les résultats d'une ACV générique et spatialisée pour l'impact de la formation de smog photochimique (Ross et Evans 2002). Les impacts de cette catégorie se font sentir d'une manière très locale et transitoire.

(Bellekom, Potting et al. 2006) ont recalculé des scores d'impacts d'acidification pour trois études en incluant des FC localisés. Une première difficulté a été rencontrée dans la désagrégation des données d'inventaire, qui n'étaient pas possibles dans tous les cas. Cela a résulté en l'impossibilité de localiser jusqu'à 20% des impacts. Les contributeurs majeurs ont été identifiés à l'aide des facteurs génériques et c'est pour ceux-là que l'information spatiale a été

recueillie et les FC localisés appliqués. Pour les trois études incluses, les conclusions n'ont pas été inversées, mais le rang des contributeurs majeurs a été perturbé. L'application de cette méthodologie perd un peu de sa pertinence dans les études choisies car le cycle de vie était dominé, dans chaque, cas par un petit nombre de procédés situés dans des endroits où les FCs localisés étaient peu éloignés des FCs génériques. Un contributeur mineur à l'impact pourrait, lorsque spatialisé, se révéler un beaucoup plus grand contributeur, étant donné la variabilité rencontrée dans la spatialisation de cette catégorie. Cette critique est appuyée par la perturbation dans le classement des contributeurs observée à même l'étude (Figure 11).



**Fig. 6:** Relative contribution of process-steps to recalculated acidification of stone wool. Note: the vertical axis does not start at zero

Figure 11 : Changements dans la contribution des procédés introduits par la spatialisation des impacts de l'acidification

## CHAPITRE 3 OBJECTIFS DE RECHERCHE

### 3.1 Besoins de recherche

La revue de littérature a fait ressortir plusieurs besoins de recherche en lien avec l'incertitude et les questions de variabilité spatiale. Ces besoins peuvent être groupés selon les acteurs qui les expriment :

- Développeurs de base de données d'ICV :
  - Création de processus élémentaires contenant de l'information spatiale plus précise
  - Intégration d'informations spatiales provenant de d'autres types de bases de données
  - Quantification de l'incertitude sur la localisation des processus élémentaires
- Développeurs de modèles d'ÉICV :
  - Quantification de l'incertitude sur les données d'entrée des modèles
  - Quantification de l'incertitude totale et de la contribution à l'incertitude des diverses sources
  - Identification du niveau de complexité minimisant l'incertitude
  - Identification de la résolution spatiale optimale
- Développeurs de bases de données d'ICV et de modèles d'ÉICV :
  - Harmonisation de la résolution spatiale
  - Harmonisation des archétypes et sous-compartiments de flux élémentaires
- Praticien de l'ACV :
  - Création d'outils de quantification, d'interprétation et de communication de l'incertitude et de la contribution à l'incertitude
  - Création d'outils d'identification de la localisation des contributeurs majeurs aux scores d'impacts
  - Prise en compte de la corrélation de l'incertitude dans les études comparatives

## 3.2 Objectifs généraux et spécifiques

Ces besoins ne peuvent évidemment pas être tous abordés dans le cadre d'une seule thèse. La présente thèse se focalise sur l'incertitude causée par l'incohérence de résolution spatiale à l'interface de l'inventaire et de l'évaluation des impacts.

<b>Objectif général</b>	Diminuer l'incertitude causée par l'incohérence des échelles spatiales à l'interface ICV/ÉICV
<b>Objectifs spécifiques</b>	Quantifier l'incertitude créée par l'incohérence des échelles spatiales
	Développer des outils de réduction des incertitudes dans l'harmonisation des échelles spatiales permettant de conserver la parcimonie
	Outiller les praticiens de l'ACV dans le choix des données et des facteurs de caractérisation pour minimiser l'incertitude

## CHAPITRE 4 RÉSUMÉ DES ARTICLES

### 4.1 Article 1 : « Quantification of the increase in characterization factor uncertainty due to spatial variability »

#### 4.1.1 Aperçu de l'article

Statut de l'article : soumission initiale à la revue Environmental Science & Technology le 21 mai 2013. Première révision majeure soumise le 12 novembre 2013.

Objectif : développement d'une procédure systématique d'agrégation des FC natifs en FC agrégés et calcul de l'augmentation d'incertitude.

Originalité : comparaison de l'incertitude des FC natifs à la variabilité spatiale à travers un indicateur d'augmentation de l'incertitude.

#### 4.1.2 Introduction : la dichotomie de résolution spatiale ICV/ÉICV

Il existe une friction entre la résolution spatiale de l'inventaire et celle des modèles d'impacts, pour les catégories dites régionales. D'un côté, les bases de données d'ICV spécifient généralement le pays où les processus élémentaires sont situés, mais comme ils représentent des moyennes nationales, ils ne sont pas localisés à l'intérieur du pays. De l'autre côté, les derniers développements en spatialisation des facteurs de caractérisation ont révélé d'importantes variations (jusqu'à plusieurs ordres de grandeur) au sein d'un même pays.

Il a été reconnu il y a plus d'une dizaine d'années que la dichotomie de résolution spatiale ICV/ÉICV augmentait l'incertitude sur les scores d'impacts (Ross et Evans 2002). Il n'était toutefois pas possible à l'époque de quantifier cette augmentation de l'incertitude, car la variabilité spatiale des FCs n'était pas encore connue.

Maintenant que les modèles d'ÉICV calculent des FCs à une résolution plus fine que celle de l'ICV et que l'incertitude sur les FC commence à être rapportée, il est possible de comparer la variabilité spatiale à l'incertitude. On peut ainsi découvrir si d'ignorer cette variabilité cause une augmentation significative de l'incertitude.

L'article 1, central à cette thèse, propose deux éléments :

- Une méthode d'agrégation des FCs natifs (native characterization factors : nCF) en des FCs agrégés (aggregated characterization factor : aCF) à une résolution spatiale cohérente à celle de l'ICV;
- La quantification de l'augmentation de l'incertitude dans le passage des nCF aux aCF, permettant d'identifier rapidement les meilleures opportunités de réduction de l'incertitude.

L'échelle dite « native » représente la délimitation jugée optimale par les développeurs des modèles d'ÉICV. C'est cette échelle qui est publiée, bien que des échelles plus fines peuvent avoir été développées.

### 4.1.3 La procédure d'agrégation

L'application de la procédure d'agrégation suppose les conditions suivantes :

- Un ensemble de nCFs dont les incertitudes ont été quantifiées de manière statistique (une PDF est disponible pour chacun d'eux);
- Un inventaire spatialisé à une échelle plus grossière (pays, continent, province/état) que celle de l'ÉICV pour la catégorie d'impacts considérée;
- Un proxy du taux d'émission de la substance considérée pour le territoire de chaque nCF, représentant la probabilité qu'une émission ait lieu sur une cellule par rapport aux autres cellules au sein du même pays.

Dans un premier temps, les proxys de taux d'émission sont normalisés :

$$\hat{w}_{i,j,k,m} = \frac{w_{i,j,k,m}}{\sum_{i=1}^{n_{j,m}} w_{i,j,k,m}} \quad \text{Équation 2}$$

où  $w_{i,j,k,m}$  est le proxy du taux d'émission du flux élémentaire  $k$  sur la zone géographique  $i$  (résolution ÉICV) pour la catégorie d'impacts  $m$ ,  $n_{j,m}$  est le nombre de zones géographiques de résolution ÉICV à agréger à l'intérieur de la zone géographique  $j$  de résolution ICV;  $\hat{w}_{i,j,k,m}$



représente donc la probabilité (entre 0 et 1) qu'une émission soit relâchée sur la section  $i$  de la zone géographique  $j$ .

Ces proxys normalisés sont ensuite utilisés dans le calcul d'une moyenne pondérée des PDF de chaque nCF (nPDF), créant la PDF du aCF (aPDF) :

$$aPDF_{j,k,m} = \sum_{i=1}^{n_{j,m}} nPDF_{i,j,m} \cdot \hat{w}_{i,j,k,m} \quad \text{Équation 3}$$

#### 4.1.4 La quantification de l'augmentation de l'incertitude

Il est possible de comparer l'incertitude de la aPDF avec celle des nPDF. La quantité d'intérêt est ici le coefficient de variation (CV = écart type / moyenne). L'indicateur d'augmentation de l'incertitude (uncertainty increase indicator, UII) est défini comme suit :

$$UII = \frac{CV_{aCFj} - \min_i(CV_{aCFi})}{\max_i(CV_{aCFi}) - \min_i(CV_{aCFi})} \quad \text{Équation 4}$$

où  $CV_{aCFj}$  est le coefficient de variation du aCF<sub>j</sub>,  $\min(CV_{aCFi})$  et  $\max(CV_{aCFi})$  sont respectivement le plus petit et le plus grand coefficient de variation des nCF<sub>i</sub> avec lesquels le aCF<sub>j</sub> a été construit.

La Figure 12 montre trois exemples fictifs d'agrégation de deux nCFs. Dans le premier, les deux nPDFs se chevauchent presque entièrement. Le poids de chaque distribution est de 50%. La aPDF résultante a donc un aCV à mi-chemin entre celui des deux nCVs, ce qui mène à un UII inférieur à 1. Dans le second exemple, les distributions ont peu de chevauchement. Les poids ont été maintenus à 50%. La aPDF résultante exhibe deux modes. Le aCV est plus élevé que celui des deux nCVs, le UII est donc supérieur à 1. Le troisième exemple est identique au second, sauf pour les poids : 90% pour la bleue, 10% pour la rouge. Le UII s'en trouve réduit, mais toujours supérieur à 1.

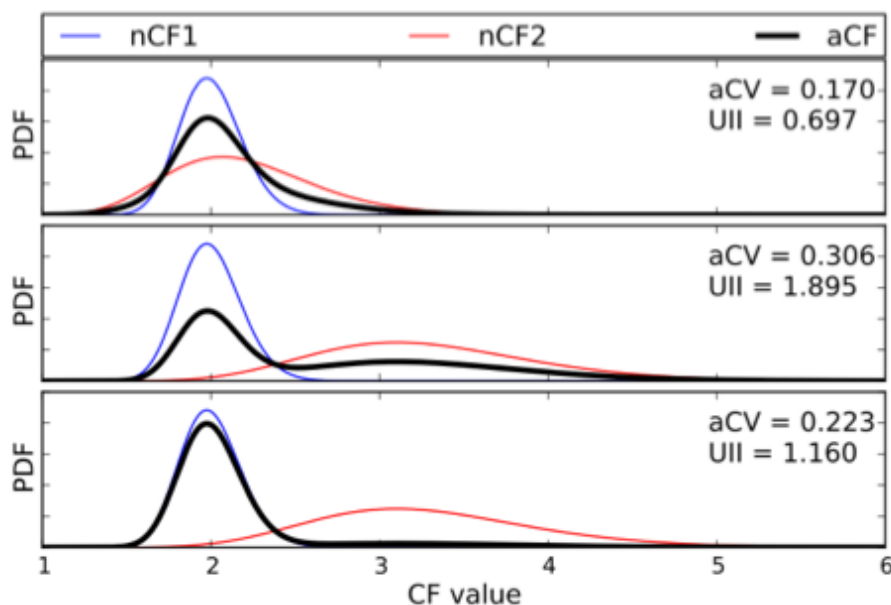


Figure 12: Illustration du calcul du aCF et du UII pour des situations simplifiées

Le UII doit être interprété prudemment. Une petite valeur de UII indique une variabilité spatiale comparable à l'incertitude. Dans ce cas, substituer un aCF par un nCF ne produirait pas de réduction significative de l'incertitude sur le score d'impacts. Des petites valeurs de UII peuvent aussi apparaître dans le cas d'agrégation de nCF se chevauchant peu, mais dont la probabilité d'émission dans les régions à nCF extrêmes est très faible.

Lorsque les nPDF deviennent de mieux en mieux différenciées, les valeurs de UII augmentent, mais il n'y a pas de changement qualitatif clair entre  $UII < 1$  et  $UII > 1$ . Des hautes valeurs de UII indiquent que la variabilité spatiale entre les nCF à agréger est significativement plus grande que l'incertitude des nCF.

#### 4.1.5 Application : l'impact de la consommation d'eau sur la santé humaine

Le modèle sélectionné pour l'illustration de l'agrégation est l'impact direct sur la santé humaine de l'utilisation de l'eau douce (Boulay, Bulle et al. 2011). Dans ce modèle, un nCF a été produit pour chaque intersection pays/bassin versant. Pour les pays complètement inclus dans un seul bassin versant, l'agrégation n'est pas pertinente ( $nCF = aCF$ ). À l'autre extrême, la Chine se divise en 22 bassins versants. L'incertitude des nCF a été calculée à l'aide d'un Monte Carlo et est représentée par des distributions lognormales. Dans ce cas-ci, on parlera d'un proxy spatial

de consommation plutôt que d'un proxy d'émission. La répartition de la consommation d'eau a été obtenue à l'aide du modèle WaterGAP 2 (Alcamo, Döll et al. 2003).

Le Tableau 3 montre les données pour les six bassins versants ukrainiens. Le premier graphique de la Figure 13 affiche les nPDF correspondant à ces bassins versants. Les points rouges sur l'axe des abscisses représentent la moyenne des distributions. Le second graphique montre les nPDF pondérées par les proxys de consommation, ce qui permet de voir quelle nPDF dominera la aPDF. Le troisième graphique représente la aPDF. Finalement, le quatrième graphique montre que la fonction de distribution cumulative (CDF) associée à la aPDF converge vers 1.

Tableau 3 : données sur les bassins versants ukrainiens pour l'eau de qualité S2a

Bassin versant	Distribution	Moyenne (DALY/m <sup>3</sup> )	GSD2	CV	Consommation d'eau (km <sup>3</sup> /année)	Poids relatif
Danube	Dirac	0	NA	0	7,94E-04	10,0%
Dnieper	Lognormale	1,61E-04	10,9	1,78	2,65E-03	33,3%
Outside main watershed*	Lognormale	1,61E-04	9,50	1,60	3,12E-03	39,2%
Dniester	Lognormale	1,97E-04	10,6	1,74	7,12E-04	8,9%
Don	Lognormale	2,88E-04	11,2	1,81	6,28E-04	7,9%
Vistula	Lognormale	4,83E-04	11,8	1,89	3,94E-05	0,5%

\* Région composite n'incluant pas de bassins versants classifiés comme « principaux ».

Le CV du aCF ukrainien est de 1,749, entre la valeur maximale (1,89) et minimale (0) des nCV. Le UII est donc de 0,853. La Figure 14 permet de visualiser les UII pour tous les pays. On peut conclure que pour la qualité d'eau S2a, la variabilité spatiale est comparable à l'incertitude dans la majorité des cas, leur UII se situant majoritairement sous 1.

#### 4.1.6 Utilisation envisagée des aCF et nCF

Lorsque possible, les nCF devraient être utilisés, car leur représentativité de la situation locale est meilleure. L'analyste doit toutefois posséder les informations nécessaires pour justifier leur utilisation, ce qui arrivera presque exclusivement pour les processus élémentaires d'avant plan. En l'absence d'informations spatiales, les aCFs devraient être utilisés.

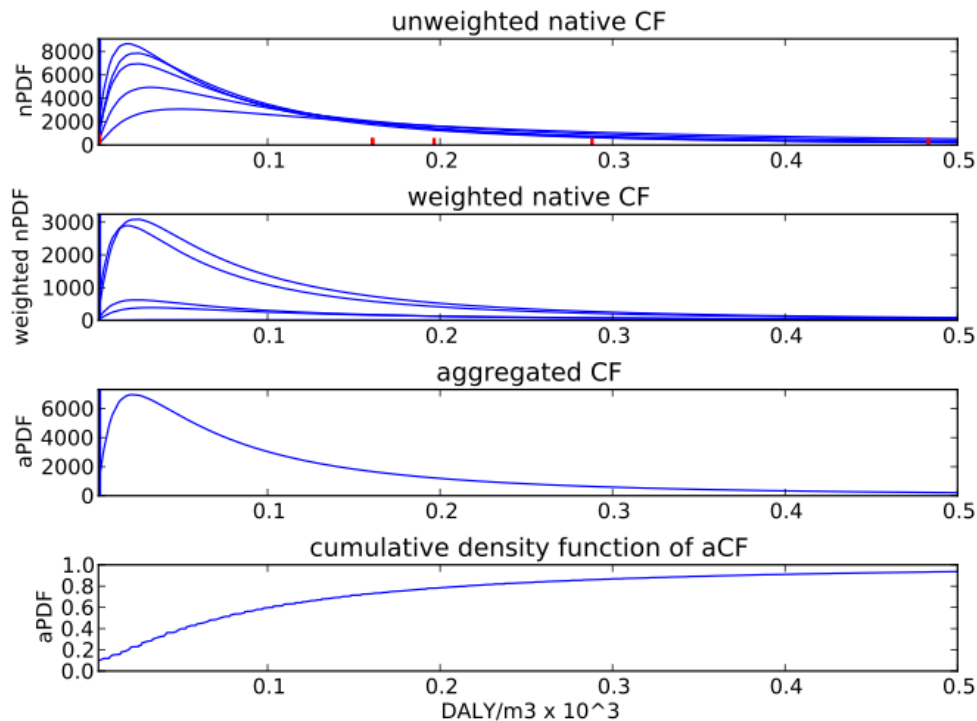


Figure 13 : les étapes de l'agrégation pour les six bassins versants ukrainiens pour l'eau de qualité S2a

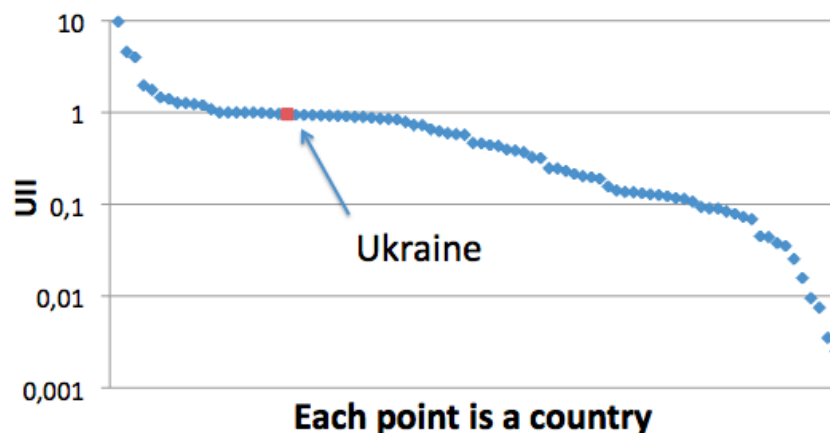


Figure 14: UII pour tous les pays, pour l'eau de qualité S2a

L'application des aCFs n'est pas idéale, mais rendue nécessaire par le manque d'informations spatiales des bases de données d'ICV. Toutefois, il est possible de minimiser leur utilisation en référant aux UII. Si les contributeurs majeurs au score d'impacts proviennent de aCFs dont le UII est élevé, une analyse de scénario s'impose comme première étape pour vérifier si le remplacement du aCF par l'un des différents nCFs a le potentiel de changer les conclusions de

l'étude. Dans l'affirmative, l'effort de cueillette de données primaires sur la situation géographique de l'usine responsable de l'émission/consommation pourrait être justifié.

De plus, les contributeurs mineurs au score d'impacts dont le UII est élevé pourraient devenir de plus importants contributeurs si le aCF était remplacé par un des nCF. Une analyse de scénario serait aussi recommandée dans cette situation. En fait, un recalcul complet ne serait pas nécessaire dans tous les cas. Si un flux élémentaire se retrouve comme un contributeur à 0,1% du score et que la valeur moyenne du plus haut nCF est de seulement un ordre de grandeur supérieur à celle du aCF, le flux élémentaire ne pourra pas dépasser une contribution de 1%, ce que l'analyste pourra juger comme méritant ou non une intervention.

Si l'incertitude diminue le pouvoir discriminant de l'étude au point de paralyser toute conclusion, l'analyste pourra vérifier si une cueillette de données primaires visant la spatialisation a le potentiel, dans le meilleur cas, de produire une réduction de l'incertitude assez grande pour parvenir à une conclusion. Les UII vont pointer vers les plus larges réductions de l'incertitude, et l'analyste pourra juger si la réduction vaut l'investissement en temps et/ou en ressources. La réduction de l'incertitude due au manque d'informations spatiales devient alors l'un des critères permettant de prioriser la cueillette de données.

## **4.2 Article 2 : « Statistical hypothesis testing tool applied to LCA: Judging the significance of differences»**

### **4.2.1 Aperçu de l'article**

Statut de l'article: soumission initiale à la revue Journal of Industrial Ecology le 12 novembre 2013.

Objectif: Introduction à l'ACV de l'outil « puissance statistique », utilisé couramment pour tester des hypothèses en épidémiologie et en contrôle de la qualité. Démonstration de l'utilité pour les analystes et les développeurs de modèles.

Originalité : Sondage auprès d'analystes comparant leur perception de l'incertitude à la mesure de puissance statistique. Démonstration de l'impossibilité de justifier une distinction entre tous les FC natifs au sein d'un pays dans la catégorie d'impact de l'utilisation de l'eau sur la santé humaine. Discussion de l'intégration de la puissance statistique comme outil d'interprétation de l'incertitude et de la communication du risque aux décideurs.

#### **4.2.2 Introduction : la puissance de test**

Les tests statistiques sont utilisés pour quantifier la certitude d'une conclusion. Même si les moyennes de deux variables aléatoires (VAs) sont telles que  $\mu_2 > \mu_1$ , la conclusion  $VA2 > VA1$  n'est pas certaine à 100%. La probabilité qu'elle soit fautive est connue sous le nom « d'erreur de type II », ou  $\beta$ . La quantité  $1 - \beta$  représente la puissance (statistical power : SP), ou la sensibilité d'un test statistique (Hines, Montgomery et al. 2003). SP représente la probabilité d'accepter une hypothèse nulle alors qu'elle est vraie. Ce concept est applicable à plusieurs situations en ACV. La méthode de calcul de la SP est tout d'abord détaillée ci-dessous. Trois applications sont ensuite discutées en détail.

Le calcul de la SP s'illustre plus facilement à l'aide d'un histogramme (voir Figure 15 et

Tableau 4 pour les informations sur l'exemple).

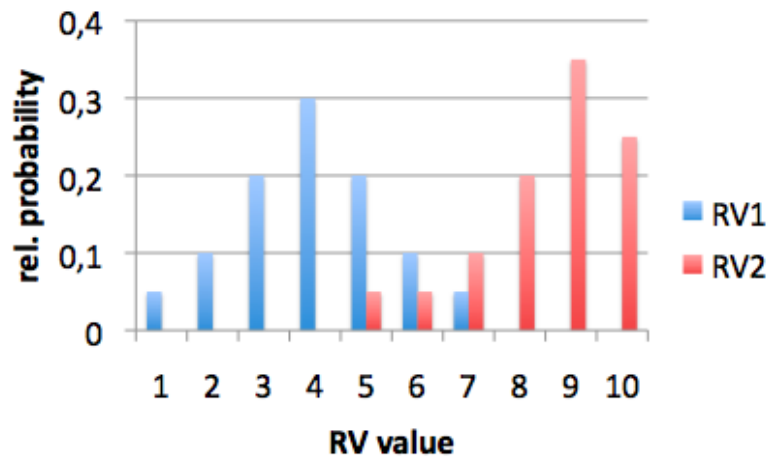


Figure 15 : Histogramme pour deux VAs (RV)

Tableau 4 : détail de calcul du SP basé sur l'exemple de la Figure 15

Valeur de la VA	P(VA2 = valeur)	P(VA1 < valeur)	P(VA2 = valeur & VA1 < valeur)
5	0,05	0,65	0,0325
6	0,05	0,85	0,0425
7	0,1	0,95	0,095
8	0,2	1	0,2
9	0,35	1	0,35
10	0,25	1	0,25
<b>Total</b>			<b>SP = 0,97</b>

La moyenne de VA2 est supérieure à celle de VA1, mais étant donné le chevauchement, il existe une probabilité de piger une valeur de VA1 supérieure à une valeur de VA2. Cette probabilité se calcule tel que détaillé au



Tableau 4. Si les VAs sont représentées par des variables continues, la formule devient :

$$SP = \int_{-\infty}^{+\infty} CDF_{RV1} \cdot PDF_{RV2} dx \quad \text{Équation 5}$$

où CDF est la fonction de densité de probabilité cumulative. En choisissant  $VA1 = VA2$ ,  $SP = 0,5$ .

Dans l'application à l'ACV, l'hypothèse nulle est que les deux quantités comparées ne sont pas différentes. On calcule ensuite la probabilité que  $VA2 > VA1$  pour rejeter l'hypothèse nulle.

### 4.2.3 Application 1 : cas génériques de distributions lognormales

La valeur de la SP est une fonction de la moyenne, de l'écart type et de la distribution de deux VAs. La distribution lognormale étant la plus utilisée en ACV, la SP a été étudiée en présence de deux distributions de cette forme. La mesure typique de l'incertitude est l'écart type géométrique (GSD2) (Heijungs et Frischknecht 2005). Pour réduire le nombre de paramètres dans cette exploration, il a été vérifié que la SP était constante lorsque le ratio  $\mu_2/\mu_1$  et les GSD2 restaient constants. Par exemple, avec  $GSD2_1 = GSD2_2 = 2$ ,  $SP = 0,9218$  peu importe les valeurs de  $\mu_2$  et  $\mu_1$ , si  $\mu_2/\mu_1 = 2$ . Aussi, pour simplifier l'analyse, on a conservé  $GSD2_1 = GSD2_2$ . Pour cette analyse,  $SP = f(\mu_2/\mu_1, GSD2)$ .

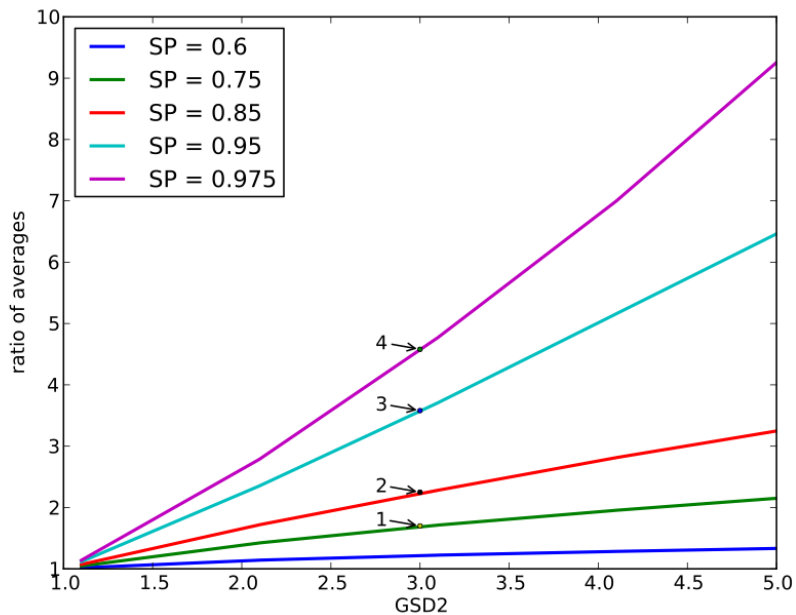


Figure 16 :  $\mu_2/\mu_1$  VS GSD2 pour des distributions lognormales. Les points 1 à 4 réfèrent aux sous-graphiques de la Figure 17

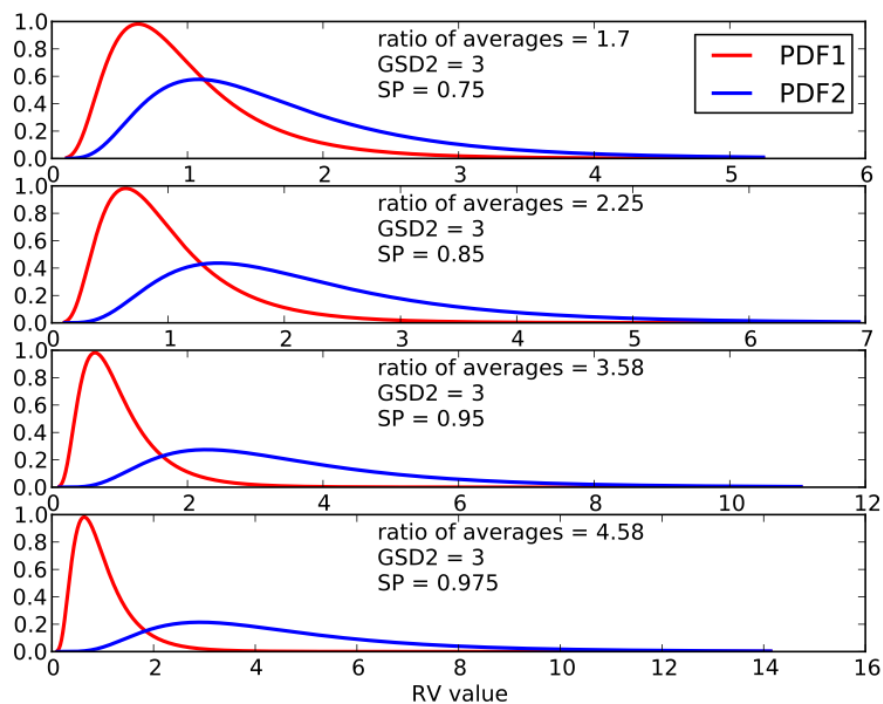


Figure 17 : quatre combinaisons de  $\mu_2/\mu_1$  et GSD2 avec leur SP

La Figure 16 s'interprète comme suit : en connaissant le GSD2 de deux VAs à comparer (axe des abscisses) ainsi que la SP désirée (courbes de niveau), le ratio des moyennes nécessaire pour

obtenir la SP désirée est lu sur l'axe des ordonnées. La Figure 17 donne un aperçu de l'allure des distributions lognormales de  $GSD2 = 3$  pour différentes valeurs de SP.

#### **4.2.4 Application 2 : comparaison de scores d'impacts et sondage d'analystes**

L'application la plus évidente de la SP est la quantification de la certitude d'une conclusion dans une ACV comparative, au niveau des scores d'impacts. Traditionnellement, des niveaux de SP de 0,9, 0,95, 0,975 ou 0,99 sont appliqués dans des domaines comme le contrôle de la qualité des procédés et en épidémiologie dans la comparaison de l'efficacité d'interventions médicales. L'objectif du sondage était de quantifier la variation du niveau de confort des praticiens face à différentes valeurs de SP. La tolérance au risque est toujours fortement associée au contexte, et il est possible que les seuils de SP acceptables en contrôle de qualité ou en épidémiologie ne soient pas adéquats en ACV.

Sept analystes ont été sondés au sujet de 40 graphiques (voir Figure 18 pour un exemple) : sur une échelle de 1 à 5 (« pas du tout confortable » à « très confortable »), quel serait votre niveau de confort face à la conclusion que le scénario 1 présente un score d'impacts significativement plus faible que le scénario 2? De plus amples informations ont aussi été fournies. La moitié des graphiques provenaient d'une ACV fictive commandée par une chaîne de cafés franchisée désirant baser le choix du matériau des gobelets sur des critères environnementaux (par exemple : carton VS styromousse). L'autre moitié des ACV seraient utilisées pour guider le gouvernement fédéral dans l'élaboration d'une politique de d'approvisionnement en énergie pour les dix prochaines années. Le paramètre de contexte servait à contrôler la sensibilité du niveau de confort face à des conséquences plus ou moins importantes des décisions appuyées par l'étude.

Aucune catégorie d'impacts n'était spécifiée, mais les analystes ont été informés que toutes les sources d'incertitudes, y compris celles sur les facteurs de caractérisation, ont été comptabilisées dans le Monte Carlo, et que la corrélation entre les systèmes avait aussi été considérée dans l'échantillonnage.

Des niveaux de PS de 0,75, 0,85, 0,9, 0,95 et 0,975 ont été choisis, et le  $GSD2$  des deux lognormales étaient de 1,5, 3,5, 5,5 et 7,5. Chaque Monte Carlo comportait 10 000 itérations. Chaque combinaison SP/ $GSD2$  a été générée pour un total de 20 graphiques par contexte fictif d'ACV. Les graphiques ont été mélangés une fois et présentés dans ce même ordre aléatoire pour

tous les analystes. Les analystes ont pu examiner les graphiques un à un, sans limite de temps, mais sans possibilité de revenir en arrière pour consulter les graphiques une seconde fois ou revoir les réponses précédentes. Les réponses étaient compilées par l'interviewer en rencontre individuelle avec chaque analyste.

Les analystes n'ont pas reçu d'explications au sujet du concept de la PS ni sur les objectifs de l'étude, et les valeurs de SP n'étaient pas affichées sur les graphiques. Une fois les 40 graphiques examinés, les analystes ont été sondés avec une série de questions ouvertes : quels signes visuels étaient recherchés pour évaluer leur niveau de confort? Pourquoi certains graphiques produisaient un plus faible niveau de confort? Quelles hypothèses ou expériences passées gardaient-ils à l'esprit en analysant les graphiques? D'autres questions ont été improvisées en fonction des réponses.

Sept analystes ont été interrogés, avec une expérience variant de deux à neuf ans. Six d'entre eux travaillent au CIRAIG, l'autre, à Quantis Montréal. Les Figure 19, Figure 20 et Figure 21 montrent respectivement les résultats généraux, l'influence du GSD2 et du contexte sur le niveau de confort. Le consensus semble assez fort pour  $SP = 0.975$  (majorité de 4 et de 5), et pour  $PS \leq 0,85$  (majorité de 1, 2 et 3). La situation est plus ambiguë pour  $SP = 0.9$  et  $0.95$ . Les analystes sont relativement peu affectés par le changement de GSD2 : pour des SP identiques, le niveau de confort diminue légèrement, mais cette tendance n'est probablement pas significative étant donné la taille de l'échantillon. Finalement, la plupart du temps, pour le même analyste et des graphiques identiques en tout point, le niveau de confort a été majoritairement identique ou légèrement inférieur pour l'ACV au sujet de la politique énergétique ( $\Delta = 0$  ou  $-1$ ).

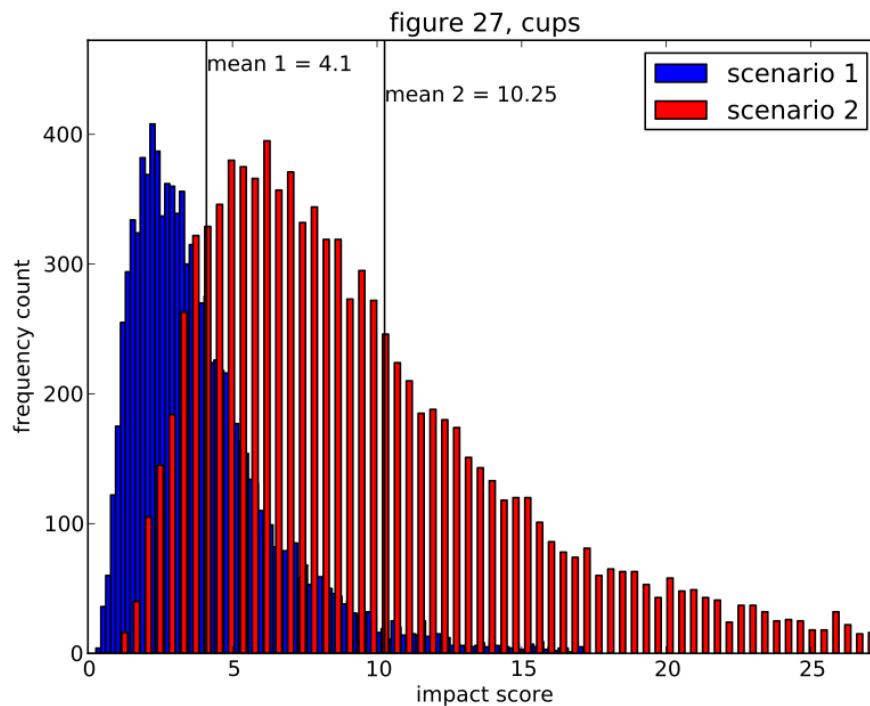


Figure 18 : exemple de figure présentée aux analystes;  $SP = 0,85$  et  $GSD2 = 3,5$

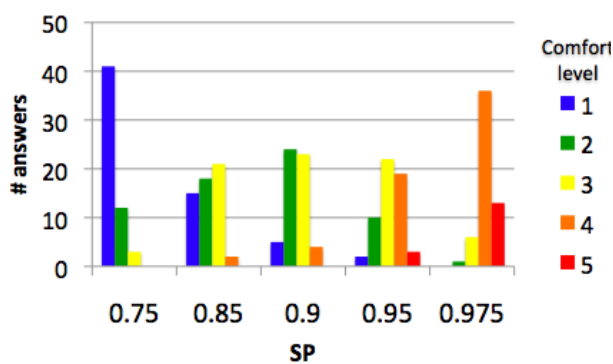


Figure 19 : résultats généraux du sondage

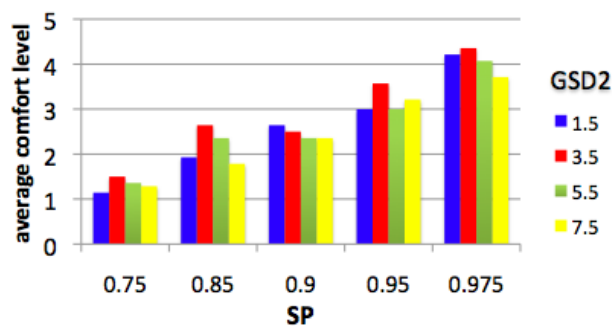


Figure 20 : Confort moyen en fonction de la SP et du GSD2

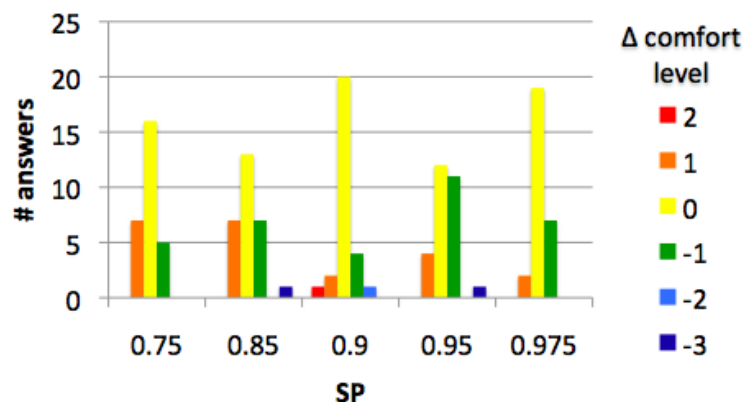


Figure 21 : Changement du niveau de confort entre le contexte « gobelet » et « politique énergétique ». Un changement négatif indique un niveau de confort plus faible pour la politique énergétique.

Du côté qualitatif, bien que des techniques d'entrevues rigoureuses et standardisées n'ont pas été appliquées, la saturation semble avoir été atteinte (Bowen 2008): tous les éléments importants ont été énoncés lors des premières entrevues et réitérés par les dernières. Toutefois, il faut noter que l'enquête souffre probablement d'un « biais organisationnel » : tous les sujets ont soit complété un diplôme de cycle supérieur au CIRAIG, soit fait leur apprentissage de l'ACV au sein du groupe. L'enquête devrait être répétée avec des analystes en provenance de d'autres organisations. Étant donné le manque de lignes directrices précises fournies par les normes ISO (Draucker, Kaufman et al. 2011), des résultats différents ne seraient pas surprenants.

Les analystes perçoivent instinctivement les plus larges valeurs de SP comme plus dignes de confiance. L'écart type des réponses est toutefois plus faible pour les valeurs extrêmes de SP du sondage, indiquant une hétérogénéité des méthodes utilisées par les analystes pour dissiper l'ambiguïté dans les situations intermédiaires, ce qu'ont aussi révélé les questions ouvertes.

La technique la plus souvent citée est l'aire de chevauchement des histogrammes. Cette technique n'est pas fiable, car cette aire dépend de la largeur des intervalles choisis pour les barres. La technique est toutefois fiable pour les variables continues, car les intervalles sont de largeur infinitésimale. D'autres techniques ont été citées, la plupart floues et sujettes à des variables n'ayant pas d'influence sur la SP. Plusieurs analystes ont avoué que, malgré leurs années d'expérience, ils se sentaient mal outillés pour justifier leur niveau de confort.

D'un côté, les experts en ACV effectuent des évaluations floues, mettant en perspective plusieurs informations comme la qualité des données, pour formuler des recommandations nuancées et discuter des limites de l'étude. Ils ont effectué l'étude eux-mêmes, ont interviewé des spécialistes, collecté les données, généré les résultats, analysé les scores sous plusieurs angles, et ont une bonne compréhension des sources de biais et d'incertitudes, même s'ils ne peuvent pas toujours traduire ces éléments en une évaluation quantitative. De l'autre côté, la SP est un simple nombre, calculé à l'aide d'une formule. La SP est un outil intéressant pour son objectivité; il devrait être inclus dans la routine d'analyse des incertitudes, mais ne devrait jamais remplacer le jugement d'un expert.

Une ligne directrice définitive n'est peut-être pas un objectif désirable dans ce contexte. Les petites études plus rapides et à faible budget devraient accepter des SP plus faibles. Certains décideurs sont plus ou moins tolérants au risque, et leur tolérance pourrait varier en fonction des conséquences de leurs décisions.

#### **4.2.5 Application 3 : création d'archétypes de facteurs de caractérisation**

Les valeurs déterministes des FCs s'étendent sur une large plage de valeurs à l'intérieur d'un pays. Si l'on tient compte de leur incertitude, peut-être qu'une différence d'un facteur 10 n'est pas statistiquement significative. La SP a été appliquée sur les FC du modèle d'impact direct sur la santé humaine de la consommation d'eau décrit au résumé de l'article 1.

Le premier graphique de la Figure 22 montre le résultat du Monte Carlo (5<sup>ème</sup> et 75<sup>ème</sup> percentiles, quartiles, médiane, moyenne et données aberrantes) pour le FC de tous les bassins versants indiens, pour l'eau de qualité S3. Les cellules, désignées par leur numéro, ont été classées par moyenne croissante de FC. Un test de SP a été effectué pour chaque paire de cellules. Le résultat est présenté sur le second graphique. Les marqueurs en étoile représentent les valeurs de SP pour deux cellules adjacentes dans l'ordre croissant de moyenne. On peut voir que deux SP de cellules adjacentes sont supérieures au seuil arbitraire de 0,95, soit la comparaison 555-375 et 518-515. Les quatorze bassins versants pourraient donc être classés en trois archétypes : faible, moyen et haut impact, à des moyennes de  $\sim 5 \times 10^{-10}$ ,  $\sim 1 \times 10^{-4}$  and  $\sim 5 \times 10^{-3}$  DALY/m<sup>3</sup>, respectivement. Bien que différents, les différents nCFs groupés dans chaque archétype ne sont pas jugés significativement distincts les uns des autres.

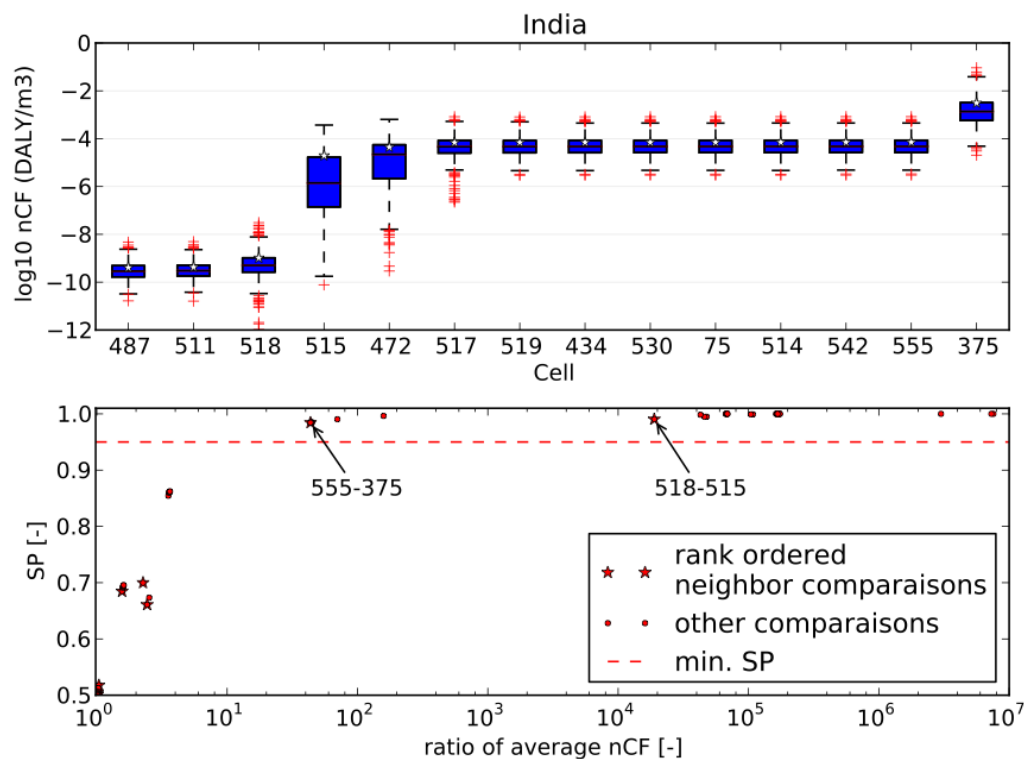


Figure 22 : Résultats du Monte Carlo et valeurs de SP pour l'eau de qualité S3 en Inde

Le même calcul a été effectué pour tous les pays où les FC sont non nuls (pays à moyen ou faible GNI par personne et problème de rareté dans au moins un bassin versant), et résumé par quatre indicateurs présentés à la Figure 23. Le premier graphique montre le plus petit ratio de moyennes nécessaire à l'intérieur d'un pays pour que la valeur de SP soit supérieure au seuil, ainsi que le plus large ratio de moyennes menant à une valeur de SP inférieure au seuil. Chaque marqueur représente un pays, classé sur la valeur du premier ratio décrit ci-dessus. Le second graphique montre le nombre de bassins versants dans chaque pays et le nombre d'archétypes justifiables par le seuil de SP choisi. On peut constater que peu importe le nombre de bassins versants, le nombre d'archétypes est le plus souvent deux, mais ne dépasse jamais trois. Le classement des pays est le même pour les deux graphiques.



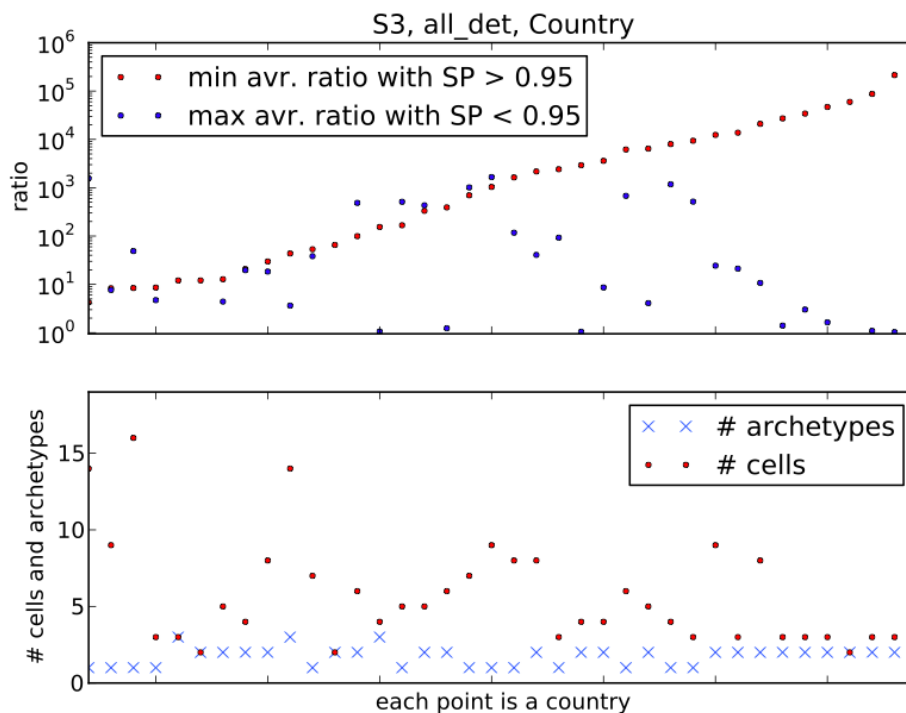


Figure 23 : survol international de la catégorie d'impacts pour la qualité d'eau S3

#### 4.2.6 Avantages et limitations de la SP

Dans l'application aux scores d'impacts, la SP fournit un résultat calculé de manière systématique, facile à communiquer et à interpréter, que l'analyste peut inclure dans une discussion. Dans une analyse de scénario, si un pire cas et un meilleur cas réalistes font passer la SP d'un côté à l'autre du seuil, le ou les paramètres impliqués doivent être identifiés comme sensibles, l'analyste devrait s'assurer de la fiabilité de ces données ou tenter de diminuer leur incertitude par une cueillette de données supplémentaires.

Du côté du développement des modèles, la SP peut être utilisée pour valider la pertinence de la démultiplication des options. Les modèles doivent viser à être « aussi complexes que nécessaire, aussi simples que possible ». La parcimonie est donc de mise. Si les FCs spatialisés ou les archétypes de flux élémentaires ne sont pas statistiquement différenciables, le fardeau nécessaire à leur utilisation n'est pas justifié. L'objectif d'augmenter le pouvoir discriminant de l'ACV devrait être atteint en minimisant la complexification.

Malheureusement, l'utilisation de la PS est entravée par le manque de moyens, en ACV, pour prendre en compte la corrélation, tant intra-système que inter-systèmes. Ce problème n'a même pas encore été adéquatement défini (Bourgault, Lesage et al. 2012). On sait déjà que de n'assumer aucune corrélation est l'option la plus conservatrice, mais qui surestime le chevauchement entre les scores d'impacts. Les valeurs de SP pourraient être artificiellement bas. Prendre en compte la corrélation est une tâche requérant beaucoup d'informations, et ne devrait être entreprise que dans les cas où la différence entre 0% et 100% de corrélation affecte les conclusions d'une étude, ce qui peut être testé à l'aide de la PS.

### **4.3 Article 3 : «Relevance and feasibility of sector-specific country-level characterization factors in LCA»**

#### **4.3.1 Aperçu de l'article**

Statut de l'article : soumission initiale à la revue Environmental Science & Technology le 20 novembre 2013.

Objectif : Démontrer qu'une agrégation basée sur un proxy de population, un proxy d'émission générique et un proxy d'émission par industrie mène à des FC agrégés très différents.

Originalité : Utilisation de données provenant d'organismes gouvernementaux de surveillance de la pollution (TRI, PRTR, NPRI) pour valider l'accord entre les proxy d'émission et de population. Développement d'outils graphiques pour l'analyse des résultats, indépendant de la catégorie d'impact, de la délimitation et du type de proxy.

#### **4.3.2 Introduction : le problème des jeux de pondération**

À l'article 1, le besoin d'une agrégation des FC natifs d'une fine échelle vers une échelle plus grossière a été démontré. La procédure demande d'attribuer des poids à chaque FC, représentant la probabilité d'une émission sur le domaine de chaque FC natif. Cette probabilité, jusqu'à présent, a été approximée à l'aide de la population vivant sur chaque cellule. Cela suppose une bonne corrélation entre la quantité de pollution émise et la population. Cette supposition se trouvera vérifiée pour certains secteurs, mais pas pour d'autres : l'extraction des ressources

naturelles (mines, foresterie, agriculture et plateforme pétrolières) ont souvent lieu dans des régions peu peuplées.

Cet article a comme objectif de valider la corrélation entre les proxys d'émissions et la population et de démontrer l'impact de cette différence sur les FC agrégés. À l'aide de bases de données d'émissions telles que le E-PRTR (European Pollution and Transfer Registry), NEI (National Emission Inventory aux États-Unis) et NPRI (National Pollutant Release Inventory au Canada), les émissions peuvent être réparties par secteur économique. La pertinence de créer des jeux de pondération et des FCs spécifiques à chaque secteur est démontrée. Finalement, étant donné la complexification engendrée par des FCs par secteur (demande d'informations, augmentation de la taille du système à résoudre), la faisabilité d'une telle entreprise est validée.

L'analyse porte sur deux catégories d'impacts régionaux pour lesquelles les FCs natifs étaient disponibles à fine échelle, soit l'acidification terrestre et l'eutrophisation de l'eau douce, tirés de la méthode IMPACT World+.

### **4.3.3 Corrélation entre émissions et population**

Les bases de données d'émissions mentionnées listent les coordonnées de latitude et longitude ainsi que la quantité de polluants émise par chaque usine. Cela permet d'associer chaque émission à une cellule de la grille du modèle d'ÉICV. La quantité émise dans chaque cellule est ensuite divisée par la quantité totale émise dans le pays pour créer le jeu de pondération par pays. L'opération est répétée avec la population de chaque cellule pour créer un autre jeu de pondération.

Un graphique de Bland-Altman (Figure 24), développé initialement dans le contexte de la médecine statistique, permet de visualiser les points où les jeux de pondération sont en accord. Sur l'axe des abscisses, la position d'un point représente la moyenne des deux valeurs à comparer pour une cellule. Sur l'axe des ordonnées, la position d'un point représente la différence entre les deux valeurs. La ligne rouge représente l'emplacement que peut prendre un point si la cellule qu'il représente contient de la population, mais aucune émission. La ligne verte représente la situation inverse.

Pour les émissions de phosphate en Angleterre, plusieurs points se trouvent près de ces droites (Figure 24). Il est toutefois difficile de juger de l'effet de ces différences sur le résultat d'intérêt :

le FC agrégé. Les FCs natifs couvrent potentiellement plusieurs ordres de grandeur. Une différence de pondération pour une cellule à très faible FC natif n'aura pas le même effet qu'au sein d'une cellule à très large FC natif. La représentation originale de Bland-Altman doit être complétée par d'autres outils.

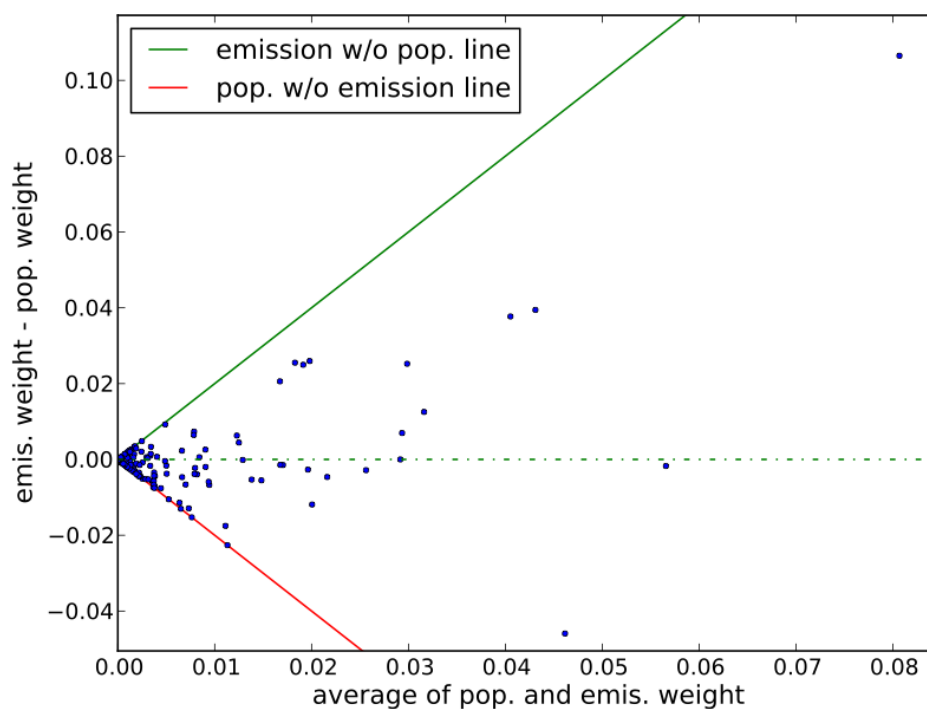


Figure 24 : Représentation de Bland-Altman pour les émission de phosphate en Angleterre

À la Figure 25, les graphiques partagent le même axe des abscisses, où la position des points représente la valeur des FCs natifs. Le premier graphique représente les poids des deux jeux de pondération, et le second, la différence entre ces poids. Le troisième graphique montre la contribution en valeur absolue de chaque cellule au FC agrégé (points), ainsi que le cumulatif de contribution de chaque cellule (lignes). C'est ce graphique qui permet de visualiser l'impact de l'accord entre les deux poids. Dans le cas ci-dessous, on voit que les cellules aux FCs natifs les plus élevés sont aussi ceux pour lesquels l'accord est le meilleur. Les plus larges différences sont présentes dans des cellules aux FCs natifs inférieurs de plus d'un ordre de grandeur aux plus larges du pays. Le résultat est un FC agrégé comparable pour les deux jeux de poids.

La Figure 26 montre la situation de l'Autriche, où la cellule la plus sensible aux émissions de phosphate se trouve au pied des Alpes, à l'Est du lac de Constance. Cette cellule est le domicile

de 3,7% de la population, mais aucune émission de phosphate n'y est rapportée. Cet unique désaccord de poids fait passer la différence des FC agrégé de 10% à près de 100%.

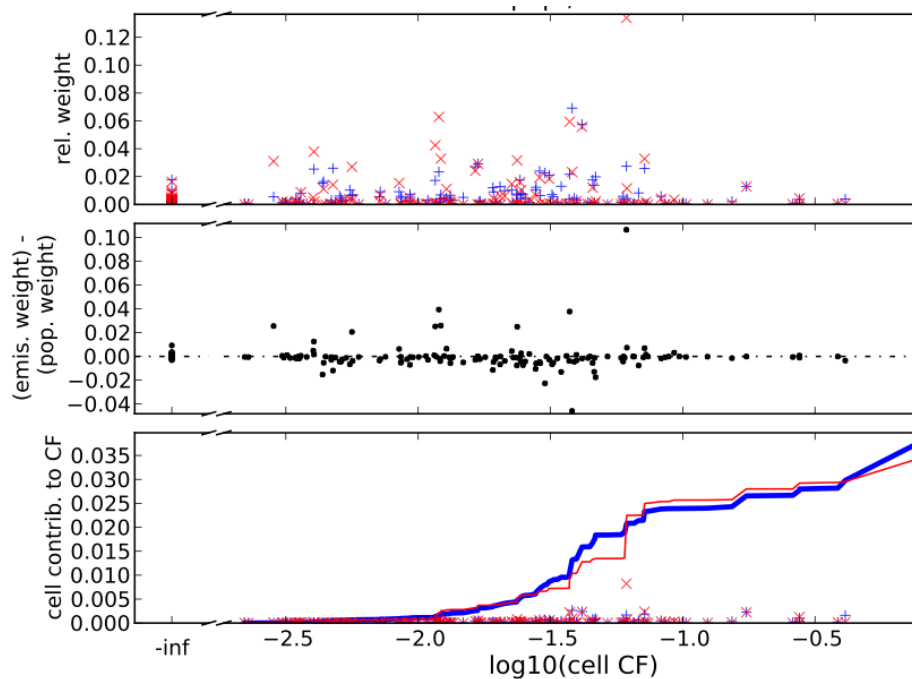


Figure 25 : Poids, différences de poids et contribution au FC agrégé pour les émissions de phosphate en Angleterre

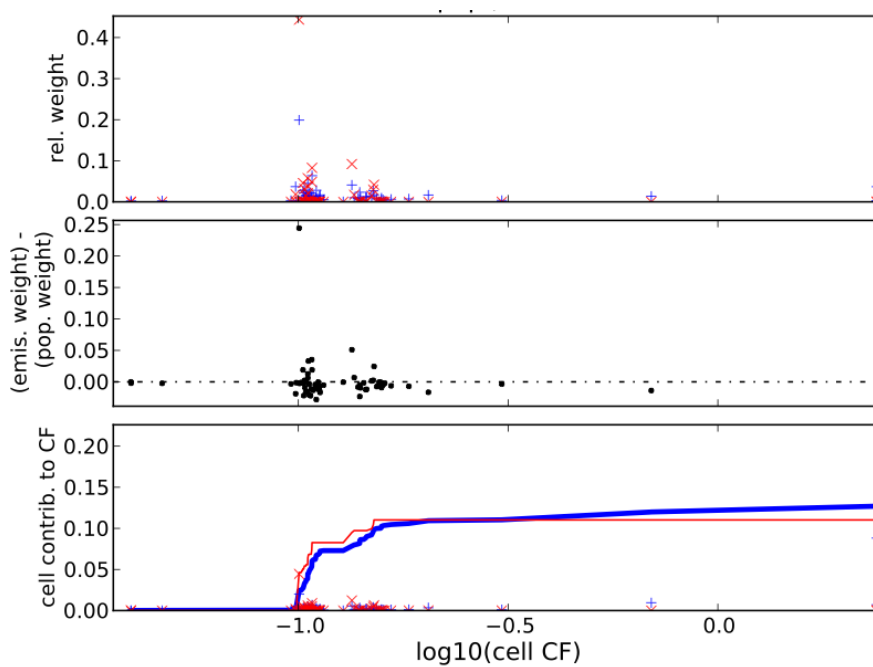


Figure 26 : Poids, différences de poids et contribution au FC agrégé pour les émissions de phosphate en Autriche

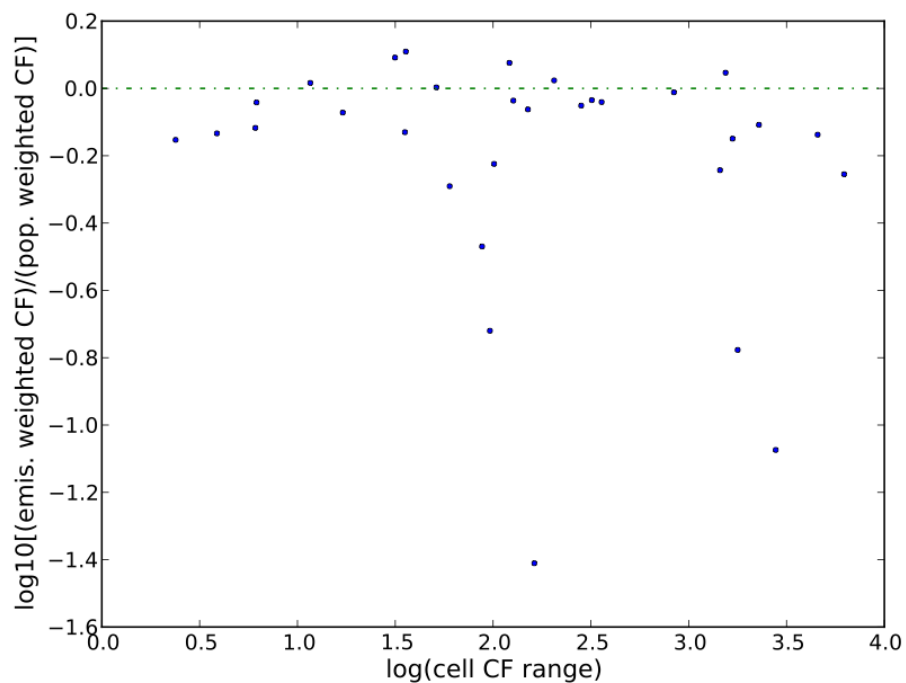


Figure 27: ratio des FC agrégés par poids d'émission et de population, en fonction de la gamme de FC, pour le phosphate (chaque point est un pays)

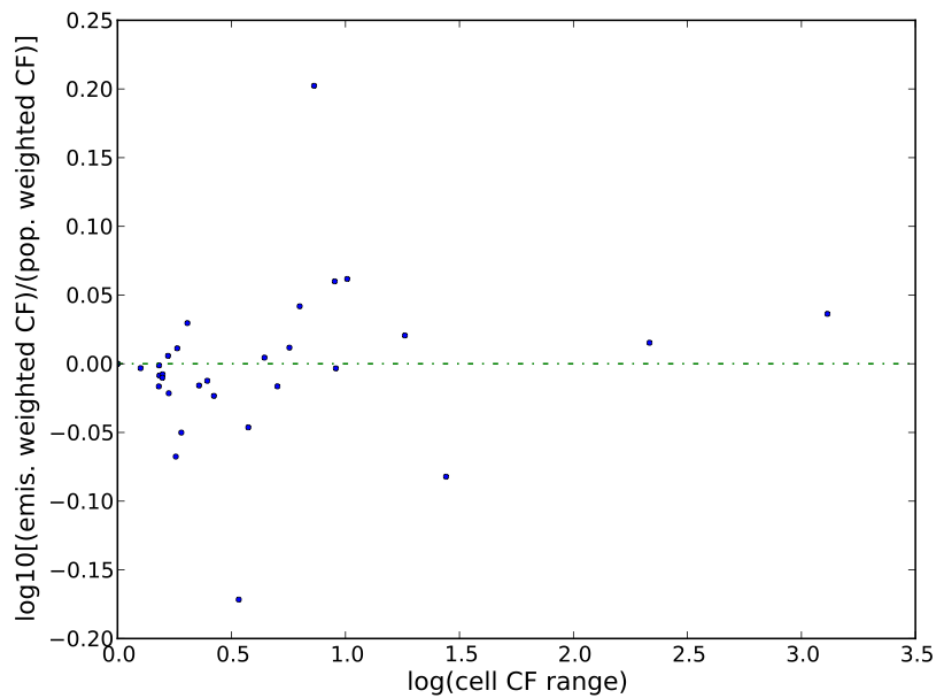


Figure 28: ratio des FC agrégés par poids d'émission et de population, en fonction de la gamme de FC, pour le  $\text{NH}_3$  (chaque point est un pays)

Les Figure 27 et Figure 28 montrent l'ensemble des résultats pour tous les pays inclus dans l'analyse. On voit que pour le phosphate, plus un pays est couvert par une large gamme de FC natifs, plus la variation entre le FC agrégé par pondération d'émission ou de population peut être importante (jusqu'à 1,5 ordre de grandeur). Cette relation ne semble pas aussi évidente pour la catégorie d'acidification terrestre, où ces différences sont plus petites.

Cette analyse fait ressortir qu'il n'y a pas de conclusion générale à tirer sur l'effet du choix du jeu de pondération : dépendamment des substances, catégories d'impacts et des pays, l'effet peut être négligeable, ou atteindre plus d'un ordre de grandeur.

#### 4.3.4 Les jeux de pondération et FCs par industrie

La Figure 29 conserve la structure inspirée par Bland-Altman, à laquelle elle superpose les FCs obtenus par des jeux de pondération par numéro ISIC (International Standard Industry Classification). Cette figure fait ressortir deux points intéressants.

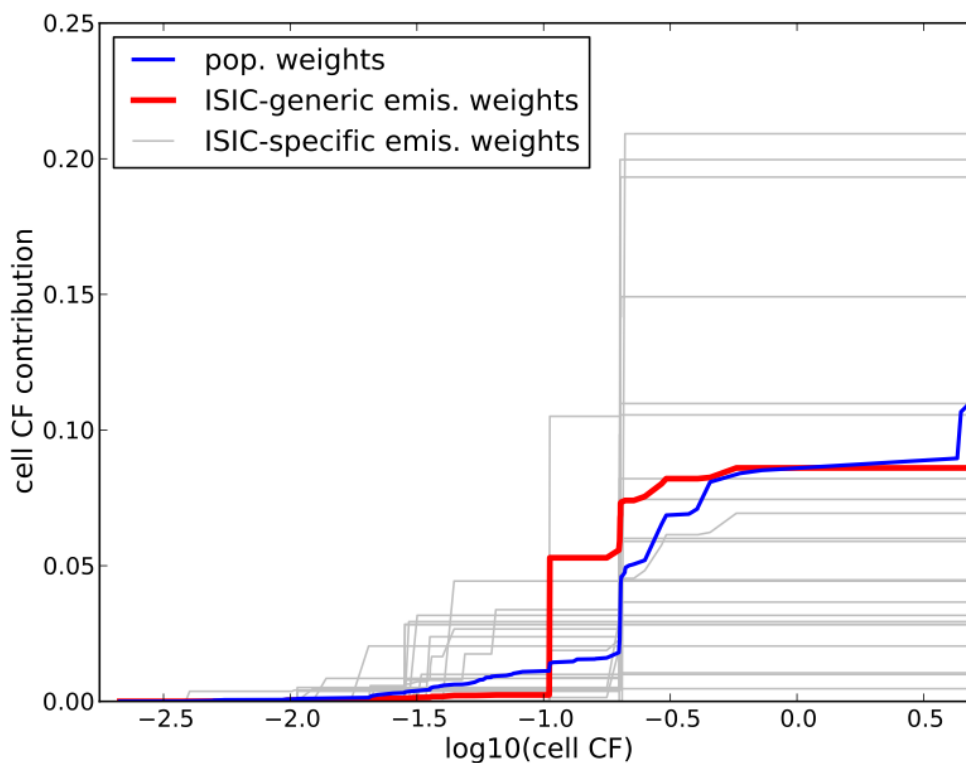


Figure 29 : FCs pondérés par population, émissions de phosphate génériques et par industrie en Italie, eutrophisation. La valeur finale des FCs par industrie dépend d'un plus petit nombre de contributeurs.

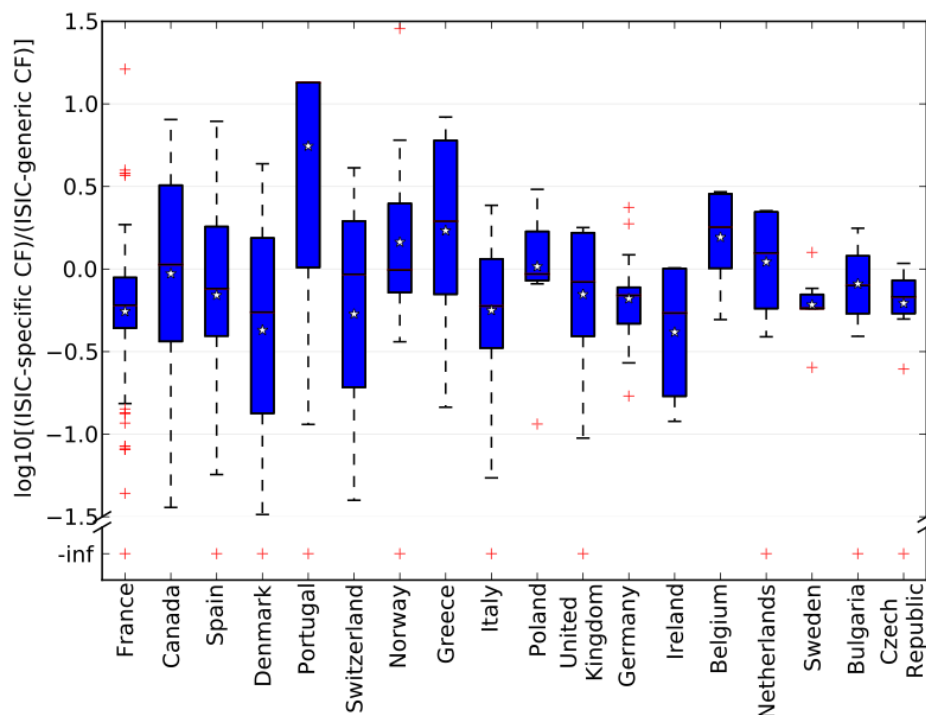


Figure 30 : Gamme de FCs par industrie divisé par le FC pondéré par émissions de phosphate, eutrophisation

Premièrement, les FCs peuvent varier considérablement d'une industrie à l'autre. Ces variations s'étendent sur deux ordres de grandeur, tel que démontré à la Figure 30. Deuxièmement, les courbes cumulatives par industrie atteignent leur maximum en quelques sauts, signe que les usines d'un secteur sont concentrées dans quelques cellules et complètement absentes des autres cellules. Si cette concentration a lieu dans les cellules aux FCs natifs les plus élevés, le FC agrégé par secteur sera beaucoup plus élevé que le FC agrégé par le jeu de pondération générique.

### 4.3.5 Faisabilité

La création de FCs par industrie implique de faire passer l'information sur l'origine des émissions du vecteur de mise à l'échelle vers l'inventaire, ainsi qu'une augmentation de la taille du système à résoudre. Le premier problème a déjà été partiellement résolu par la diagonalisation



du vecteur de mise à l'échelle (Mutel et Hellweg 2009). L'effet de l'augmentation de la taille du système sur le temps de calcul a été testé et s'avère ne pas augmenter de manière à entraver l'opérationnalisation. Les deux sections suivantes traitent de ces deux problèmes.

#### 4.3.5.1 Formalisme mathématique

Le Tableau 5 résume la séquence de calcul et la structure des matrices employées. Quelques nouveautés ont été introduites pour opérationnaliser les FCs par industrie. Premièrement, un masque (une matrice) doit être construit afin de lier les processus élémentaires aux secteurs économiques. Il permet l'agrégation de l'inventaire par secteur économique par simple produit matriciel, et mène à une matrice d'inventaire à deux dimensions (flux élémentaires x secteurs) plutôt qu'à un vecteur d'inventaire à une dimension (flux élémentaires x 1). Les FCs doivent donc se présenter dans une matrice de même dimension que l'inventaire et être multipliés élément-par-élément plutôt qu'à l'aide de la multiplication matricielle. Une matrice est créée pour chaque catégorie d'impacts.

Tableau 5: matrices de la structure computationnelle adaptée

Nom	Symbole	Contenu	Dimension	Formule
Matrice technologique	<b>A</b>	Entrées de la technosphère par processus élémentaire	$n \times n$	Fournie par la base de donnée ICV
Matrice intervention	<b>B</b>	Flux élémentaires par processus élémentaires	$m \times n$	Fournie par la base de donnée ICV
Vecteur de demande finale	<b>f</b>	Flux de références du système	$n \times 1$	Construit par l'analyste
Matrice de mise à l'échelle	<b>S</b>	Quantité totale de chaque processus élémentaire	$n \times n$	$S = \text{diag}(A^{-1}f)$
Matrice d'inventaire par processus	<b>G</b>	Totaux des flux élémentaires émis et consommés	$m \times n$	$G = BS$
Masque processus-secteur	<b>U</b>	Si PÉ $i$ fait parti du secteur $j$ : $U_{ij} = 1$ ; sinon, zéro	$n \times p$	Fournie par la base de donnée ICV
Matrice d'inventaire par secteur	<b>G_ISIC</b>	Totaux des flux élémentaires émis et consommés	$m \times p$	$G\_ISIC = GU$
Matrice de FC par secteur	<b>Q</b>	Facteurs de caractérisations	$m \times p$	Fourni par la méthode d'ÉICV
Matrice d'impact	<b>H</b>	Contribution de chaque secteur et flux élémentaire au score	$m \times p$	$H =$ multiplication élément par élément de $G\_ISIC$ et $Q$

Le résultat est une matrice à deux dimensions (flux élémentaires x secteurs) que l'on peut sommer par ligne et par colonne pour obtenir la somme des impacts. La multiplication par la matrice de FCs doit être répétée pour chaque catégorie d'impacts.

#### **4.3.5.2 Test de temps de calcul**

Il n'est pas nécessaire d'inverser la matrice technologique pour obtenir le vecteur de mise à l'échelle. Cette opération est coûteuse en temps de calcul et peut être évitée par un solveur de système d'équations linéaires et des matrices creuses (« sparse matrices »). Des routines optimisées sont disponibles dans la plupart des langages de programmation.

Afin de tester le temps de calcul, un système fictif a été construit pour imiter le pire cas imaginé par Heijungs (Heijungs 2012) : 50 000 processus élémentaires et 60 000 000 de flux élémentaires. Le détail des hypothèses posées pour la construction des matrices est disponible à la section 2 du matériel complémentaire de l'article. Le temps de calcul ne dépend pas de la valeur d'un coefficient, mais de sa présence ou absence. L'obtention du vecteur de mise à l'échelle s'effectue en environ une minute et la multiplication de la matrice de mise à l'échelle et de la matrice intervention demande de deux à trois secondes. L'augmentation de la taille du système ne semble donc pas poser de problème au niveau des temps de calcul.

### **4.3.6 Discussion**

#### **4.3.6.1 Avantages de l'agrégation par industrie**

Puisque les FCs agrégés par industries comportent un plus petit nombre de contributeurs, il est attendu que l'opération d'agrégation créera une plus faible augmentation d'incertitude que pour les FCs agrégés par émission. Toutefois, il n'a pas été possible de tester cet effet car les FCs pour l'eutrophisation et l'acidification n'ont pas encore d'estimés fiables de leur incertitude.

L'agrégation par industrie permet de se débarrasser de l'hypothèse d'accord entre la population et le taux d'émission. Les FCs agrégés par industrie sont plus représentatifs, plus adéquats, que s'ils étaient basés sur une pondération par émission. Cette affirmation est particulièrement pertinente pour les secteurs d'industries primaires, qui se situent plus souvent dans des régions moins peuplées.

#### **4.3.6.2 Limites de l'analyse**

Les données utilisées sont de bonne qualité mais ne sont pas nécessairement complètes. Par exemple, les bases de données n'exigent pas de rapport d'émissions pour les petites entreprises et de faibles quantités émises. L'ACV nous enseigne qu'un grand nombre de contributeurs à un score d'impacts peuvent, mis ensemble, représenter une plus grande contribution que des usines à l'avant-plan. Ainsi, il est difficile d'évaluer la taille et la direction du biais causé par l'absence des données.

Les bases de données rapportent les totaux d'émissions diffuses et de pertes de nutriment, sans les désagréger par secteur économique. Ils peuvent être utilisés pour générer des FCs agrégés par substance, mais pas par secteur, ce qui limitera la précision que l'on pourrait atteindre en agriculture (types de récoltes).

E-PRTR contenait beaucoup d'erreurs au niveau des coordonnées des usines. Près de 1% des usines étaient situées hors de leur pays d'origine. En l'absence d'informations pour effectuer les corrections, ces usines ont été éliminées de l'analyse.

Malgré ces limites, l'analyse montre que la variabilité inter-secteurs des FCs est significative. Les méthodes d'analyse et de visualisation proposées ne dépendent pas des sources de données et pourront être réutilisées si de nouvelles sources ont besoin d'être évaluées.

#### **4.3.6.3 Sources complémentaires de proxy**

Certaines bases de données se spécialisent sur les activités de certaines industries. Par exemple, (Mutel, Pfister et al. 2012) ont utilisé eGRID pour obtenir de l'information sur la localisation, les performances et les émissions des centrales électriques américaines. CropScale, produit par le National Agricultural Statistics Service aux États-Unis, publie des statistiques géolocalisées de haute résolution au sujet des différentes récoltes. En croisant cette source d'informations avec des moyennes d'épandage de fertilisants et de pesticides, on pourrait obtenir des proxys d'émissions désagrégés par type de culture.

La localisation des usines et la nature de leur production sont des informations faciles à obtenir auprès des associations de producteurs d'un secteur. Même si ces usines ne rapportent pas d'émissions, elles peuvent être estimées à partir de d'autres usines générant les mêmes produits.

## 4.4 Article 4 : « Systematic disaggregation : a hybrid LCI computation algorithm enhancing interpretation phase in LCA »

### 4.4.1 Aperçu de l'article

Statut de l'article : Publié le 24 avril 2012 dans la revue International Journal of LCA.

Objectif : Développement d'un algorithme de calcul des impacts du cycle de vie permettant d'accéder à tous les niveaux de détails nécessaires pour l'interprétation et l'application de FC spatialisés.

Originalité : Alliance de deux approches de calcul : séquentielle et matricielle, permettant un calcul rapide et une structure de résultat adaptée aux besoins des analystes. s

### 4.4.2 L'approche matricielle : intérêts et limites

L'opération la plus déroutante dans le calcul de l'inventaire du cycle de vie est l'inversion de la matrice technologique (**A**) :

$$f = As \Leftrightarrow s = A^{-1}f \quad \text{Équation 6}$$

Le contenu de la matrice technologique est facile à interpréter : chaque coefficient  $a_{ij}$  représente combien d'unités produites par le processus élémentaire  $i$  sont requises *directement* pour la production d'une unité du processus élémentaire  $j$ . Le contenu de la matrice inverse est lui aussi facile à interpréter : chaque coefficient  $a^{-1}_{ij}$  représente combien d'unités produites par le processus élémentaire  $i$  sont requises *directement et indirectement* pour la production d'une unité du processus élémentaire  $j$ . En multipliant  $A^{-1}$  par le vecteur de demande finale  $f$ , on obtient le vecteur de mise à l'échelle, représentant combien d'unités de chaque processus élémentaire l'économie devra produire pour soutenir la création de l'unité fonctionnelle.

Tableau 6 : Variables et définitions dans la représentation matricielle en ACV, tiré de (Bourgault, Lesage et al. 2012)

Name	Symbol	Content	Dimensions	Formula
Technology matrix	<b>A</b>	Consumption of intermediary products by other unit processes	$n \times n$	Normally supplied by the database
Intervention matrix	<b>B</b>	Emissions and natural resource consumption of unit processes	$m \times n$	Normally supplied by the database
Demand vector	$f$	Reference flow for the system	$n \times 1$	Set up by LCA practitioner
Scaling vector	$s$	Total quantity of intermediary products	$n \times 1$	$f = As \Leftrightarrow s = A^{-1}f$
Inventory	$g$	Total emissions and natural resource consumption	$m \times 1$	$g = Bs$
Characterization matrix	<b>Q</b>	Characterization factors	$k \times m$	Normally supplied by impact method
Impact vector	$h$	Total score for each impact category	$k \times 1$	$h = Qg$

Pour des matrices de petite taille, il est possible d'obtenir explicitement la formule reliant les  $a^{-1}_{ij}$  aux  $a_{ij}$ . Malheureusement, la longueur des formules croît hors de toute proportion pratique pour un modèle représentant une chaîne d'approvisionnement d'une taille réaliste. Une routine à l'intérieur d'un logiciel pourra calculer les coefficients, mais la formule explicite restera inaccessible.

Cette situation limite l'une des principales tâches de la phase d'interprétation : révéler comment se répartissent les impacts le long de la chaîne d'approvisionnement. Dans la structure matricielle développée par Heijungs (Heijungs et Suh 2002), tous les impacts sont agrégés en un seul chiffre (dimension 0). La diagonalisation du vecteur de mise à l'échelle proposée par (Mutel et Hellweg 2009) désagrège l'impact pour chaque processus élémentaire en un vecteur (dimension 1). La même diagonalisation pourrait être appliquée au niveau du vecteur d'inventaire pour produire un score d'impacts désagrégé pour chaque flux élémentaire. On pourrait aussi appliquer cette diagonalisation aux deux vecteurs simultanément pour produire un tableau d'impacts de dimension 2.

Ces opérations sont utiles, mais ne règlent pas le problème fondamental de l'inversion de la matrice : l'impossibilité d'accéder à la séquence d'offre et demande liant une occurrence d'un processus élémentaire à l'unité fonctionnelle (« structural path assessment »). Pour l'analyse des incertitudes, l'inversion de matrice ne laisse d'autre choix qu'une méthode d'approximation de premier ordre (Taylor, voir section 2.2.5.2), ou les méthodes stochastiques (Monte Carlo, voir section 2.2.5.1).

### 4.4.3 L'approche séquentielle : intérêts et limites

La méthode séquentielle consiste à suivre une à une les branches de la chaîne d'approvisionnement pour quantifier toutes les occurrences de chaque processus élémentaire. Dans ce cas, il est possible d'exprimer les impacts en fonction des coefficients des matrices technologique et d'intervention, ainsi que des facteurs de caractérisations :

$$impact = \sum_{i=1}^{\infty} f_j \cdot (a_{k,j} \cdot \dots \cdot a_{m,n}) \cdot b_{p,n} \cdot FC_p \quad \text{Équation 7}$$

où chaque terme  $i$  représente l'impact d'une substance émise par une occurrence d'un processus élémentaire. Chaque terme est le produit d'un facteur de caractérisation pour la substance considérée ( $FC_p$ ), un facteur d'émission de la substance  $p$  par le processus élémentaire  $n$  ( $b_{p,n}$ ), un flux de référence  $f_j$  contenu dans l'unité fonctionnelle, et le produit des coefficients de la matrice technologique reliant le flux de référence au processus élémentaire émettant le flux élémentaire.

La structure de cette formule épouse fidèlement la forme de ce qu'elle prétend modéliser : une somme d'impacts, répartis le long d'une chaîne d'approvisionnement; elle est donc facile à interpréter. Malheureusement, il s'agit d'une somme infinie. Les contributions individuelles de la majorité des termes sont négligeables, mais lorsque rassemblées, elles représentent une fraction considérable (Wiedmann, Lenzen et al. 2007). Aucune règle générale ne permet de localiser rapidement les contributions individuelles significatives.

### 4.4.4 Combinaison des deux approches : la désagrégation systématique

L'algorithme de désagrégation systématique permet de conserver les avantages de la méthode séquentielle, tout en calculant une quantité limitée de termes. Le calcul procède comme suit :

- Un flux de référence est sélectionné;
- L'inversion de matrice est utilisée pour calculer l'impact total de la section sous-jacente de la chaîne d'approvisionnement par rapport à l'impact de tout le système;
- Si cette proportion est plus petite qu'un critère de désagrégation établi par l'utilisateur, l'algorithme ne poursuit pas la désagrégation dans cette section de la chaîne d'approvisionnement;

- Si cette proportion est plus grande que le critère, chaque fournisseur sera examiné comme les précédents;
- À mesure que l'algorithme s'enfonce dans la chaîne d'approvisionnement, les contributions à l'impact diminuent et deviennent éventuellement toutes plus petites que le critère de désagrégation.

L'algorithme aura donc identifié les contributeurs majeurs à l'impact total. Le reste de l'impact est réparti sur une infinité de contributeurs individuellement insignifiants, dont l'identité est de peu d'intérêt pour le commanditaire d'une ACV. Avec le recul, cet algorithme aurait dû être baptisé « désagrégation sélective ».

La valeur ajoutée de la désagrégation systématique réside dans cet accès au détail du chemin entre un impact et l'unité fonctionnelle. Cet accès permet :

- La compréhension rapide de la provenance des impacts dans la chaîne d'approvisionnement;
- La validation de l'adéquation entre le modèle de la base de données génériques et la chaîne d'approvisionnement qu'il prétend représenter, lorsque des informations primaires sont disponibles;
- Le remplacement ponctuel de coefficients jugés inadéquats, sans forcer ce choix sur les autres branches de la chaîne d'approvisionnement;
- L'ajustement ponctuel de l'incertitude des coefficients;
- L'échantillonnage corrélé ou non corrélé de coefficients apparaissant dans plusieurs branches de la chaîne d'approvisionnement, lors d'un Monte Carlo;
- La visualisation facile et intuitive de la contribution à l'incertitude de chaque terme;
- Le calcul de contribution à l'incertitude de chaque coefficient à l'incertitude d'un terme;
- La focalisation d'une analyse uniquement sur les termes ayant une influence significative sur le résultat.

L'algorithme en lui-même est très simple et d'un intérêt limité. Son application est mécanique et sans complication. Sa puissance réside dans les possibilités ouvertes pour la phase d'interprétation. Le Chapitre 5 montre comment il se combine avec les autres outils développés dans le cadre de cette thèse afin de diminuer l'incertitude.

## CHAPITRE 5 DISCUSSION GÉNÉRALE

L'intégration des outils proposés concerne soit le développement du modèle (ICV et ÉICV), soit son application. Cette division est le résultat de l'analyse effectuée directement sur les FCs, en-dehors de tout contexte d'application. Ce choix permet de tirer certaines conclusions valides pour toutes les applications. Par exemple, si un UII est faible dans un pays, il ne créera aucune préoccupation, peu importe l'ACV réalisée. Les UII élevés pointent vers un manque de données créant *potentiellement* un biais dans le résultat. La seconde partie de l'intégration des outils, du côté de l'analyste, permet de vérifier si les biais ont le potentiel d'influencer les conclusions d'une étude.

Le Tableau 7 relie les objectifs spécifiques de la thèse avec les outils et les acteurs impliqués. Au sein de chaque section au sujet d'un objectif spécifique, la synergie des outils est exposée. La discussion générale se termine par un retour sur les besoins et pistes de recherches soulignés par la revue de littérature.

Tableau 7 : Outils et acteurs impliqués dans l'atteinte des objectifs spécifiques

Objectif spécifique	Outil(s)	Acteur(s)
Quantification de l'incertitude due à l'incohérence d'échelle spatiale	<ul style="list-style-type: none"> <li>• UII (article 1)</li> </ul>	développeurs ÉICV
Réduction de l'incertitude + parcimonie de complexification	<ul style="list-style-type: none"> <li>• Méthode d'agrégation et UII (article 1)</li> <li>• PS (article 2)</li> <li>• jeux de pondération (article 3)</li> </ul>	développeurs ÉICV et ICV
Outils pour analyste	<ul style="list-style-type: none"> <li>• UII (article 1)</li> <li>• PS (article 2)</li> <li>• désagrégation (article 4)</li> </ul>	Praticien

### 5.1 Objectif 1 : Quantification de l'incertitude

#### 5.1.1 L'agrégation comme source d'incertitudes

Le manque d'informations spatiales empêche la prise en compte de la variabilité du modèle : on tente de représenter une gamme de valeurs par une seule valeur, ce qui augmente l'incertitude. Cette augmentation est représentée par l'indicateur d'augmentation de l'incertitude (UII). Son



calcul consiste simplement à une comparaison de l'incertitude des facteurs de caractérisation avant et après l'agrégation. Son application permet l'identification rapide des catégories d'impacts, substances et régions pour lesquels la variabilité spatiale *pourrait* être un enjeu dans une ACV.

### 5.1.2 Parallèle entre UII et PS

La puissance statistique (PS) et le UII sont deux mesures négativement corrélées au chevauchement des variables aléatoires (VAs) comparées : plus le chevauchement est élevé, plus la PS et le UII seront faibles. Le UII est toutefois plus sophistiqué : il peut comparer plus de deux VAs à la fois, ainsi que donner plus ou moins de poids à certaines.

La Figure 31 montre une certaine corrélation entre UII et SP. Cinq cents paires de VAs de forme lognormale ont été générées. La moyenne de la première a été fixée à 1, et la moyenne de la seconde a été choisie aléatoirement entre 1,1 et 100. Le GSD2 des deux variables a aussi été choisi aléatoirement 1,1 et 8. Pour chaque paire de courbes, le UII et la PS ont été calculés. Un poids de 50% a été choisi pour les deux courbes dans le calcul du UII.

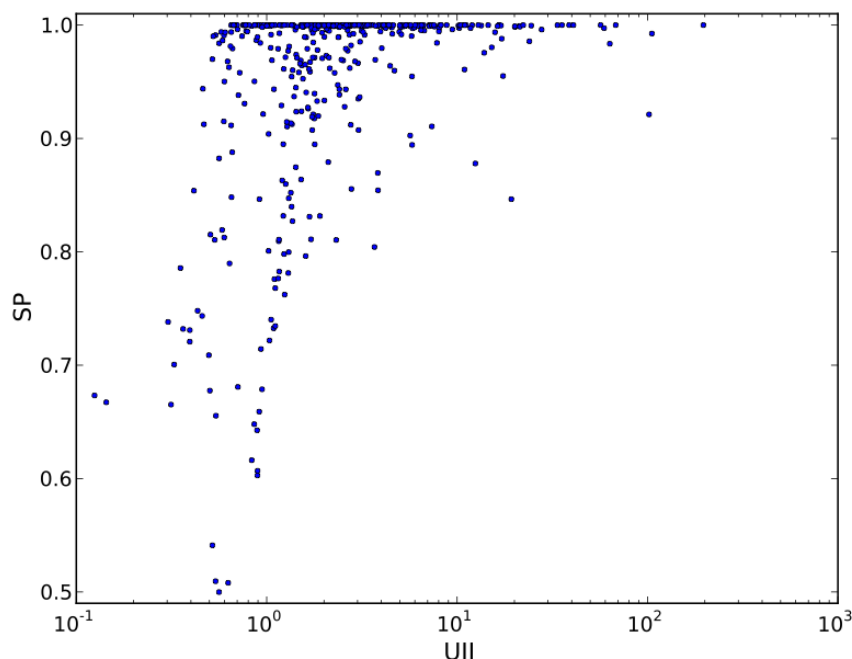


Figure 31 : SP en fonction de UII pour deux variables aléatoires

D'un côté, la valeur de UII n'est pas plafonnée et peut théoriquement atteindre de très grandes valeurs lorsque les courbes sont loin l'une de l'autre. De l'autre côté, la SP prend indirectement en compte la distance entre les courbes, mais au-delà d'une certaine distance, toutes les courbes évaluées ont un SP de 1. Ainsi, la variabilité de UII obtenus aux valeurs de  $SP = 1$  est plus grande qu'avec des SP plus près de 0,5.

Bien que similaires, les deux quantités ne sont pas interchangeables, et leur signification est différente. Le UII répond à la question « Quelle est l'augmentation de l'incertitude dans une opération d'agrégation? », tandis que la notion d'agrégation est absente du calcul de la SP.

L'utilisation des deux quantités devrait être considérée dans la discussion autour de l'agrégation. Cette tâche revient aux développeurs des modèles d'ÉICV, qui ont une connaissance intime des paramètres géographiques (landscape parameters) et des caractéristiques physico-chimiques des polluants ayant une influence sur les FCs.

Le degré de complexification choisi devrait toutefois respecter le principe de parcimonie, ce qui amène au second objectif spécifique de cette thèse.

## **5.2 Objectif 2 : réduction de l'incertitude et parcimonie de complexification**

L'incertitude créée par l'agrégation de petites zones en zones plus grandes peut être limitée par l'agrégation en zones où la variabilité spatiale et l'incertitude sont comparables, au lieu de tirer une ligne arbitraire. Augmenter le nombre de zones signifie complexifier le modèle. Il est possible de vérifier la valeur ajoutée, en regard de l'incertitude, de découper une zone plus grande en plusieurs zones plus petites.

### **5.2.1 Synergie entre PS et UII**

La PS peut être utilisée pour évaluer la différence entre des FCs agrégés à l'aide de différents jeux de pondération. L'article 3 a fait ressortir que la moyenne pondérée par population ou proxy d'émission menait à des FCs relativement proches dans la plupart des cas. En utilisant les deux jeux de pondération dans une agrégation telle que présentée à l'article 1, on disposerait de deux PDFs, que l'on peut ensuite soumettre au test de PS.

Il serait aussi intéressant de calculer la PS entre un FC agrégé et chaque FC natif qui le compose. Si la PS est faible entre un FC agrégé et certains natifs, ils pourraient être utilisés de manière interchangeable, minimisant ainsi le nombre de FCs à conserver dans un modèle d'ÉICV.

### 5.2.2 Synergie entre UII et agrégation par industrie

Le manque d'informations géographiques pour la sélection d'un FC natif est analogue à ne pas utiliser l'information sur l'identité des processus élémentaires responsables des émissions. Dans cette analogie, les FCs natifs sont remplacés par des FCs spécifiques aux secteurs, et le jeu de pondération représentant la distribution spatiale des émissions est remplacé par la quantité normalisée d'émissions par secteur. Calculer un UII en agrégeant les FCs spécifiques aux secteurs pourrait guider le choix des pays et substances pour lesquels la différenciation en secteurs serait acceptable.

### 5.2.3 Synergie entre PS et agrégation par industrie

L'application de la PS au développement des FCs spatialisés a utilisé une représentation telle que montrée à la Figure 32. Cette représentation pourrait être adaptée pour évaluer la différence entre les FC spécifiques aux secteurs, au sein d'un même pays.

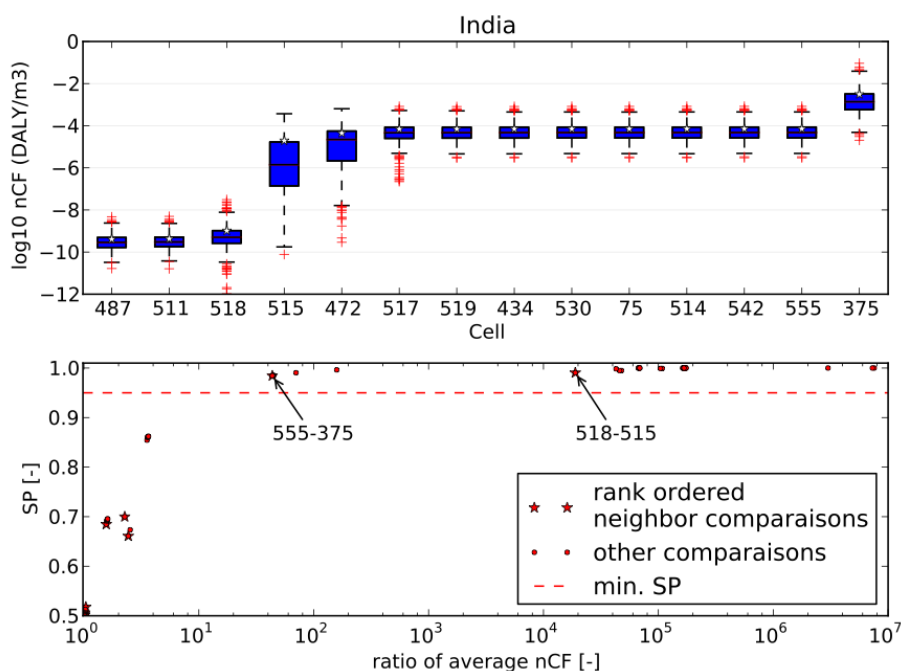


Figure 32 : Monte Carlo et PS pour les FC de l'eau de qualité S3 en Inde.

## **5.3 Objectif 3 : intégration à la pratique de l'analyste**

### **5.3.1 Spatialisation ponctuelle**

La spatialisation ponctuelle consiste à la recherche d'informations sur le lieu de certaines activités de production afin de remplacer leurs FC agrégés par des FC natifs. Cette étape ne sera pas nécessaire pour toutes les ACVs. La Figure 33 montre l'algorithme de décision déterminant si une étude nécessite une spatialisation ponctuelle. Cette opération sera nécessaire en présence des caractéristiques suivantes : les principaux contributeurs à l'incertitude et au score d'impacts sont caractérisés par des FCs à UII élevés; les scores des systèmes à comparer sont peu différenciés; une combinaison de remplacement de FCs agrégés par des FCs natifs pourrait inverser les conclusions. Cet algorithme ne demande pas de jugements subtils et repose uniquement sur la comparaison systématique de valeurs quantitatives. Son application serait plus efficace si elle était effectuée par un logiciel plutôt que manuellement, par un analyste.

Si la spatialisation ponctuelle est requise, il faut prioriser la collecte de données : quels sont les remplacements maximisant la différenciation des scores d'impacts? Cette optimisation, elle aussi, pourrait être complètement automatisée. L'analyste se verrait proposer une liste ordonnée d'activités de transformation à localiser. Les ressources nécessaires pour l'obtention de l'information spatiale pourraient être évaluées avant de se lancer dans leur recherche.

La désagrégation systématique intervient à cette étape du calcul. Dans le tableau de résultat produit par cette méthode, chaque flux élémentaire émis par chaque occurrence de processus élémentaire est rapporté séparément, et le chemin détaillé entre l'usine et le flux de référence est accessible. Ainsi, le FC natif sélectionné pourrait être appliqué à un seul flux.

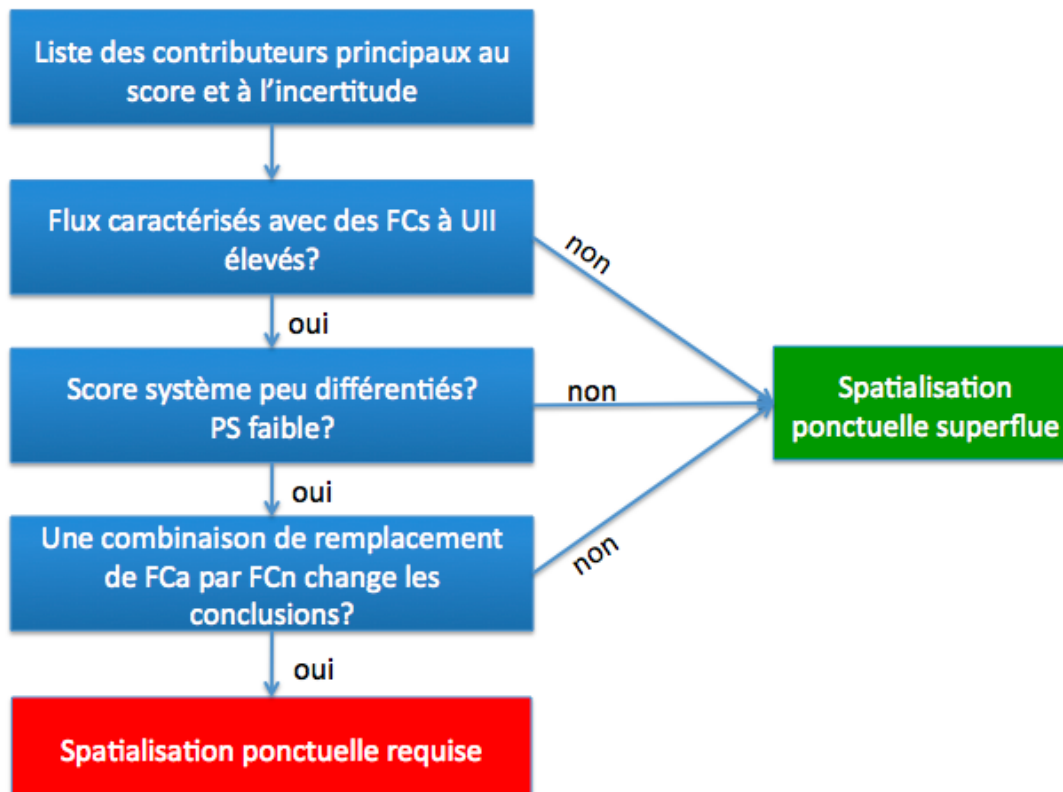


Figure 33 : Identification du besoin de spatialisation ponctuelle

La Figure 34 montre le schéma d'application de la spatialisation ponctuelle. L'analyste débute avec le flux ayant le plus de potentiel de diminution de l'incertitude. Il obtient l'information permettant de choisir le FC natif. Celui-ci ne créera pas nécessairement toute la diminution de l'incertitude escomptée. L'effet du remplacement est évalué et les remplacements se poursuivent jusqu'à ce qu'une conclusion soit obtenue.

La spatialisation ponctuelle est donc un processus mettant en branle les concepts de UII, PS et de désagrégation systématique. On peut en dresser un schéma général, mais le détail de son application et ses résultats seront différents en fonction de chaque ACV.

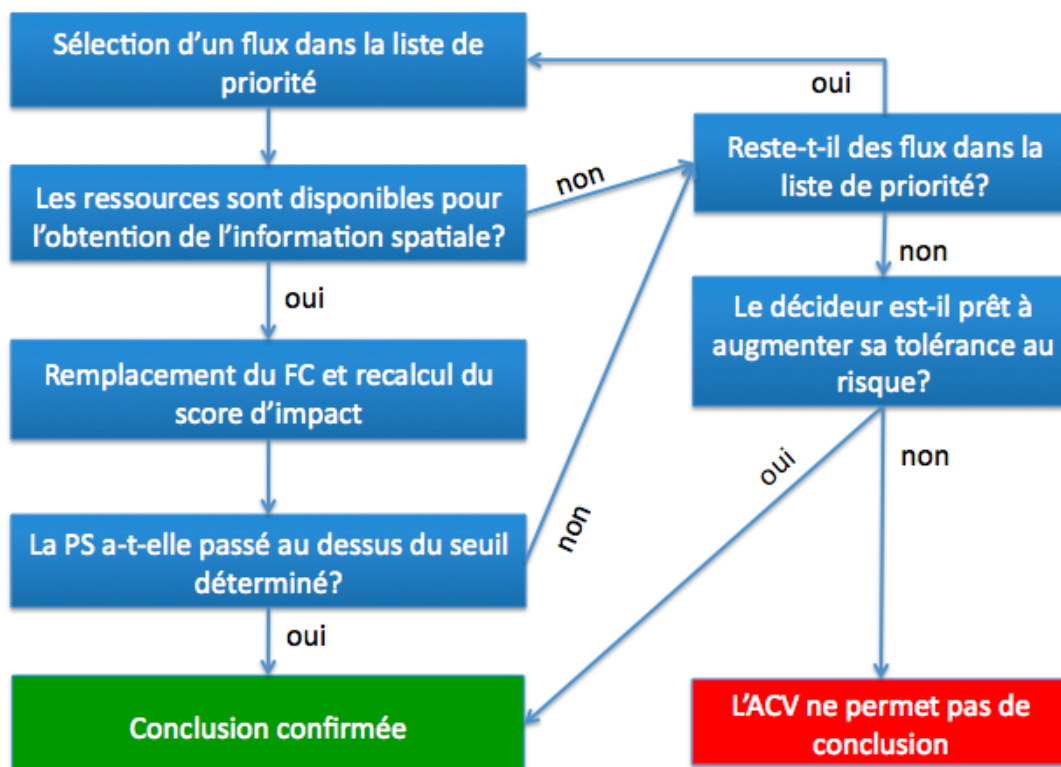


Figure 34 : Schéma d'application de la spatialisation ponctuelle

## 5.4 Retour sur les besoins

La revue de littérature a fait ressortir certaines lacunes et pistes de recherche. La suite d'outils développée répond à plusieurs de ces besoins, brièvement rappelés ici.

### 5.4.1 Quantification de l'incertitude et de la compréhension de sa provenance

La quantification de l'incertitude est requise par les normes ISO. Au-delà des besoins normatifs, les praticiens doivent être en mesure de comprendre la signification et la provenance des incertitudes afin de renseigner les décideurs. La technique de réagrégation et le UII font partie des outils nécessaires pour répondre à ces besoins.

### 5.4.2 Interprétation de l'incertitude

Une poignée d'études a utilisé des tests statistiques pour discuter de la confiance envers les conclusions. D'autres se sont contentées de pointer vers un chevauchement des scores. La

puissance statistique de test est maintenant disponible pour être intégrée à la pratique des analystes.

### **5.4.3 L'incertitude comme guide de développement des modèles**

Van Zelm et Huijbregts ont utilisé l'incertitude pour évaluer la pertinence d'une complexification du modèle d'impacts des pesticides. Bien qu'ils n'aient pas formulé explicitement le concept, leur contribution sous-entend leur vision de l'incertitude comme un guide pointant vers les opportunités d'amélioration du modèle. Les concepts de UII et SP sont des outils concrets à utiliser dans ce contexte.

### **5.4.4 La nature itérative de l'ACV**

Les normes ISO ainsi que plusieurs auteurs ont souligné la nature itérative de l'application de l'ACV. La spatialisation ponctuelle, mettant en branle tous les outils développés dans cette thèse, représente un exemple du mouvement de va-et-vient entre la cueillette de données et l'interprétation des résultats. Un praticien bien renseigné sur les bénéfices d'une cueillette de données supplémentaires pourra focaliser ses énergies dans la bonne direction. Ce va-et-vient nécessite aussi la prise en compte des pratiques des analystes dans le développement des modèles.

## CHAPITRE 6 CONCLUSION ET RECOMMANDATIONS

Cette thèse démontre que des outils mathématiques simples peuvent réduire l'incertitude en ACV. Ces outils permettent de disséquer les sources d'incertitudes tant au niveau de la construction des modèles que dans leur application. Dans le premier cas, une meilleure compréhension de l'incertitude indique quel niveau de complexification est nécessaire. Dans le cas de l'application par les praticiens, les outils présentés pointent vers les données à remplacer pour minimiser l'incertitude.

À ce point-ci de l'application des outils, il n'est pas possible de formuler des recommandations générales au sujet des pays, substances ou catégories d'impacts nécessitant une attention particulière. Le développement au cas par cas sera nécessaire, ce qui représente plus de travail. Toutefois, s'il est effectué en tenant compte des pratiques de l'analyste, il ne représentera pas nécessairement un alourdissement du quotidien des analystes.

Le reste de cette thèse souligne les leçons tirées qui devraient être diffusées à toute la communauté.

### **6.1 L'incertitude comme guide d'amélioration du modèle et des données**

Il est important que la communauté de l'ACV change sa perception du rôle de l'incertitude. Jusqu'à présent, l'état d'esprit a été réactif : on lui porte attention après coup, du bout des lèvres, parce que ISO le demande. Une attitude proactive et positive est bien plus productive. L'incertitude pointe vers les opportunités d'amélioration de modèles et de la qualité des données. Elle renseigne sur la valeur ajoutée des complexifications et force à confronter la réalité avant de se lancer dans des échafaudages superflus.

L'incertitude a montré que le manque d'informations spatiales au niveau de l'inventaire ainsi que l'utilisation de FCs agrégés n'était pas problématiques dans plusieurs cas. Pour les autres cas, l'augmentation de l'incertitude est maintenant quantifiée et les analystes disposent d'un outil simple pour en mitiger les effets. Pour en arriver à cette conclusion, il a fallu créer des jeux de



pondération. Cette entrée de données, à mi-chemin dans la séquence de calcul, a pour rôle de rapprocher le modèle d'ACV de la réalité qu'il prétend représenter.

L'article sur la création de FCs par industrie applique une approche de « pertinence et faisabilité ». L'incertitude est l'un des aspects à prendre en compte dans l'évaluation de la pertinence d'une complexification.

## **6.2 La coordination de haut niveau dans la complexification de l'ACV**

Cette thèse a été rendue nécessaire par le manque de coordination de la recherche en inventaire et en évaluation des impacts. La recherche en ACV est effectuée par un groupe aux compétences et intérêts hétérogènes. Les chercheurs en ICV ne sont pas nécessairement habilités à évaluer les implications de leurs décisions sur l'ÉICV, et vice versa.

Ceci dit, un « excès » de collaboration pourrait mener à une paralysie des avancées. Par exemple, avant de se lancer dans la création de FCs spatialisés à fine échelle, un chercheur en ÉICV aurait pu constater que l'inventaire ne permet pas leur opérationnalisation dans son état actuel et décider de ne pas les développer.

Il n'y a aucun doute que si la pertinence d'une complexification est démontrée, elle devrait être mise de l'avant. Toutefois, la faisabilité de son implémentation et l'impact sur l'amont et l'aval du calcul devraient faire partie de la discussion à son sujet.

Cette collaboration s'opère déjà. Deux initiatives en cours en témoignent. Les responsables de la méthode IMPACT World+ se sont joints à des chercheurs en inventaire pour proposer un format d'échange de données permettant une association fluide des flux élémentaires avec leurs FCs pour la version 3 de ecoinvent. Les mêmes acteurs ont aussi produit une fiche de spécification à l'intention des développeurs de logiciels d'ACV pour assurer l'intégration de la méthode dans les futures versions du logiciel.

### **6.3 Développement d'outils pragmatiques d'application**

Il est utopique de penser que le développement du modèle de l'ACV règlera tous les problèmes d'incertitudes. La diversité de situations à laquelle il est appliqué rend l'exercice d'abstraction beaucoup trop ardu. Il est plus pragmatique de créer un ensemble d'outils cohérent qu'un être humain pourra utiliser, adapter et interpréter en regard de chaque situation unique.

C'est pourquoi chaque nouveau développement de l'ACV doit être pensé en fonction de son interaction avec les praticiens. Il faut prendre en compte les pratiques actuelles, les contraintes de temps et de ressources, les outils informatiques et la structure de pensée des praticiens. En poussant la construction d'une complexification jusqu'à un point où elle s'insérera dans la pratique, on bénéficie d'un retour plus rapide sur ses impacts réels dans la prise de décision environnementale. Une boucle serrée de rétroaction recherche-application-recherche est le meilleur moyen de focaliser les efforts où ils feront une différence.

### **6.4 Développements à venir**

La désagrégation systématique s'avère un outil très fertile. Elle est essentielle dans l'application des FCs temporels (travaux de Annie Levasseur, Didier Beloin-St-Pierre et Ariane Pinsonneault). L'algorithme permet aussi un échantillonnage corrélé ou non corrélé pour les paramètres apparaissant à plusieurs endroits au sein du même système de produit.

D'une manière générale, la corrélation intra-système et inter-systèmes devrait être investiguée. La situation actuelle de cette question est comparable à celle de l'incertitude et du manque d'informations spatiales il y a quelques années. On savait que ce manque d'informations augmentait l'incertitude, mais il n'était pas possible de le quantifier. Présentement, on sait que l'hypothèse de corrélation inter-systèmes totale dans les analyses comparatives sous-estime l'incertitude, mais on n'a aucune idée de la prévalence et de la gravité des problèmes causés par ce biais. La PS serait utile dans l'identification de ces situations problématiques.

## BIBLIOGRAPHIE

- Alcamo, J., P. Döll, et al. (2003). "Development and testing of the WaterGAP 2 global model of water use and availability." Hydrological Sciences Journal **48**(3): 317–337.
- Baitz, M., S. Albrecht, et al. (2013). "LCA's theory and practice: like ebony and ivory living in perfect harmony?" The International Journal of Life Cycle Assessment **18**(1): 5-13.
- Bellekom, S., J. Potting, et al. (2006). "Feasibility of Applying Site-dependent Impact Assessment of Acidification in LCA." The International Journal of Life Cycle Assessment **11**(6): 417-424.
- Bennett, D. H., T. McKone, et al. (2002). "Defining Intake Fraction." Environmental Science & Technology **36**(9): 206A-211A.
- Berthoud, A., P. Maupu, et al. (2011). "Assessing freshwater ecotoxicity of agricultural products in life cycle assessment (LCA): a case study of wheat using French agricultural practices databases and USEtox model." The International Journal of Life Cycle Assessment **16**(8): 841-847.
- Björklund, A. (2002). "Survey of approaches to improve reliability in LCA." The International Journal of Life Cycle Assessment **7**(2): 64-72.
- Boulay, A.-M., C. Bulle, et al. (2011). "Regional Characterization of Freshwater Use in LCA: Modeling Direct Impacts on Human Health." Environmental Science & Technology **45**(20): 8948–8957.
- Boulay, A.-M., M. Margni, et al. (2009). Regionalisation of impacts from water use Paper presented at the Life Cycle Assessment. LCA IX. Boston.
- Bourgault, G., P. Lesage, et al. (2012). "Systematic disaggregation: a hybrid LCI computation algorithm enhancing interpretation phase in LCA." International journal of Life Cycle Assessment **17**(6): 774-786.
- Bourgault, G., P. Lesage, et al. (2012). Uncertainty and correlation: Thinking ahead. LCA XII. Tacoma.
- Bowen, G. (2008). "Naturalistic inquiry and the saturation concept: a research note." Qualitative Research **8**(1): 137-152.
- Ciroth, A. (2004). "Uncertainties in life cycle assessment." The International Journal of Life Cycle Assessment **9**(3): 141-142.
- Ciroth, A. (2006). "Validation – The Missing Link in Life Cycle Assessment. Towards pragmatic LCAs." The International Journal of Life Cycle Assessment **11**(5): 295-297.
- Ciroth, A., G. Fleischer, et al. (2004). "Uncertainty calculation in life cycle assessments." The International Journal of Life Cycle Assessment **9**(4): 216-226.
- Clavreul, J., D. Guyonnet, et al. (2012). "Quantifying uncertainty in LCA-modelling of waste management systems." Waste Management **32**(12): 2482-2495.

- Cooper, J., M. Noon, et al. (2012). "Parameterization in Life Cycle Assessment inventory data: review of current use and the representation of uncertainty." The International Journal of Life Cycle Assessment **17**(6): 689-695.
- Damasio, A. (1997). L'erreur de Descartes: la raison des émotions. Paris, Odile Jacob.
- Draucker, L., S. Kaufman, et al. (2011). "Moving Forward on Product Carbon Footprint Standards." Journal of Industrial Ecology **15**(2): 169-171.
- Edwards, B. (1999). Drawing on the right side of the brain, Tarcher.
- Fréchette-Marleau, S., V. Bécaert, et al. (2008). "Evaluating the variability of aquatic acidification and photochemical ozone formation characterization factors for Canadian emissions." The International Journal of Life Cycle Assessment **13**(7): 593-604.
- Geisler, G., S. Hellweg, et al. (2005). "Uncertainty Analysis in Life Cycle Assessment (LCA): Case Study on Plant-Protection Products and Implications for Decision Making." The International Journal of Life Cycle Assessment **10**(3): 184-192.
- Hammonds, J., F. Hoffman, et al. (1994). An Introductory Guide to Uncertainty Analysis in Environmental and Health Risk Assessment. US Department of Energy: 1-47.
- Hauschild, M., M. Goedkoop, et al. (2013). "Identifying best existing practice for characterization modeling in life cycle impact assessment." The International Journal of Life Cycle Assessment **18**(3): 683-697.
- Heijungs, R. (2010). "Sensitivity coefficients for matrix-based LCA." The International Journal of Life Cycle Assessment **15**(5): 511-520.
- Heijungs, R. (2012). Spatial differentiation, GIS-based regionalization, hyperregionalization, and the boundaries of LCA. Environment and Energy. FrancoAngeli: 167-175.
- Heijungs, R. and R. Frischknecht (2005). "Representing Statistical Distributions for Uncertain Parameters in LCA. Relationships between mathematical forms, their representation in EcoSpold, and their representation in CMLCA." The International Journal of Life Cycle Assessment **10**(4): 248-254.
- Heijungs, R. and S. Suh (2002). The Computational Structure of Life Cycle Assessment. Leiden, Kluwer Academic Publishers.
- Hertwich, E. G., T. E. McKone, et al. (2000). "A Systematic Uncertainty Analysis of an Evaluative Fate and Exposure Model." Risk Analysis **20**(4): 439-454.
- Hines, W., D. Montgomery, et al. (2003). Probability and Statistics in Engineering, Hoboken, NJ: Wiley.
- Huijbregts, M. (1998). "Application of uncertainty and variability in LCA. Part I: A General Framework for the Analysis of Uncertainty and Variability in Life Cycle Assessment." The International Journal of Life Cycle Assessment **3**(5): 273-280.
- Huijbregts, M. (1998). "Application of Uncertainty and Variability in LCA. Part II: Dealing with parameter uncertainty and uncertainty due to choices in life cycle assessment." The International Journal of Life Cycle Assessment **3**(6): 343-351.

- Huijbregts, M., W. Gilijamse, et al. (2003). "Evaluating Uncertainty in Environmental Life-Cycle Assessment. A Case Study Comparing Two Insulation Options for a Dutch One-Family Dwelling." Environmental Science & Technology **37**(11): 2600-2608.
- Huijbregts, M., G. Norris, et al. (2001). "Framework for modelling data uncertainty in life cycle inventories." The International Journal of Life Cycle Assessment **6**(3): 127-132.
- Humbert, S., R. Manneh, et al. (2009). "Assessing regional intake fractions in North America." Science of The Total Environment **407**(17): 4812-4820.
- Ibáñez-Forés, V., M.-D. Bovea, et al. (2011). "Life cycle assessment of ceramic tiles. Environmental and statistical analysis." The International Journal of Life Cycle Assessment **16**(9): 916-928.
- Imbeault-Tétrault, H., O. Jolliet, et al. (2013). "Analytical Propagation of Uncertainty in Life Cycle Assessment Using Matrix Formulation." Journal of Industrial Ecology **17**(4): 485-492.
- Jolliet, O., R. Helmes, et al. (2010). Global high resolution freshwater eutrophication impact characterization. LCA X. Portland, Oregon, USA.
- Jolliet, O., M. Saadé, et al. (2005). Analyse du cycle de vie: comprendre et réaliser un écobilan. Lausanne, Presses Polytechniques et Universitaires Romandes.
- Kalos, M. and P. Whitlock (2008). Monte Carlo Methods. Weinheim, Wiley-VCH.
- Klöpffer, W. (2013). "Comment on the editorial note by Baitz et XXI aliis." The International Journal of Life Cycle Assessment **18**(1): 14-16.
- Lenzen, M., K. Kanemoto, et al. (2012). "Mapping the Structure of the World Economy." Environmental Science & Technology **46**(15): 8374-8381.
- Lenzen, M., D. Moran, et al. (2013). "Building Eora: a global multi-region input-output database at high country and sector resolution." Economic Systems Research **25**(1): 20-49.
- Lloyd, S. and R. Ries (2007). "Characterizing, Propagating, and Analyzing Uncertainty in Life-Cycle Assessment: A Survey of Quantitative Approaches." Journal of Industrial Ecology **11**(1): 161-179.
- MacLeod, M., D. H. Bennett, et al. (2004). "Dependence of Intake Fraction on Release Location in a Multimedia Framework." Journal of Industrial Ecology **8**(3): 89-102.
- Manneh, R., M. Margni, et al. (2010). "Spatial Variability of Intake Fractions for Canadian Emission Scenarios: A Comparison between Three Resolution Scales." Environmental Science & Technology **44**(11): 4217-4224.
- Margni, M. D. (2010). Research Program Description. Montréal: 1-12.
- Mattila, T., P. Leskinen, et al. (2012). "Uncertainty in environmentally conscious decision making: beer or wine?" The International Journal of Life Cycle Assessment **17**(6): 696-705.
- McKay, M. and W. Conover (1979). "A comparison of three methods for selecting values of input variables in the analysis of output from a computer code." Technometrics **21**(2): 239-245.

- Mutel, C., S. Pfister, et al. (2012). "GIS-Based Regionalized Life Cycle Assessment: How Big Is Small Enough? Methodology and Case Study of Electricity Generation." Environmental Science & Technology **46**(2): 1096-1103.
- Mutel, C. L., L. de Baan, et al. (2013). "Two-Step Sensitivity Testing of Parametrized and Regionalized Life Cycle Assessments: Methodology and Case Study." Environmental Science & Technology **47**(11): 5660-5667.
- Mutel, C. L. and S. Hellweg (2009). "Regionalized Life Cycle Assessment: Computational Methodology and Application to Inventory Databases." Environmental Science & Technology **43**(15): 5797-5803.
- Nigge, K. (2001). "Generic spatial classes for human health impacts, Part I." The International Journal of Life Cycle Assessment **6**(5): 257-264.
- Nigge, K. (2001). "Generic spatial classes for human health impacts, Part II." The International Journal of Life Cycle Assessment **6**(6): 334-338.
- Norris, G. (2002). "Life cycle emission distributions within the economy: Implications for life cycle impact assessment." Risk Analysis **22**(5): 919-930.
- Norris, G. A. (2002). "Impact Characterization in the Tool for the Reduction and Assessment of Chemical and Other Environmental Impacts." Journal of Industrial Ecology **6**(3-4): 79-101.
- Owens, J. (1999). "Why life cycle impact assessment is now described as an indicator system." The International Journal of Life Cycle Assessment **4**(2): 81-86-86.
- Owens, J. W. (1997). "Life-Cycle Assessment: Constraints on Moving from Inventory to Impact Assessment." Journal of Industrial Ecology **1**(1): 37-49.
- Owsianiak, M., R. K. Rosenbaum, et al. (2013). "Addressing Geographic Variability in the Comparative Toxicity Potential of Copper and Nickel in Soils." Environmental Science & Technology **47**(7): 3241-3250.
- Pennington, D., M. Margni, et al. (2005). "Multimedia fate and human intake modeling: Spatial versus nonspatial insights for chemical emissions in Western Europe." Environmental Science & Technology **39**(4): 1119-1128.
- Peters, G. P. (2007). "Efficient algorithms for life cycle assessment, input-output analysis, and Monte-Carlo analysis." The International Journal of Life Cycle Assessment **12**(6): 373-380.
- Potting, J. and M. Hauschild (1997). "Predicted environmental impact and expected occurrence of Actual Environment Impact: Part I." The International Journal of Life Cycle Assessment **2**(3): 171-177.
- Potting, J. and M. Hauschild (2006). "Spatial Differentiation in Life Cycle Impact Assessment: A decade of method development to increase the environmental realism of LCIA." The International Journal of Life Cycle Assessment **11**: 11-13.
- Potting, J., W. Schöpp, et al. (1998). "Site-dependent life-cycle impact assessment of acidification." Journal of Industrial Ecology **2**(2): 63-87.

- Rack, M., S. Valdivia, et al. (2013). "Life Cycle Impact Assessment—where we are, trends, and next steps: a late report from a UNEP/SETAC Life Cycle Initiative workshop and a few updates from recent developments." The International Journal of Life Cycle Assessment **18**(7): 1413-1420.
- Reap, J., F. Roman, et al. (2008). "A survey of unresolved problems in life cycle assessment. Part I: goal and scope and inventory analysis." The International Journal of Life Cycle Assessment **13**(4): 290-300-300.
- Reap, J., F. Roman, et al. (2008). "A survey of unresolved problems in life cycle assessment. Part II: impact assessment and interpretation." The International Journal of Life Cycle Assessment **13**(5): 374-388.
- Rochat, D., M. Margni, et al. (2006). "Continent-specific Intake Fractions and Characterization Factors for Toxic Emissions: Does it make a Difference?" The International Journal of Life Cycle Assessment **11**: 55-63.
- Röös, E., C. Sundberg, et al. (2011). "Uncertainties in the carbon footprint of refined wheat products: a case study on Swedish pasta." The International Journal of Life Cycle Assessment **16**(4): 338-350.
- Ross, S. and D. Evans (2002). "Excluding site-specific data from the lca inventory: how this affects life cycle impact assessment." The International Journal of Life Cycle Assessment **7**(3): 141-150.
- Ross, S., D. Evans, et al. (2002). "How LCA studies deal with uncertainty." The International Journal of Life Cycle Assessment **7**(1): 47-52.
- Roy, P.-O. (2012). Modélisation environnementale régionalisée à l'échelle mondiale de l'acidification terrestre et aquatique dans le cadre de l'évaluation des impacts du cycle de vie. Ph.D., École Polytechnique de Montreal.
- Saltelli, A. and P. Annoni (2010). "How to avoid a perfunctory sensitivity analysis." Environmental Modelling & Software **25**(12): 1508-1517.
- Seppälä, J., M. Posch, et al. (2006). "Country-dependent Characterisation Factors for Acidification and Terrestrial Eutrophication Based on Accumulated Exceedance as an Impact Category Indicator." The International Journal of Life Cycle Assessment **11**(6): 403-416.
- Sills, D. L., V. Paramita, et al. (2012). "Quantitative Uncertainty Analysis of Life Cycle Assessment for Algal Biofuel Production." Environmental Science & Technology **47**(2): 687-694.
- Sleeswijk, A. W. (2010). Regional LCA in a global perspective: A basis for spatially differentiated environmental life cycle assessment. Ph.D., Leiden University, .
- Spatari, S. and H. L. MacLean (2010). "Characterizing Model Uncertainties in the Life Cycle of Lignocellulose-Based Ethanol Fuels." Environmental Science & Technology **44**(22): 8773-8780.
- Stratton, R. W., H. M. Wong, et al. (2011). "Quantifying Variability in Life Cycle Greenhouse Gas Inventories of Alternative Middle Distillate Transportation Fuels." Environmental Science & Technology **45**(10): 4637-4644.

- Tessum, C. W., J. D. Marshall, et al. (2012). "A Spatially and Temporally Explicit Life Cycle Inventory of Air Pollutants from Gasoline and Ethanol in the United States." Environmental Science & Technology **46**(20): 11408-11417.
- van Zelm, R. and M. A. J. Huijbregts (2013). "Quantifying the Trade-off between Parameter and Model Structure Uncertainty in Life Cycle Impact Assessment." Environmental Science & Technology **47**(16): 9274-9280.
- Venkatesh, A., P. Jaramillo, et al. (2011). "Uncertainty in Life Cycle Greenhouse Gas Emissions from United States Natural Gas End-Uses and its Effects on Policy." Environmental Science & Technology **45**(19): 8182-8189.
- Weber, C. L. (2012). "Uncertainty and Variability in Product Carbon Footprinting." Journal of Industrial Ecology **16**(2): 203-211.
- Wegener Sleeswijk, A. (2011). "Regional LCA in a global perspective. A basis for spatially differentiated environmental life cycle assessment." The International Journal of Life Cycle Assessment **16**(2): 106-112.
- Weidema, B., C. Bauer, et al. (2013). Overview and Methodology: Data Quality Guidelines for the Ecoinvent Database Version 3. Dübendorf, Switzerland, Ecoinvent Center.
- Werner, F. and R. Scholz (2002). "Ambiguities in decision-oriented Life Cycle Inventories: The Role of mental models." The International Journal of Life Cycle Assessment **7**(6): 330-338.
- Wiedmann, T., M. Lenzen, et al. (2007). "Examining the global environmental impact of regional consumption activities — Part 2: Review of input–output models for the assessment of environmental impacts embodied in trade." Ecological Economics **61**(1): 15-26.



**ANNEXE 1 – Article 1 : « Quantification of the increase in  
characterization factor uncertainty due to spatial  
variability»**

# 1 **Quantification of the increase in characterization factor uncertainty due to spatial** 2 **variability**

3 Guillaume Bourgault, Pascal Lesage, Anne-Marie Boulay, Réjean Samson

## 4 **Abstract**

5 Life Cycle Inventory calculated using commonly available LCI databases usually have a  
6 country-scale resolution. The spatial resolution of some regionalized impact assessment  
7 models (LCIA) is much finer. This incoherence of spatial resolution forces the life cycle  
8 modelers to choose a characterization factor (CF) that does not have the same domain of  
9 spatial validity as that of the LCI result. For some impact categories, CF can span several  
10 orders of magnitude within a given country. This situation introduces an additional  
11 uncertainty that has not yet been quantified. The present paper proposes a systematic way  
12 of aggregating a set of CF from a fine “native” delineation into a set with any spatial  
13 delineation (country, continent, etc.) It introduces an uncertainty increase indicator (UII)  
14 measuring the effect of the aggregation on the uncertainty of CF. The methodology is  
15 applied to a freshwater use direct impact on human health model. In some cases, the lack of  
16 spatial information forces a practitioner to use a CF with a coefficient of variation up to 10  
17 times larger. However, in most cases examined, the intrinsic uncertainty of CF before  
18 aggregation is so large as to cloud the uncertainty contribution of aggregation.

## 19 **1 Introduction**

20 In their evaluation of best practices in life cycle impact assessment (LCIA) (Hauschild,  
21 Goedkoop et al. 2013), Hauschild et al. included a criterion on geographic differentiation  
22 and global coverage of a model, since supply chains usually span several continents. The  
23 latest developments in LCIA have taken this guideline into account, and models with global  
24 coverage and a very fine spatial resolution were generated. It was shown that the value of  
25 characterization factors (CF) can vary significantly over space (see Table 1 for examples  
26 and references).

27 From the inventory perspective, new process-based life cycle inventory (LCI) databases  
28 such as ecoinvent v3 (Weidema, Bauer et al. 2013) and environmentally extended multi-

29 region input-output models such as Eora (Lenzen, Kanemoto et al. 2012) are significant  
30 steps towards representing global supply chains. Despite this extension of the geographic  
31 scope, the spatial resolution of an LCI database is not as fine as the one developed by the  
32 latest LCIA models for so-called regional impact categories. In a typical life cycle assessment  
33 (LCA), foreground processes can be easily geolocated, making it simple to identify the  
34 appropriate spatially differentiated CF. However, it will often be relevant to extend  
35 regionalized LCIA to the background unit processes, which are usually represented by  
36 national generic data. The applicability of the spatialized CFs is therefore compromised.

37 Hauschild et al. also set out an applicability criterion: “The characterization factors [should  
38 be] straightforward to apply for general LCA practitioners and in most market-relevant LCA  
39 software tools” (Hauschild, Goedkoop et al. 2013). We are confident that new structures  
40 will be implemented in software tools to ensure the application of spatialized CFs if these  
41 structures consider the interface with the inventory. To ensure this applicability, a  
42 harmonization step is required.

43 A simple method to use spatially differentiated CFs in the existing LCA framework is to  
44 aggregate CFs (with a usually finer delineation) over areas consistent with the delineations  
45 of the LCI (usually coarser). This aggregation procedure will almost consistently increase  
46 the uncertainty of the CFs since the aggregated CFs represent a larger territory, leading to  
47 potentially greater variability (Rack, Valdivia et al. 2013).

48 The purpose of this paper is twofold. First, a systematic procedure to quantify the increase  
49 of uncertainty caused by the aggregation of CFs from a finer to a coarser spatial delineation  
50 is proposed. Second, the increased uncertainty created by the aggregation is calculated and  
51 compared with the uncertainty of the native-scale CFs for one impact category, namely the  
52 direct impacts of freshwater consumption on human health (Boulay, Bulle et al. 2011).

53 To quantify additional uncertainty, we introduced a simple *uncertainty increase indicator*  
54 (UII) that can be calculated for any regional impact category, regardless of the initial and  
55 final levels of aggregation and which is equally valid for midpoint and endpoint indicators.  
56 In an LCA, the indicator helps to determine the processes for which more precise

57 information on emission location would significantly reduce uncertainty, helping to  
58 prioritize data collection.

59 **Table 1: Examples of the spatial resolutions of regional impact categories**

Impact category	Spatial resolution	Spatial variability (CF <sub>max</sub> /CF <sub>min</sub> )	References
Eutrophication	0.5° x 0.5° grid	10 <sup>9</sup>	(Helmes, Huijbregts et al. 2012)
Human toxicity	North American subwatershed and 2° x 2.5° air grid	10 <sup>7</sup> - 10 <sup>9</sup> (variation on intake fractions)	(Manneh, Margni et al. 2010)
Acidification	European country	100 - 1000	(Potting, Schöpp et al. 1998), (Potting, Schöpp et al. 1998; Posch, Seppälä et al. 2008)
	States in USA	~8	(Norris 2002)
	2° x 2.5° grid	200 - 25 000	(Roy, Deschênes et al. 2012)
Water use on human health	Country/watershed intersection (808 CF)	10 <sup>4</sup>	(Boulay, Bulle et al. 2011)
Land use	193 ecoregions in Canada	8	(Saad, Margni et al. 2011)
Thermal pollution	Specific to each lake/river	To be defined	(Verones, Hanafiah et al. 2012)

## 60 2 Methodology

### 61 2.1 Abandoning the term *generic*

62 We suggest the use of a new set of terms better suited than *generic* and *spatialized* to refer  
63 to the CFs created with the methodology set out in this paper.

64 LCIA model developers choose spatial resolution based on different factors: the rate of  
65 change over space of sensitive parameters, data availability, etc. A set of CFs is then set out  
66 according to what is known as the *native resolution* or *LCIA-defined resolution*. Note that  
67 while the resolution may not be the finest developed (Mutel, Pfister et al. 2012), it was  
68 deemed most relevant by developers: as fine as necessary, as coarse as possible.

69 Each of the CFs has a domain of spatial validity: it is deemed representative of the potential  
 70 impact of an elementary flow in a certain area (a grid cell, a watershed, etc.). Best practices  
 71 should lead method developers to report the uncertainty of these CFs, and the uncertainty  
 72 should include the variability of the CFs within the chosen domain.

73 This methodology aggregates many native CFs to create a single CF with a larger domain of  
 74 spatial validity. Correspondingly, the spatial variability might be greater, leading to greater  
 75 uncertainty. Both the initial CFs and the CF resulting from aggregation are spatialized. For  
 76 this reason, the terms *native* CF (nCF) and *aggregated* CF (aCF) are used in this paper. We  
 77 avoid the term *generic*, proposed by the generic/site-dependant/site-specific framework  
 78 set out by Potting et al (Potting, Schöpp et al. 1998) and used by many authors (Huijbregts,  
 79 Lundi et al. 2003; Hauschild, Jolliet et al. 2011; Struijs, Beusen et al. 2011; Roy, Deschênes et  
 80 al. 2012; Kounina, Margni et al. 2013; Owsianiak, Rosenbaum et al. 2013; Rack, Valdivia et  
 81 al. 2013), because of the arbitrary resolution separating site-dependent and generic CFs  
 82 and the fact that the nature of the aggregated CFs is not radically different from that of the  
 83 native CFs.

## 84 2.2 Aggregation procedure

85 We propose the design of the probability density function (PDF) of an aCF based on the  
 86 weighted average of the PDF of nCF:

$$aPDF_{j,k,m} = \sum_{i=1}^{n_{j,m}} nPDF_{i,j,m} \cdot \hat{w}_{i,j,k,m} \quad \text{equation 1}$$

87 where  $k$  is a specific elementary flow,  $m$  is an impact category,  $aPDF_{j,k,m}$  is the PDF of the aCF  
 88 of geographical zone  $j$  (LCI resolution),  $nPDF_{i,k,m}$  is the PDF of the nCF of geographical zone  $i$   
 89 (native LCIA resolution), and  $n_{j,m}$  is the number of  $i$  geographical zones within geographical  
 90 zone  $j$ . The weighting factor of each nPDF,  $rel\_w_{i,j,k,m}$ , can be calculated by:

$$\hat{w}_{i,j,k,m} = \frac{w_{i,j,k,m}}{\sum_{i=1}^{n_{j,m}} w_{i,j,k,m}} \quad \text{equation 2}$$

91 where  $w_{i,j,k,m}$  is an absolute measure of the average rate of emission of substance  $k$  over an  
92 arbitrary time frame and over the geographical zone  $i$  or a proxy of that absolute average  
93 rate. Therefore,  $rel\_w_{i,j,k,m}$  is deemed representative of the likelihood that an emission will  
94 occur in a specific geographical zone  $i$  within geographical zone  $j$ .

95 This aggregation procedure is already a topic of interest in risk assessment. While experts  
96 do not agree on the range of values a parameter should take, methods were developed to  
97 reach a compromise. The linear opinion pool can be used to combine the views of many  
98 experts in a single probability distribution and is analogous to equation 1 (Clemen et  
99 Winkler 1999). In the opinion pool, weights represent the quality or trust in the experts. In  
100 the CF aggregation context, the weights represent the likelihood of emission in each cell.

101 There are other mathematical aggregation methods, such as the logarithmic opinion pool  
102 (Clemen et Winkler 1999) in which the weights are used as exponents for each PDF, which  
103 are then multiplied. In section 2.6, we show that this aggregation procedure is ill suited to  
104 capturing the increase in uncertainty. Bayesian approaches are more difficult to apply  
105 because they require information on the experts' precision, bias and dependence. Using a  
106 simpler method requiring less information is preferable if it fulfills the purpose of capturing  
107 the uncertainty increase.

108 The linear opinion pool has already been suggested in LCA. In a streamlined LCA, when only  
109 limited information on an intermediate flow (iron, for example) is available, Olivetti et al.  
110 (Olivetti, Patanavanich et al. 2013) proposed that the statistical distribution of the impact  
111 should be the combination of the impact distribution of the different possibilities (cast iron,  
112 sand casted iron and pig iron).

113 In their work on sensitivity testing in LCA, Mutel et al. use the kernel density estimator  
114 approach to aggregate CFs (Mutel, de Baan et al. 2013), a method that is virtually identical  
115 to the linear opinion pool. In the study?, CFs for damages to biodiversity by cocoa  
116 plantations are estimated using data from cocoa, coffee and other agroforestry practices.  
117 The coffee and agroforestry data sources were attributed lower weights, and the authors  
118 assumed the uncertainty of the CFs.

### 119 **2.3 Weighting datasets**

120 Weighting sets should strive to be indicative of the probability that a given nCF is the  
121 appropriate one to use when spatial information is only available at the aCF level. The basis  
122 on which weighting sets are built can range from very simple (e.g. uniform weighting or  
123 area-based weighting) to very data intensive (e.g. emission-based weighting using total  
124 rates of emission over a certain area or even industry-specific emission-based weighting  
125 sets, leading to industry-specific aCFs). Population density has already been used as an  
126 emission proxy in an aggregation process (Manneh, Margni et al. 2010). Model developers  
127 should carry out sensitivity analyses based on the weighting sets, ideally striking an  
128 acceptable balance between reliability and complexity.

### 129 **2.4 Delineation of LCA and LCIA**

130 Considering current LCI data, the geographic zone  $j$  in equations 1 and 2 will typically be at  
131 a country or continent scale. However, the equations are valid for any combination of LCI  
132 and LCIA spatial resolutions. The method could even accommodate eventual situations in  
133 which the delineation is specific to the elementary flows contributing to a given impact  
134 category.

### 135 **2.5 nCF uncertainty**

136 In their evaluation of LCIA methods, Hauschild et al. included the following criterion  
137 (Hauschild, Goedkoop et al. 2013): CF uncertainty should be reported quantitatively with  
138 statistical distributions. Only the midpoint global warming potential model by the IPCC  
139 provides full quantitative uncertainty estimates. This is not a regional impact category, so it  
140 does not constitute an application example in this paper.

141 Now that explicit criteria have been published for best practices regarding uncertainty in  
142 LCIA, future regionalized models are expected to include this type of information.  
143 Considering the current shortage of suitable models to apply the aggregation procedure,  
144 uncertainty estimates were specifically generated for this purpose (see section 3.3). It will  
145 suffice to illustrate the relevance of the approach.

146 **2.6 Uncertainty increase indicator**

147 Is the aCF more uncertain than the nCF it is based on? To answer the question  
 148 quantitatively, it is necessary to use a relative indicator of the uncertainty, independent of  
 149 the shape of the distribution: the coefficient of variation (CV = standard deviation divided  
 150 by the average). Since nCFs covering very different values may be aggregated, it does not  
 151 make sense to compare the absolute value of their uncertainty: variance or standard  
 152 deviation values attached to greater nCFs will appear greater.

153 The uncertainty increase indicator (UII) is defined as:

$$UII = \frac{CV_{aCFj} - \min_i(CV_{aCFi})}{\max_i(CV_{aCFi}) - \min_i(CV_{aCFi})} \quad \text{equation 3}$$

154 The CV of every nCF to be aggregated is evaluated, and the minimum and maximum CVs are  
 155 determined. Then, the CV of the aCF is calculated from its PDF. The UII is an indication of  
 156 the value of the CV of aCF relative to the minimum and maximum CVs of the nCFs.

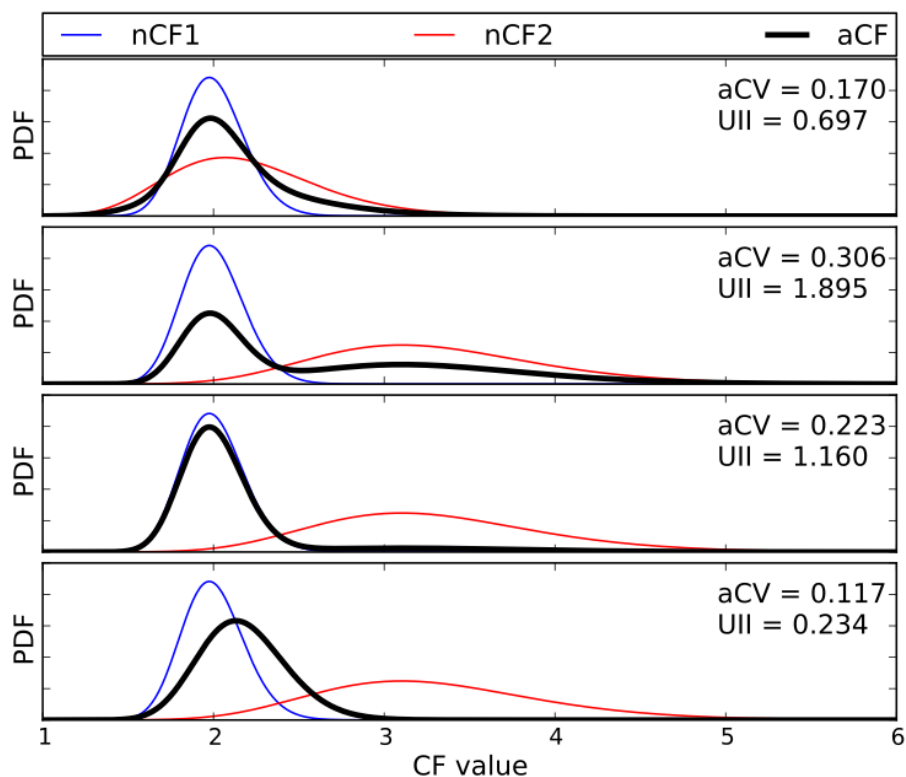
157 Figure 1 shows four fictional examples of two nCFs aggregated into an aCF. In the top  
 158 subplot, the two nPDFs overlap, and the relative weight of both distributions is 50%. The  
 159 mean and shape of the resulting aCF are similar to those of the initial nCFs. The CV of the  
 160 aCF will be between that of the two nCFs.

161 The fact that the aCF can have a lower CV than some of the nCFs from which it was  
 162 calculated may seem counterintuitive, but it is indeed possible. The aggregation process is  
 163 similar to mixing two samples into one larger sample. If both samples have relatively close  
 164 averages, the variance of the larger sample will find itself between the variance of the initial  
 165 samples. This situation leads to an UII less than 1.

166 In the second subplot, the weights are kept at 50%, but there is less overlap between the  
 167 two distributions. The CV of the aCF is then higher than the CV of either nCFs, leading to an  
 168 UII superior to 1. In the third subplot, the aggregation is made with the same nCFs as the  
 169 second subplot but with a weight of 90% for the leftmost distribution. The UII in this case is  
 170 also greater than 1 but lower than in the second case.

171





172

173 **Figure 1: Examples of UII results. CV1 = 0.091, CV2 = 0.204 for all subplots**

174 The fourth subplot starts with the same distributions and weights as the second but uses  
 175 the logarithmic opinion pool mentioned in section 2.2. The resulting distribution does not  
 176 adequately capture the difference between the distributions. Therefore, the UII is lower  
 177 and fails to point to potential inaccuracies created by its use.

178 The UII must be interpreted carefully. Lower UII values indicate that the spatial variability  
 179 of the CFs aggregated over region  $j$  is comparable to the uncertainty of each cell. In this case,  
 180 substituting the aCF by the nCF would not produce a significant reduction in the relative  
 181 uncertainty of the impact scores. A UII value close to 1 means that the uncertainty of the  
 182 aCF is close to the uncertainty of the most uncertain nCF, and a value close to 0 indicates  
 183 that the uncertainty is comparable to the least uncertain nCF. Lower UII values may also  
 184 indicate that the nCF are different but that emissions are less likely to occur in regions with  
 185 extreme nCF values.

186 As the nPDFs become more differentiated, the value of the UII increases, but there is no  
 187 clear qualitative change between  $UII < 1$  and  $UII > 1$ . Higher UII values indicate that, within

188 a geographic zone, spatial variability is significant as compared to the uncertainty of the  
 189 model. The situation could justify supplementary data collection to locate emissions more  
 190 specifically.

## 191 **2.7 Equivalence with bootstrapping method**

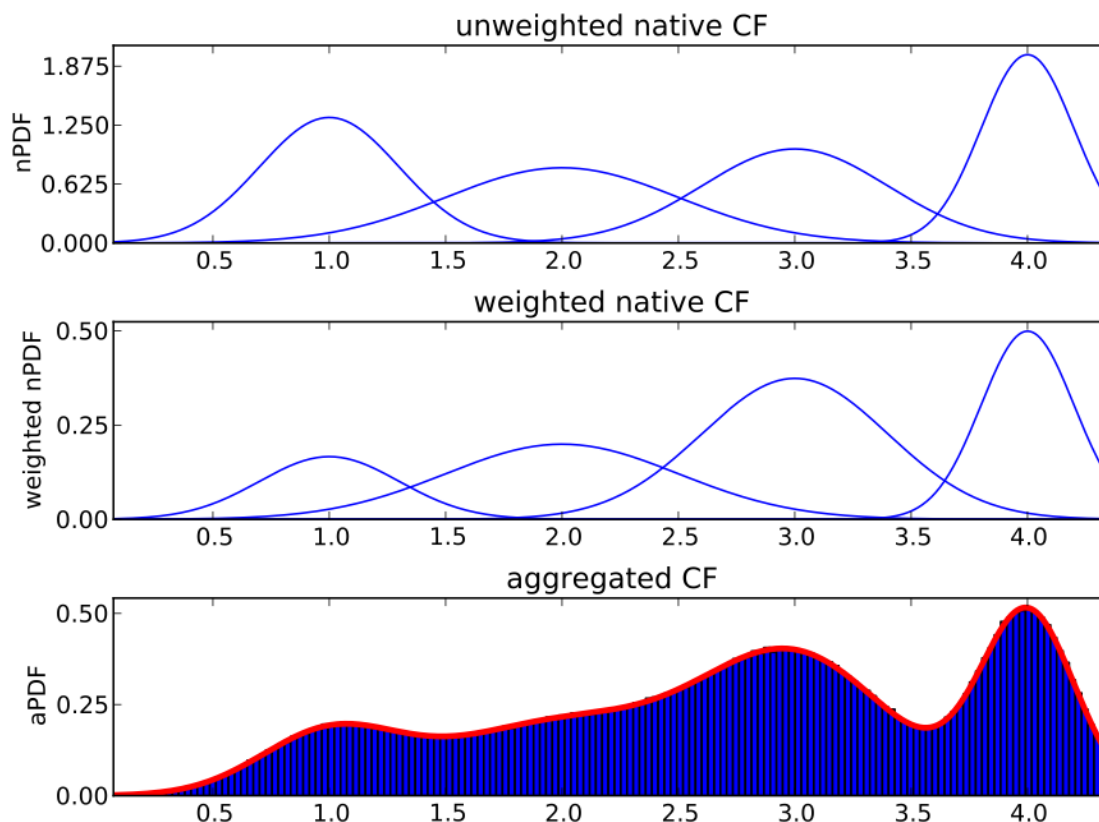
192 A bootstrapping method can also be used to assess the effect of the multiple nCF choices  
 193 (Efron et Tibshirani 1991). Huijbregts et al. used a similar method to evaluate the effect of  
 194 discrete normative choices or scenarios on the uncertainty of an LCA(Huijbregts, Gilijamse  
 195 et al. 2003).

196 For each bootstrapping iteration, the first step is to randomly select an nCF. The likelihood  
 197 of selecting each nCF is equivalent to the relative weight, as defined in equation 2. Then, the  
 198 distribution of the chosen nCF is sampled.

199 **Table 2** shows the data used in a fictional example with four nCFs with normal  
 200 distributions, and the top subplot of Figure 2 displays the PDFs. The second subplot shows  
 201 the PDFs multiplied by their relative weight. The third subplot, which presents the results  
 202 of the bootstrapping approach and the proposed procedure, indicates that both approaches  
 203 are equivalent.

204 **Table 2: nCF characteristics for the bootstrapping example**

Average	Standard deviation	Weight	Relative weight
1	0.3	1	12.5%
2	0.5	2	25.0%
3	0.4	3	37.5%
4	0.2	2	25.0%



205

206 **Figure 2: The PDF of the aCF perfectly fits the distribution of the bootstrapping**  
 207 **simulation. In the third subplot, the red line is the result of the linear opinion pool,**  
 208 **and the histogram is the result of the bootstrapping approach**

### 209 3 Application example

#### 210 3.1 Description of the model

211 The aggregation procedure and the calculation of the UII are demonstrated using a model  
 212 used to calculate the direct potential impacts on human health of freshwater use (Boulay,  
 213 Bulle et al. 2011). Since the focus of this paper is the aggregation of nCFs and not the  
 214 characterization model per se, we only provide the information required to understand the  
 215 aggregation process. For further details on the impact model used in the demonstration,  
 216 please refer to the original publication.

217 The calculated CFs depend on the quality of the water consumed. The quality classification  
 218 depends on chemical and biological criteria and divides water into eight categories. Ground  
 219 water and surface water are also differentiated, and rainwater is considered to be in its own

220 category. The result is 17 water qualities treated as distinct elementary flows, each with its  
221 own nCF.

222 The impact of water use also depends on water scarcity. When there is no scarcity for a  
223 certain type of water, there is no direct impact on human health. If there is scarcity, the  
224 impact depends on the level of competition among potential users in the same watershed.  
225 Scarcity data is defined at a watershed resolution. Scarcity constitutes the midpoint of the  
226 model.

227 The capacity to adapt to a lack of water is modeled as dependent on the gross national  
228 income (GNI) per capita. It is assumed that developed countries will deploy technological  
229 solutions to compensate for a lack of water and therefore suffer no direct human health  
230 impact, whereas, in countries in which GNI per capita is under a specified threshold,  
231 adaptation is partial or non-existent. In the latter cases, direct human health impacts will be  
232 caused by the lack of water for agriculture, aquaculture and human consumption. The GNI  
233 per capita data is defined at a country geographic resolution. Lack of water for each user  
234 ultimately results in endpoint damages quantified in disability-adjusted life years (DALY).  
235 In this paper, the endpoint nCFs, in DALY/m<sup>3</sup>, are used in the aggregation procedure.

### 236 **3.2 Delineation**

237 The resulting delineation is 808 cells of heterogeneous shapes and areas corresponding to  
238 the intersection of national borders and watersheds. For certain countries, the methodology  
239 is unnecessary since a single watershed covers the entire territory. In such cases, the aCF is  
240 equal to the nCF. Larger countries may contain up to 22 watersheds.

### 241 **3.3 Uncertainty of characterization factors**

242 Supporting information on the evaluation of the uncertainty is provided. Thirteen input  
243 parameters were identified as randomly distributed. Their distributions were determined  
244 through expert judgements findings or statistical data. A Monte Carlo simulation was then  
245 performed with a stratified sample of one thousand iterations. A lognormal was fitted to the  
246 results of the Monte Carlo.

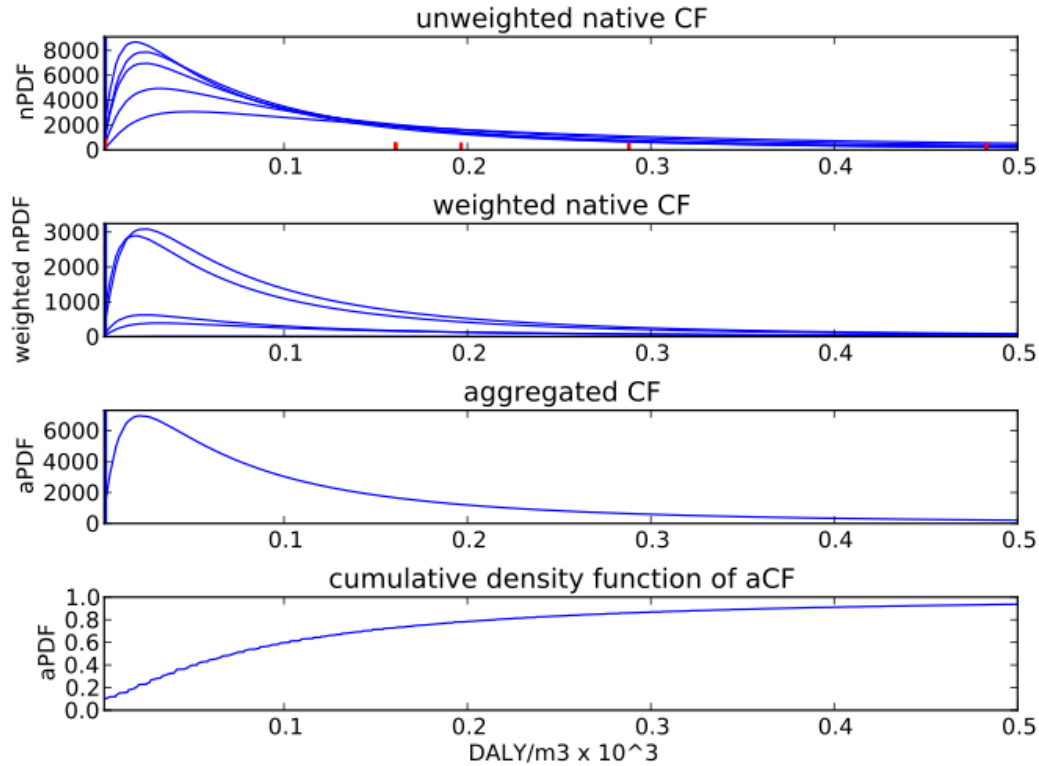
### 247 3.4 Weighting set

248 The WaterGAP 2 global water use and availability model (Alcamo, Döll et al. 2003) provides  
249 data on water withdrawal rates at a global scale. Water withdrawal is deemed  
250 representative of the likelihood of water consumption in every cell. Unfortunately, only the  
251 total water withdrawal is available, without a breakdown by water quality. It is therefore  
252 necessary assume the same weighting set for every water quality.

### 253 3.5 Creation of the aPDF

254 Aggregation was carried out for every country. Ukraine was chosen as geographic zone  $j$  in  
255 this detailed illustration. The geographic zones  $i$  are the six watersheds covering the  
256 country. Figure 3, subplot 1 shows the PDF of the six corresponding nCFs for surface water  
257 quality 2a (*Good*: requiring standard treatment to become drinking water). The red dots  
258 near the x-axis represent the deterministic values of the nCF. One watershed has an nCF of  
259 0 DALY/m<sup>3</sup>, and this value is 100% certain: reliable data indicates no scarcity of S2a water  
260 in the Danube watershed. The PDF is represented by a Dirac impulse function, which is  
261 almost invisible on the plot because it is confounded with the y-axis. Additional data on the  
262 watersheds is detailed in Table 3. GSD2 is the squared geometric standard deviation, a  
263 relative and dimensionless measure of variance for lognormal distributions (Heijungs et  
264 Frischknecht 2005).

265 Figure 3, subplot 2 shows the same set of nPDFs multiplied by their relative weights  
266  $rel\_w_{i,j,k,m}$ . In Figure 3, subplot 3, the weighted nPDFs are added, resulting in the aPDF. The  
267 mode at 0 DALY/m<sup>3</sup> is of infinite height and infinitesimal width, so it is difficult to interpret  
268 the influence and meaning of the mode on the aPDF.



269

270 **Figure 3: Four successive steps in the aggregation process for the six Ukrainian**  
 271 **watersheds and S2a water quality**

272

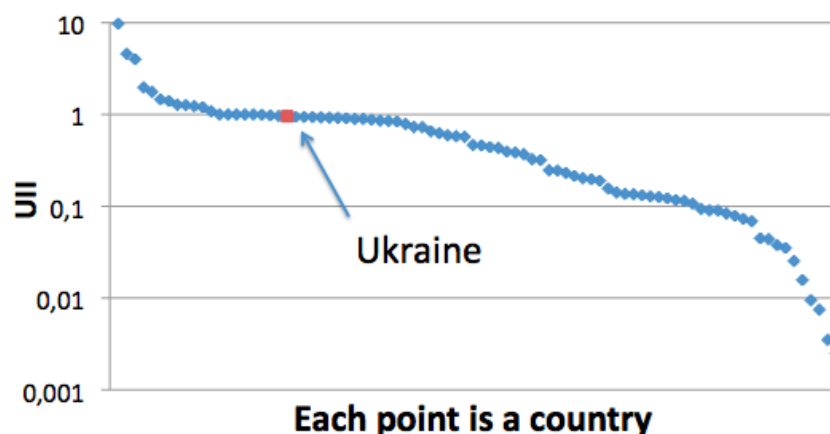
273 **Table 3: Data for the six Ukrainian watersheds for S2a water quality**

Watershed	Statistical distribution	Average (DALY/m <sup>3</sup> )	GSD2	CV	Water withdrawal (km <sup>3</sup> /yr)	Relative weight
Danube	Dirac	0	NA	0	7.94E-04	10.0%
Dnieper	Lognormal	1.61E-04	10.9	1.78	2.65E-03	33.3%
Outside main watershed*	Lognormal	1.61E-04	9.50	1.60	3.12E-03	39.2%
Dniester	Lognormal	1.97E-04	10.6	1.74	7.12E-04	8.9%
Don	Lognormal	2.88E-04	11.2	1.81	6.28E-04	7.9%
Vistula	Lognormal	4.83E-04	11.8	1.89	3.94E-05	0.5%

274 \*This region is a composite region that is not included in a watershed classified as *main watershed*.

275 Examining the cumulative density function (CDF) is more enlightening. As expected, the  
 276 curve in Figure 3, bottom subplot converges to 1 at infinity. The Dirac impulse is the reason  
 277 why the CDF curve does not intersect the y-axis at 0 but at 10.0%.

278 The CV of aCF in Ukraine is 1.75. The minimum and maximum nCF CVs used in the UII  
 279 formula are 0 and 1.89, respectively. The UII for S2a water in Ukraine is 0.85, indicating that  
 280 spatial variability and uncertainty are of comparable magnitude. Carrying out the exercise  
 281 for all countries shows that this is generally the case for S2a water (see Figure 4).



282

283 **Figure 4: Uncertainty increase indicator (UII) for quality S2a water use**

284 High UIIs are associated with countries in which there is a combination of nCFs with low CV  
 285 but with deterministic values spread out over several orders of magnitude. For example, in  
 286 Guinea, which has the highest UII, only one of the 15 watersheds has a water scarcity  
 287 problem, leading to an nCF of zero for 14 watersheds. The average aCF lies between 0 and  
 288 the average nCF of the only watershed with a scarcity problem. This average is not  
 289 representative of the situation for any of the watershed. In other words, a high UII points to  
 290 a lack of spatial information that forces the model into an uneasy compromise. Whether this  
 291 compromise can be accepted in the context of a given LCA will depend on the contribution  
 292 and sensitivity of the flow and the goal and scope of the study.

## 293 4 Discussion

### 294 4.1 Intended use of aCF and UII

295 Ideally, nCF would systematically be used in an LCA. In reality, the spatial resolution of the  
 296 life cycle inventory will almost exclusively restrict the use of nCF to emissions from  
 297 foreground unit processes whose locations can be easily determined, and other elementary  
 298 flows will, by default, be characterized with aCF.

299 Combined with interpretation methods such as contribution and sensitivity analyses, the  
300 UII can help to determine whether it is worth spending resources to better localize certain  
301 unit processes to allow the use of nCF. Scenario analysis in which the study outcome is  
302 calculated using extreme individual nCF could help confirm whether substituting the aCF by  
303 any of the nCFs could actually affect the conclusion of the study.

304 The UII can also be used to guide uncertainty minimization. Certain methods make it  
305 possible to identify major contributors to uncertainty (Huijbregts, Gijlamse et al. 2003;  
306 Mutel, de Baan et al. 2013). If aCFs are identified in this type of analysis, high UIIs can be  
307 used to filter cases in which providing spatial information to choose a specific nCF could  
308 reduce uncertainty more substantially.

#### 309 **4.2 Elementary flow archetypes**

310 When more finely delineated LCIA models with uncertainty become available, the  
311 aggregation procedure may reveal a limited number of markedly differentiated modes in  
312 the aCF. This would be the sign that nCF variability is driven by parameters that are poorly  
313 predicted by spatial proximity. In such situations, it may be advisable to create elementary  
314 flow archetypes.

315 Elementary flow archetypes were proposed without the use of the approaches outlined in  
316 this paper. However, using deterministic nCF values could lead to the erroneous conclusion  
317 that there are clearly differentiated archetypes within a region. The proposed approach,  
318 which enable the consideration of uncertainties, paints a more nuanced picture and help  
319 prevent the expense of resources to create differentiations in LCI that would lose their  
320 relevance due to the uncertainty in LCIA.

321 This highlights the importance of integrating uncertainty assessment from the onset of any  
322 effort aiming to increase the complexity of LCA. A possible criterion in judging the relevance  
323 of new complexity is its ability to increase the discriminating power of LCA. When the  
324 uncertainty of the CF is robustly and systematically estimated, we may observe that it is  
325 greater than the range of values covered by the deterministic nCF values. The example  
326 presented in this paper demonstrates that this may well be the case, since high UII values  
327 are the exception rather than the norm. If this situation proves to be the case in other



328 regional impact categories, the lack of spatial harmonization between LCI and LCIA may not  
329 be of concern after all. The current methodology will still be relevant in the construction of  
330 aCF, but it will show that the uncertainty reduction efforts should be focused on the LCIA  
331 models themselves.

332 The UII could be used as a criterion to determine the flow and country combinations in an  
333 LCI database that must be delineated at a finer spatial scale. Adding regions within  
334 countries, guided by the needs of the LCIA models, would increase the size of the database  
335 only with information that enhances the discriminating power of LCA and avoids a burden  
336 that would make no practical difference on the outcome of an LCA.

### 337 **4.3 Archetype aggregation**

338 In LCIA, reporting multiple CF values for each emission compartment is now a standard  
339 practice. Typical sub-compartments include population density (high or low) and water  
340 types (lake, river, groundwater). New archetypes may be based on other characteristics  
341 deemed relevant, such as stack height, soil type, etc. One should expect a lag between the  
342 identification of these sensitive parameters by LCIA model developers and their eventual  
343 integration into LCI databases. In many cases, the characteristics can remain unspecified  
344 due to the generic nature of the datasets.

345 The situation is analogous to the lack of spatial information and would require the creation  
346 of aggregated CFs with regards to the archetypes. The aggregation could occur using the  
347 approach outlined in this paper as soon as uncertainty information on the CFs is available.  
348 Weighting sets may, however, be more challenging to develop. For example, if an LCIA  
349 model distinguishes between three ranges of stack heights, the relative frequency of each  
350 range could be estimated through the survey of the designers of such devices or through  
351 data collection on a sample of plants. The approaches presented here could be used to  
352 develop a *generic stack height* aCF, and the resulting UII could be used in actual LCA models  
353 to help determine whether it is worth trying to obtain data on actual stack height for  
354 specific unit processes.

## 355 **5 Acknowledgements**

356 The authors acknowledge the industrial partners of the International Chair in Life Cycle  
 357 Assessment (a research unit of CIRAIG [Interuniversity Research Centre for the Life Cycle of  
 358 Products, Processes and Services]) for their financial support: Alcan, Arcelor-Mittal, Bell  
 359 Canada, Cascades, Eco-Entreprises-Québec/Recyc-Québec, Groupe EDF/GDF-SUEZ, Hydro-  
 360 Québec, Johnson & Johnson, Mouvement des caisses Desjardins, RONA, Total and Veolia  
 361 Environnement.

362 We also wish to thank Cécile Bulle, Manuele Margni, Olivier Jolliet and Christopher Mutel  
 363 for the support and input they have provided.

## 364 **6 References**

- 365 Alcamo, J., P. Döll, et al. (2003). "Development and testing of the WaterGAP 2 global model  
 366 of water use and availability." Hydrological Sciences Journal **48**(3): 317-337.
- 367 Boulay, A.-M., C. Bulle, et al. (2011). "Regional Characterization of Freshwater Use in LCA:  
 368 Modeling Direct Impacts on Human Health." Environmental Science & Technology  
 369 **45**(20): 8948-8957.
- 370 Boulay, A.-M., C. Bulle, et al. (2011). "Regional characterization of freshwater use in LCA:  
 371 Modeling direct impacts on human health " Environmental science and technology  
 372 policy analysis **45**(20): 8948-8957.
- 373 Clemen, R. T. and R. L. Winkler (1999). "Combining Probability Distributions From Experts  
 374 in Risk Analysis." Risk Analysis **19**(2): 187-203.
- 375 Efron, B. and R. Tibshirani (1991). "Statistical Data Analysis in the Computer Age." Science  
 376 **253**: 390-395.
- 377 Hauschild, M., M. Goedkoop, et al. (2013). "Identifying best existing practice for  
 378 characterization modeling in life cycle impact assessment." The International Journal  
 379 of Life Cycle Assessment **18**(3): 683-697.
- 380 Hauschild, M., O. Jolliet, et al. (2011). "A bright future for addressing chemical emissions in  
 381 life cycle assessment." The International Journal of Life Cycle Assessment **16**(8):  
 382 697-700.
- 383 Heijungs, R. and R. Frischknecht (2005). "Representing Statistical Distributions for  
 384 Uncertain Parameters in LCA. Relationships between mathematical forms, their  
 385 representation in EcoSpold, and their representation in CMLCA." The International  
 386 Journal of Life Cycle Assessment **10**(4): 248-254.
- 387 Helmes, R., M. Huijbregts, et al. (2012). "Spatially explicit fate factors of phosphorous  
 388 emissions to freshwater at the global scale." International journal of Life Cycle  
 389 Assessment **17**(5): 646-654.
- 390 Huijbregts, M., W. Gilijamse, et al. (2003). "Evaluating Uncertainty in Environmental Life-  
 391 Cycle Assessment. A Case Study Comparing Two Insulation Options for a Dutch One-  
 392 Family Dwelling." Environmental Science & Technology **37**(11): 2600-2608.

- 393 Huijbregts, M., S. Lundi, et al. (2003). "Geographical scenario uncertainty in generic fate and  
394 exposure factors of toxic pollutants for life-cycle impact assessment." Chemosphere  
395 **51**(6): 501-508.
- 396 Kounina, A., M. Margni, et al. (2013). "Review of methods addressing freshwater use in life  
397 cycle inventory and impact assessment." The International Journal of Life Cycle  
398 Assessment **18**(3): 707-721.
- 399 Lenzen, M., K. Kanemoto, et al. (2012). "Mapping the Structure of the World Economy."  
400 Environmental Science & Technology **46**(15): 8374-8381.
- 401 Manneh, R., M. Margni, et al. (2010). "Spatial Variability of Intake Fractions for Canadian  
402 Emission Scenarios: A Comparison between Three Resolution Scales." Environmental  
403 Science & Technology **44**(11): 4217-4224.
- 404 Mutel, C., S. Pfister, et al. (2012). "GIS-Based Regionalized Life Cycle Assessment: How Big Is  
405 Small Enough? Methodology and Case Study of Electricity Generation." Environmental  
406 Science & Technology **46**(2): 1096-1103.
- 407 Mutel, C. L., L. de Baan, et al. (2013). "Two-Step Sensitivity Testing of Parametrized and  
408 Regionalized Life Cycle Assessments: Methodology and Case Study." Environmental  
409 Science & Technology **47**(11): 5660-5667.
- 410 Norris, G. A. (2002). "Impact Characterization in the Tool for the Reduction and Assessment  
411 of Chemical and Other Environmental Impacts." Journal of Industrial Ecology **6**(3-4):  
412 79-101.
- 413 Olivetti, E., S. Patanavanich, et al. (2013). "Exploring the Viability of Probabilistic Under-  
414 Specification To Streamline Life Cycle Assessment." Environmental Science &  
415 Technology **47**(10): 5208-5216.
- 416 Owsianiak, M., R. K. Rosenbaum, et al. (2013). "Addressing Geographic Variability in the  
417 Comparative Toxicity Potential of Copper and Nickel in Soils." Environmental  
418 Science & Technology **47**(7): 3241-3250.
- 419 Posch, M., J. Seppälä, et al. (2008). "The role of atmospheric dispersion models and  
420 ecosystem sensitivity in the determination of characterisation factors for acidifying  
421 and eutrophying emissions in LCIA." The International Journal of Life Cycle  
422 Assessment **13**(6): 477-486.
- 423 Potting, J., W. Schöpp, et al. (1998). "Site-dependent life-cycle impact assessment of  
424 acidification." Journal of Industrial Ecology **2**(2): 63-87.
- 425 Rack, M., S. Valdivia, et al. (2013). "Life Cycle Impact Assessment—where we are, trends,  
426 and next steps: a late report from a UNEP/SETAC Life Cycle Initiative workshop and  
427 a few updates from recent developments." The International Journal of Life Cycle  
428 Assessment **18**(7): 1413-1420.
- 429 Roy, P.-O., L. Deschênes, et al. (2012). "Life Cycle Impact Assessment of Terrestrial  
430 Acidification: Modeling Spatially Explicit Soil Sensitivity at the Global Scale." Environmental  
431 Science & Technology **46**(15): 8270-8278.
- 432 Saad, R., M. Margni, et al. (2011). "Assessment of land use impacts on soil ecological  
433 functions: development of spatially differentiated characterization factors within a  
434 Canadian context." International journal of Life Cycle Assessment **16**(3): 198-211.
- 435 Struijs, J., A. Beusen, et al. (2011). "Characterization factors for inland water eutrophication  
436 at the damage level in life cycle impact assessment." The International Journal of Life  
437 Cycle Assessment **16**(1): 59-64.

- 438 Verones, F., M. M. Hanafiah, et al. (2012). "Characterization Factors for Thermal Pollution in  
439 Freshwater Aquatic Environments." Environ. Sci. Technol **44**(24): 9364-9369.
- 440 Weidema, B., C. Bauer, et al. (2013). Overview and Methodology: Data Quality Guidelines for  
441 the Ecoinvent Database Version 3. Dübendorf, Switzerland, Ecoinvent Center.
- 442

1 Supporting information for “**Quantification of the increase in characterization factor**  
2 **uncertainty due to spatial variability**”

3 **1 Introduction**

4 The purpose of this supporting material is to show the detailed calculation and assumptions  
5 used to obtain the uncertainty on the endpoint characterization factors (CF) of the direct  
6 impact of water consumption on human health of Impact World+<sup>1</sup>. The aim was to generate  
7 plausible nPDF curves for the demonstration of the reaggregation procedure. While it was  
8 necessary to rely on expert judgment for estimating the uncertainty of some underlying  
9 parameters, we believe the use of a rigorous process led to realistic uncertainty estimates.

10 **2 Overview of the model**

11 **2.1 Midpoint characterization factor**

12 **2.1.1 Water scarcity**

13 Water scarcity ( $\alpha_i^*$ ) in a given watershed is calculated with the following parameters:

14 **Table 1: Parameter for water scarcity calculation**

Symbol	Name	Unit	Note
CU	Consumed water	km <sup>3</sup> /yr	The consumptive use calculated by the WaterGap model <sup>2</sup>
$f_g$	Fraction of usage dependent on groundwater	[-]	Obtained from the WaterGap model
Q90	Statistical low flow	km <sup>3</sup> /yr	Flow exceeded 9 months out of 10
$P_i$	Proportion of available water of category $i$	[-]	Calculated by the impact category developers <sup>1</sup>
GWR	Groundwater resource availability	km <sup>3</sup> /yr	Obtained from the WaterGap model <sup>2</sup>

$$\alpha_{surface,i}^* = \frac{CU \cdot (1 - f_g)}{Q90} \cdot \frac{1}{P_i} \quad \text{equation 1}$$

$$\alpha_{ground,i}^* = \frac{CU \cdot f_g}{GWR} \cdot \frac{1}{P_i} \quad \text{equation 2}$$

16 Water categories range from 1 (excellent) to 5 (unusable except for transport or  
 17 hydropower), and are based on bacteriologic and toxicological criteria. Intermediate  
 18 categories are 2a, 2b, 2c, 2d, 3 and 4. Each of the 8 categories can be found as surface water  
 19 or groundwater. The proportion of each water category in a watershed ( $P_i$ ) is determined  
 20 by an upper threshold. This means that within the same watershed, the sum of the  $P_i$  over  
 21 all categories will be higher than 1. Table 2 shows that for example, 3% of the water meets  
 22 or exceeds the criteria of water quality 2b, 20% of the water meets or exceeds the criteria of  
 23 water quality 2c, etc.

24 **Table 2: fictional example of  $P_i$  distribution in a cell**

Water quality	$P_i$
1	0.01
2a	0.03
2b	0.03
2c	0.2
2d	0.3
3	0.5
4	0.8
5	1

25 Water scarcity is an indication of how much water is consumed relative to what is naturally  
 26 replenished within a watershed. This quantity exceeds 1 in certain watershed.

### 27 2.1.2 Stress index

28 The water scarcity is based on consumption-to-availability ratios. The last step toward the  
 29 midpoint is to transform the water scarcity into a stress index ( $\alpha_i$ ), based on a withdrawals-  
 30 to-availability ratios. The correlation between the two quantities is based on a regression  
 31 and follows an s-curve with the equation:

$$\alpha_i = \frac{K}{(1 + (Q \cdot \exp(-B(\alpha_i^* - M)))^{1/v}} \quad \text{equation 3}$$

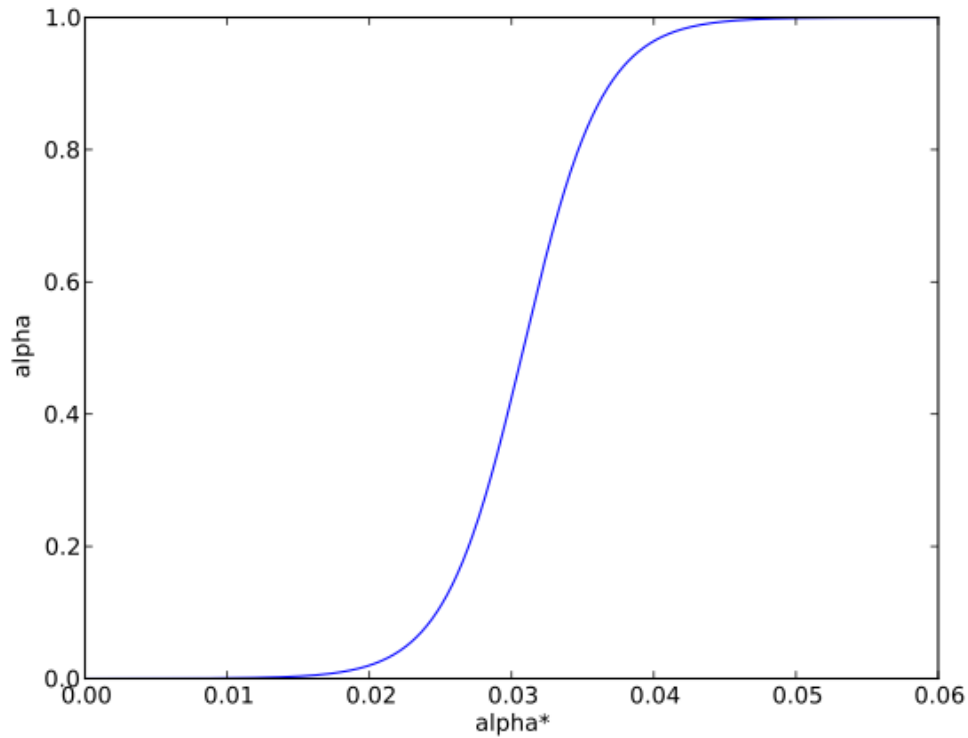
32 When  $\alpha_i^*$  is below 0.00977 or above 0.0548, the value of the stress index  $\alpha_i$  is forced to 0 or  
 33 1, respectively. The value of the midpoint CF is necessarily between 0 and 1.

34 Table 3 shows the parameter values for equation 3:

35

**Table 3:  $\alpha^*_i$  to  $\alpha_i$  correlation parameters**

Parameter	value
K	1
Q	0.3065
B	0.5683
M	2.2
v	0.005498



36

37

**Figure 1:  $\alpha^*_i$  to  $\alpha_i$  correlation****38 2.2 Endpoint characterization factor**

39 The passage from midpoint to endpoint takes into account three other parameters: the  
40 adaptation capacity (AC), the user distribution ( $U_{i,j}$ ) and an effect factor ( $EF_j$ ):

$$CF_i = \alpha_i \cdot (1 - AC) \cdot \sum_{j=1}^{10} U_{i,j} \cdot EF_j \quad \text{equation 4}$$

41 The subscript  $j$  refers to the 10 different types of users potentially affected by a lack of  
42 water.

### 43 **2.2.1 Adaptation capacity**

44 Because water is an essential resource for human life, it will be compensated before any  
45 impact on human health is felt by the population. Compensation will occur through import  
46 or treatment of lower quality water, if available. In that case, no direct human health  
47 impact is felt; only indirect impacts, outside the scope of this work, will be caused by the  
48 compensation activities themselves.

49 In high income countries, defined here as those where GNI (Gross national income per  
50 capita) is above 11 455 \$/cap, compensation is assumed to always occur, and the AC is set  
51 to 1. Direct human health impact of water consumption is therefore null for those  
52 countries, regardless of water availability. In low-income countries, where GNI is below  
53 936 \$/cap, financial resources are assumed unavailable for compensation, and AC is set to  
54 0. Middle-income countries have an AC value based on a linear interpolation between the  
55 two thresholds.

### 56 **2.2.2 User distribution**

57 The user distribution represents the proportion of water consumed by user  $j$ . For direct  
58 human health impact, users are agriculture, domestic, or fisheries. The other users  
59 (industrial, cooling, hydropower and recreational) do not have a direct human health  
60 impact when deprived of water. The distribution is different for each water quality within a  
61 watershed. The sum of  $U_{i,j}$  over all users is therefore equal to 1. Data is based on the  
62 Watergap model<sup>2</sup>.

### 63 **2.2.3 Effect factors**

64 Effect factors represent the impact, in disability adjusted life year (DALY), of the lack of one  
65 cubic meter of water on the human population. For domestic use<sup>3</sup>, the EF is 0.00311  
66 DALY/m<sup>3</sup>. For fisheries and agriculture, each missing cubic meter means a certain amount  
67 of calories not produced for human consumption, and each of those translates into DALY  
68 caused by malnutrition. Those values are 6.17E-3 m<sup>3</sup>/kcal and 1.25E-3 m<sup>3</sup>/kcal,  
69 respectively<sup>4</sup>, and the malnutrition effect is 1.246E-7 DALY/kcal<sup>5</sup>. For fisheries, the effect is  
70 direct, i.e. a deficit of one calorie in fisheries equals a one calorie deficit for a human being.  
71 For agriculture, the portion of agricultural calories fed to animals, which are then fed to



72 humans, had to be taken into account. The proportion retained was 40%<sup>6</sup>. Then, the  
 73 number of calories necessary to feed animals to produce one calorie for human was  
 74 evaluated to 7.22<sup>7</sup>. The effect factor for agriculture and fisheries are therefore:

$$EF_{fisheries} = \frac{1.246 \cdot 10^{-7} \text{ DALY / kcal}}{6.17 \cdot 10^{-3} \text{ m}^3 / \text{kcal}} = 2.02 \cdot 10^{-5} \text{ DALY / m}^3 \quad \text{equation 5}$$

$$EF_{agri,uncor.} = \frac{1.246 \cdot 10^{-7} \text{ DALY / kcal}}{1.25 \cdot 10^{-3} \text{ m}^3 / \text{kcal}} = 9.97 \cdot 10^{-5} \text{ DALY / m}^3 \quad \text{equation 6}$$

$$EF_{agri} = (1 - 0.4)EF_{agri,uncor.} + \frac{0.4EF_{agri,uncor.}}{7.22} = 6.53 \cdot 10^{-5} \text{ DALY / m}^3 \quad \text{equation 7}$$

### 75 3 Uncertainty on input parameters

#### 76 3.1 Water scarcity

77 The CU and Q90 were determined by expert judgment to be lognormally distributed with a  
 78 GSD2 of 2 and 1.1, respectively. The mean of those parameters depend on the cell.

79 There was no indication about the uncertainty of  $P_i$  parameter. This parameter is bound  
 80 between 0 and 1, so the beta distribution was retained. This function accepts four  
 81 parameters: a minimum, a maximum, and two shape parameters, a and b. When  $a = b$ , the  
 82 distribution is symmetrical. The a parameter determines the shape of the left side of the  
 83 distribution, and b, the right side. When  $a > b$ , the average is skewed to the right, and vice  
 84 versa. When  $a < 1$ , a mode appears at the minimum, and when  $b < 1$ , a mode appears at the  
 85 maximum. This situation was avoided by forcing both a and b above 1.

86 The beta was retained in this particular case for its flexibility. The default interval for  $P_i$   
 87 was chosen to be  $[P_i - 0.1, P_i + 0.1]$ . In the cases where  $P_i$  was between 0.1 and 0.9, a and b  
 88 were fixed to 2, and the distribution was symmetrical and bell shaped. If  $P_i$  was lower than  
 89 0.1, the minimum was set to zero, and the maximum to  $P_i + 0.1$ . The parameter a was set to  
 90 2, and the parameter b was adjusted until the average of the now asymmetric distribution  
 91 was equal to  $P_i$ . The symmetrical procedure was followed if  $P_i$  was above 0.9.

92 Another constraint had to be taken into account in the determination of the  $P_i$  distribution.  
 93 Values of  $P_i$  for each water quality within the same cell can vary independently, but as  
 94 shown in Table 2,  $P_1 < P_{2a} < P_{2b} < \dots$ . If  $P_{2a} - P_1 < 0.2$ , their distribution would overlap,  
 95 creating a non zero probability that  $P_{2a} < P_1$ , which we know is not possible. In this  
 96 situation,  $\max(P_1)$  and  $\min(P_{2a})$  were both set to  $(P_1 + P_{2a}) / 2$ . Each pair of  $P_i$  was tested for  
 97 potential adjustment. When boundaries were determined for all  $P_i$ , the values of a and/or b  
 98 were adjusted to restore the average of each distribution to the deterministic  $P_i$  value.

### 99 3.2 Stress index

100 The  $\alpha^*_i$  to  $\alpha_i$  correlation presented another challenge. The equation of the correlation was  
 101 deemed uncertain, but all the parameters, calculated with a regression, represent no  
 102 physical properties that would allow for an expert judgment on their range. It was decided  
 103 that for each value of  $\alpha^*_i$ , there would be a beta distribution of possible associated  $\alpha_i$ . The  
 104 first step was to establish a rule for the upper and lower bounds of the distributions.

105 Expert judgment determined that for  $\alpha^*_i > 0.11$ ,  $\alpha_i = 1$ , and  $\alpha^*_i < 0.01$ ,  $\alpha_i = 0$ . There was no  
 106 uncertainty within those zones. Values of  $\alpha_i$  always remain bound between 0 and 1. In the  
 107 [0.01, 0.11] interval, two linear relationships were established, representing potential  
 108 candidate for maximum and minimum of beta distributions (dashed lines on Figure 2).  
 109 Then, for each value of  $\alpha^*_i$ , the retained maximum and minimum were determined  
 110 according the following relationships:

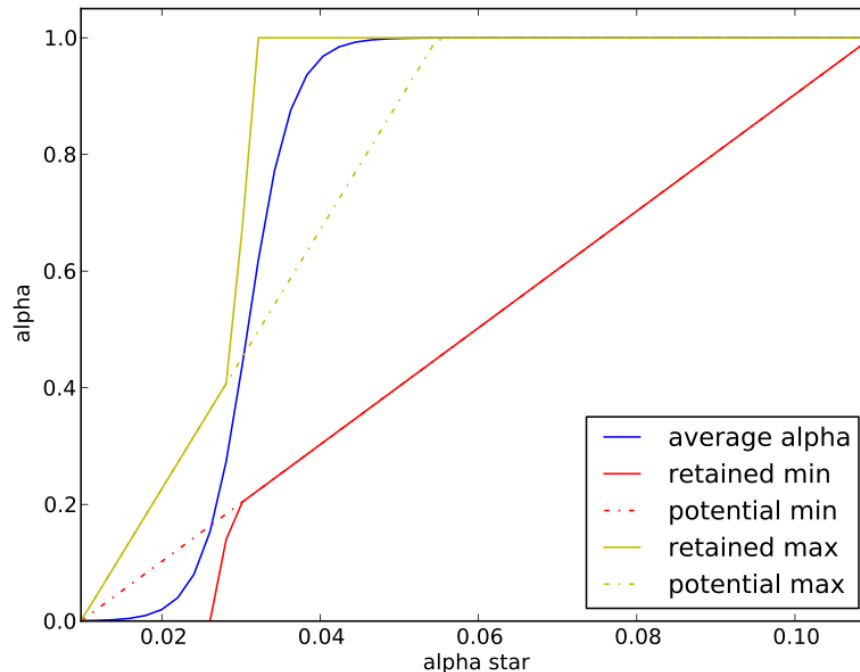
$$\Delta = \max(\text{abs}(\text{average} - \text{potential\_min}), \text{abs}(\text{average} - \text{potential\_max})) \quad \text{Equation 8}$$

$$\text{retained\_max} = \min(\text{average} + \Delta, 1) \quad \text{Equation 9}$$

$$\text{retained\_min} = \max(\text{average} - \Delta, 0) \quad \text{Equation 10}$$

111  
 112 In the Monte Carlo simulation, for each iteration, an  $\alpha^*_i$  has been generated, and a beta  
 113 distribution was fitted to the average, minimum and maximum with the same procedure  
 114 described earlier. It is then sampled once to generate the  $\alpha_i$ , and the calculation sequence is  
 115 carried on towards the endpoint.

116



117

118 **Figure 2: average, minimum and maximum bounds for beta distributions in the  $\alpha_i^*$  to**  
 119  **$\alpha_i$  correlation**

### 120 3.3 User distribution

121 Two constraints had to be taken into account simultaneously: each  $U_{i,j}$  is bound between 0  
 122 and 1, and the sum of  $U_{i,j}$  over  $j$  (all users within a cell) should be equal to 1. The beta  
 123 distribution was retained for the user distribution uncertainty. The distributions were  
 124 given the arbitrary bounds  $[\max(U_{i,j} - 0.1, 0), \min(U_{i,j} + 0.1, 1)]$ , and fitted as previously  
 125 explained with other beta distributed parameters. Within a cell, all the  $U_{i,j}$  were sampled  
 126 independently. If their sum was outside the  $[0.99, 1.01]$  interval, the iteration was rejected.  
 127 New iterations were generated until 1000 of them were accepted.

### 128 3.4 Effect factor

129 Table 4 shows the distribution types and parameters used in the calculation of the  
 130 uncertainty of the effect factor. Malnutrition<sup>5</sup> and domestic<sup>3</sup> water shortage were based on  
 131 more detailed data, disaggregated at the country level. Instead of using a different value for  
 132 each country during the calculation, it was decided to use a single, uncertain value, with a  
 133 lognormal distribution.

134 Less reliable data were available for the water requirement for fisheries and agriculture.  
 135 Only a range of maximum and minimum values were available<sup>4</sup>. In order to stay  
 136 conservative, a uniform distribution was assumed between those two extreme values.

137 The portion of agricultural calories fed to animals is a percentage bound between 0 and  
 138 100%. To avoid non-zero probability density outside this interval, the beta function was  
 139 chosen. In order to get a symmetrical distribution, the parameter a and b were forced to  
 140 the same value of 2, yielding a bell shaped distribution. Data indicated a range of vegetal  
 141 calories per animal calories between 4 and 11<sup>7</sup>. A uniform distribution was also used.

142 **Table 4: Uncertainty on parameters of the effect factor**

Parameter	Type of distribution	Parameter(s) of the distribution	Note
Malnutrition (DALY/kcal)	lognormal	average = 1.25E-07 DALY/kcal GSD2 = 5.39 [-]	
Water requirement, fisheries (m <sup>3</sup> /kcal)	uniform	min = 5.71E-04 m <sup>3</sup> /kcal max = 1.18E-02 m <sup>3</sup> /kcal	
Water requirement, agriculture (m <sup>3</sup> /kcal)	uniform	min = 7.14E-04 m <sup>3</sup> /kcal max = 1.79E-03 m <sup>3</sup> /kcal	
Domestic water shortage (DALY/kcal)	lognormal	average = 3.11E-03 DALY/kcal GSD2 = 12.9 [-]	
Vegetal calories per animal calories [-]	uniform	min = 4.0 [-] max = 11.0 [-]	
Proportion of agricultural calories fed to livestock [-]	beta	min = 30% max = 50% a = 2 b = 2	expert judgment

#### 143 **4 Results**

144 A Monte Carlo was carried on with stratified sampling: percentage point functions were  
 145 calculated for each distribution at [0.0005, 0.0015, ..., 0.9985, 0.9995] (a total of 1000  
 146 iterations) and were then randomly shuffled. A routine was coded in Python with the help  
 147 of the Numpy, Scipy and Random packages. The routine calculated the endpoint CF for each  
 148 cell, water quality and iteration.

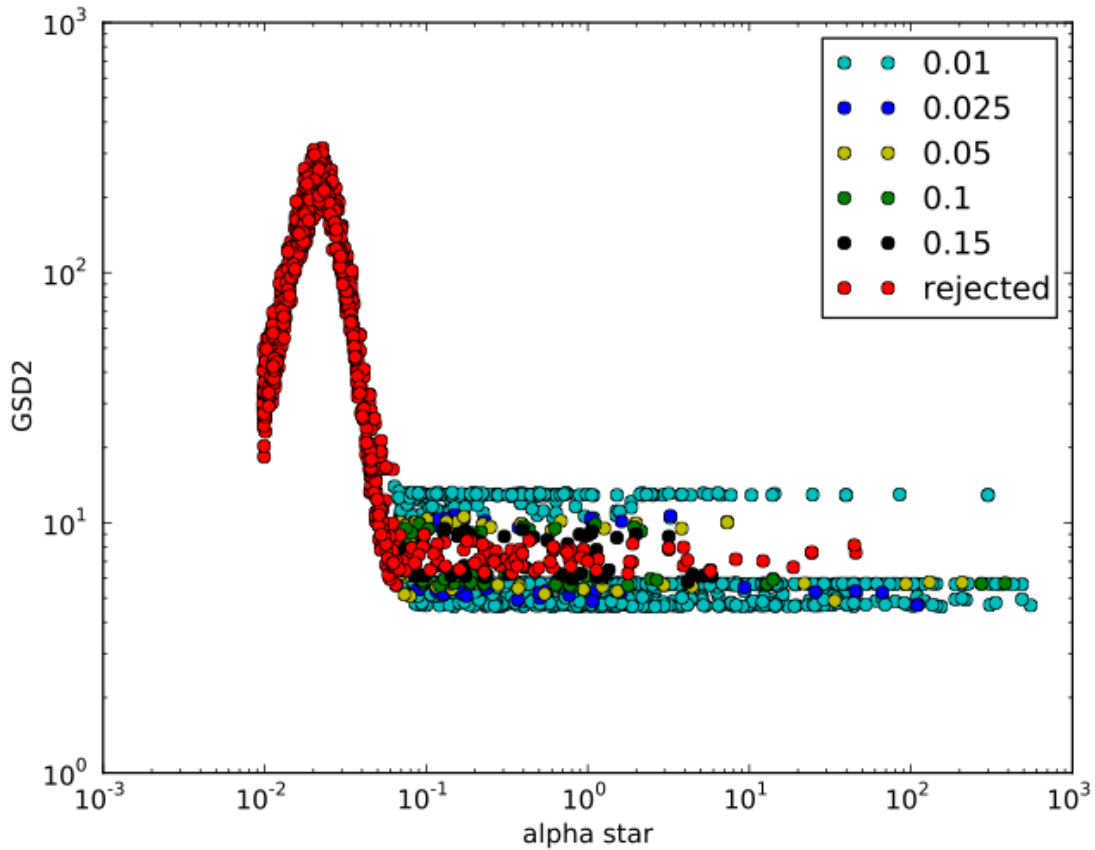
149 Distributions of CF were highly skewed, and since many parameters were represented with  
 150 a lognormal distribution, the first attempt was to test for lognormality. An Anderson-  
 151 Darling test was performed on the natural logarithm of the CF results, for normality. The  
 152 GSD2 of the lognormal was calculated with the variance of the sample.

153 Out of the 12 928 CF of the model, 10 797 were either “NA” (no water of this quality  
 154 available in this cell) or equal to zero because of absence of scarcity, or full adaptation  
 155 capacity of high-income countries. Table 5 shows the breakdown of CF according to the  
 156 level of significance of the Anderson-Darling test, and the minimum, average and maximum  
 157 GSD2 of the associated lognormal.

158 **Table 5: Anderson-Darling test results for lognormality**

<b>Significance level</b>	<b># of CF</b>	<b>Minimum GSD2</b>	<b>Average GSD2</b>	<b>Maximum GSD2</b>
1%	1132	4.62	6.59	13.96
2.5%	40	4.68	7.12	10.75
5%	59	4.88	7.34	10.57
10%	54	5.56	7.06	9.91
15%	65	5.83	11.04	233.94
Rejected (>15%)	781	5.61	89.27	315.22
<b>Total</b>	<b>2131</b>			

159  
 160 Exploration revealed an interesting relationship between the spread of the CFs and the  
 161 value of the intermediate variable  $\alpha^*_i$ . Figure 3 shows the plot of the calculated GSD2  
 162 against the deterministic value of  $\alpha^*_i$  for all 2131 CF. The color of the markers refers to the  
 163 significance level of the Anderson-Darling test.

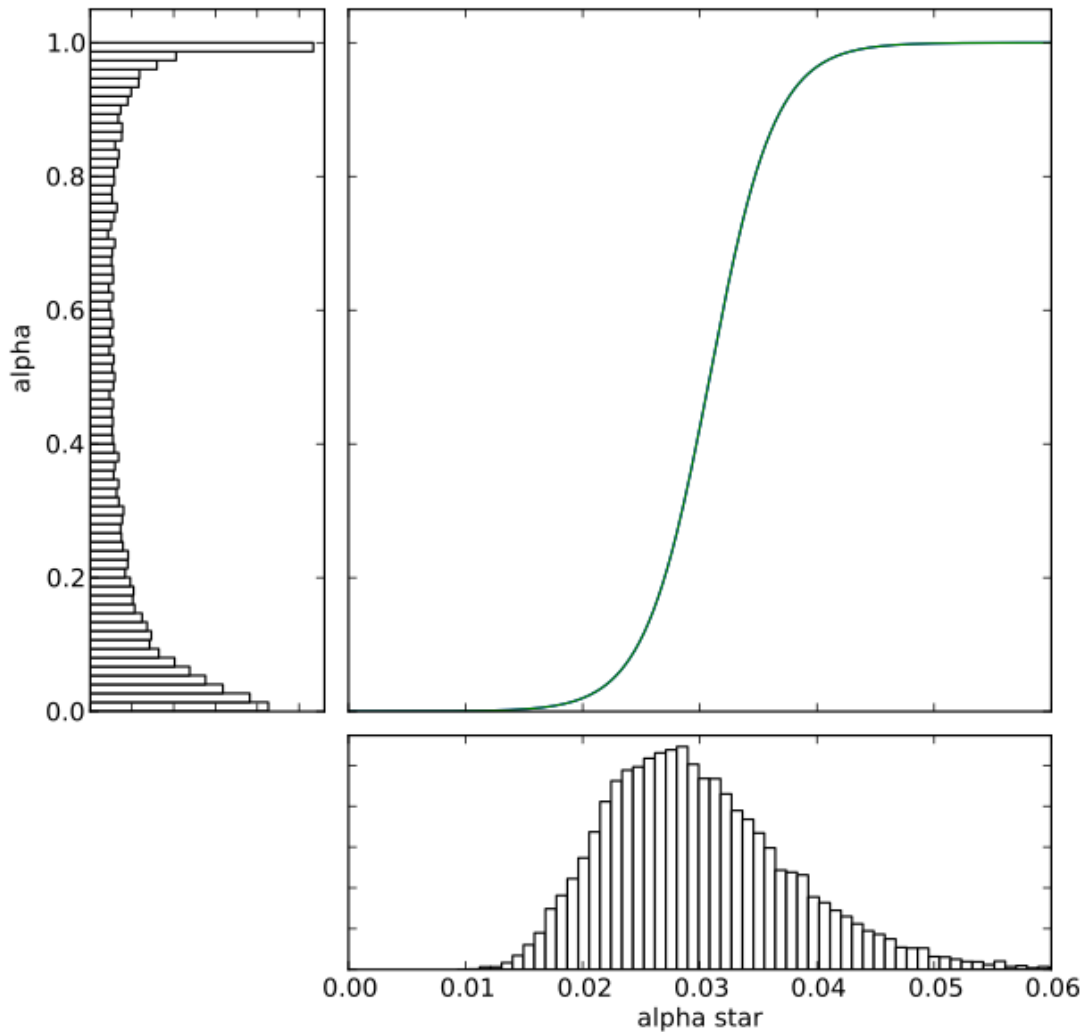


164

165

**Figure 3: spread VS deterministic  $\alpha^*_i$  values for all water qualities**

166 This pattern appeared because of the steep increase in the  $\alpha^*_i$  to  $\alpha_i$  correlation. If  $\alpha^*_i$   
 167 happened to be in the transition zone, the resulting distribution of  $\alpha_i$  became bimodal, with  
 168 modes at 0 and 1 (see Figure 4), or, at the very least, skewed. This is the most likely cause  
 169 of the two effects observed in the transition zone: increase spread, and inability to  
 170 satisfactorily fit a lognormal at the endpoint CF.



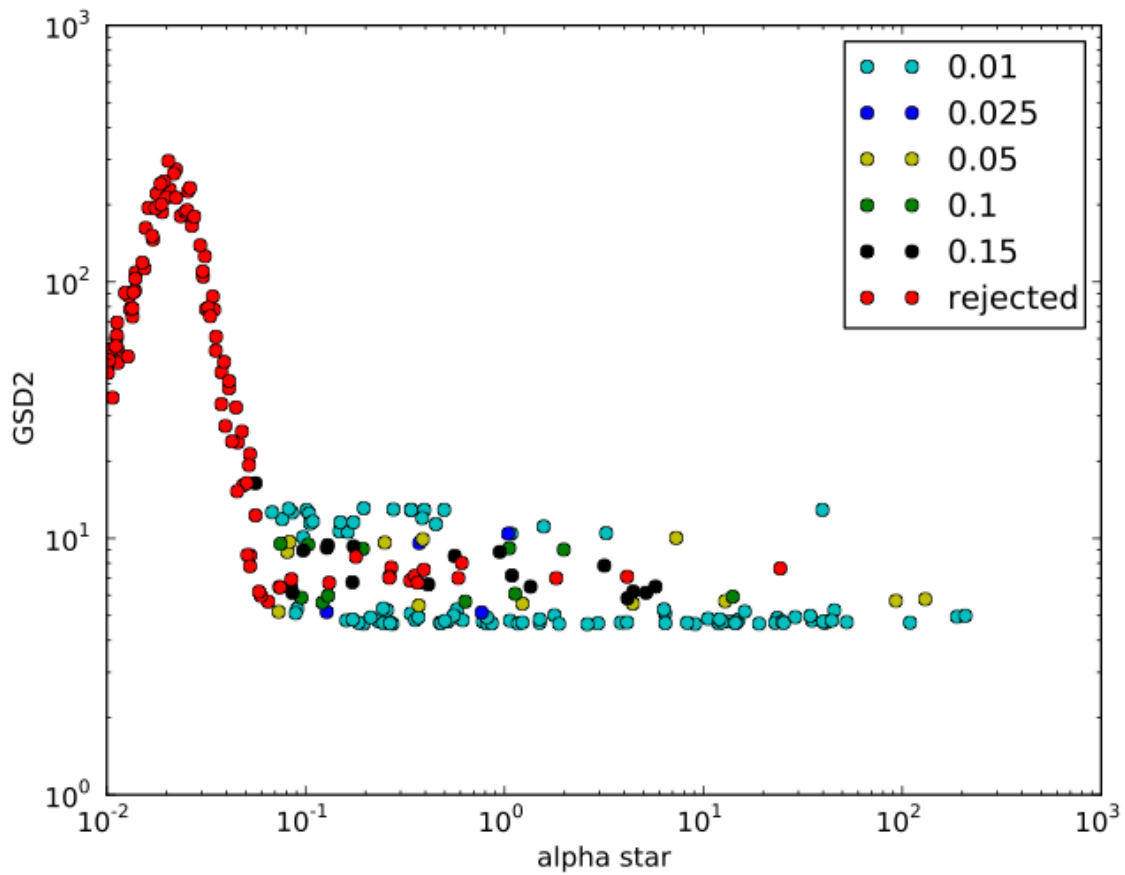
171

172 **Figure 4: distortion caused by the  $\alpha^*_i$  to  $\alpha_i$  correlation, when  $\alpha^*_i$  is in the transition**  
 173 **zone**

174 When  $\alpha^*_i$  is larger than the upper boundary of the transition zone, all uncertainty disappear,  
 175 as  $\alpha_i$  becomes “flattened” to the maximum value of 1. The uncertainty on the endpoint CF is  
 176 then entirely controlled by uncertainty of effect factors and user distribution.

177 Horizontal stripes in Figure 3 also seemed to indicate a pattern. To discover where it came  
 178 from, the same plot was separately created for each of the water quality. Two distinct type  
 179 of graphs appeared. The first pattern appeared with water of quality 1, 2a and 2b, as in  
 180 Figure 5; the other pattern appeared with water quality 2c, 2d, 3, 4 and 5, as in Figure 6.  
 181 Higher quality of water is consumed by all possible users, hence, the larger range of GSD2  
 182 values above the transition zone. Lower quality water is only used for agriculture, except

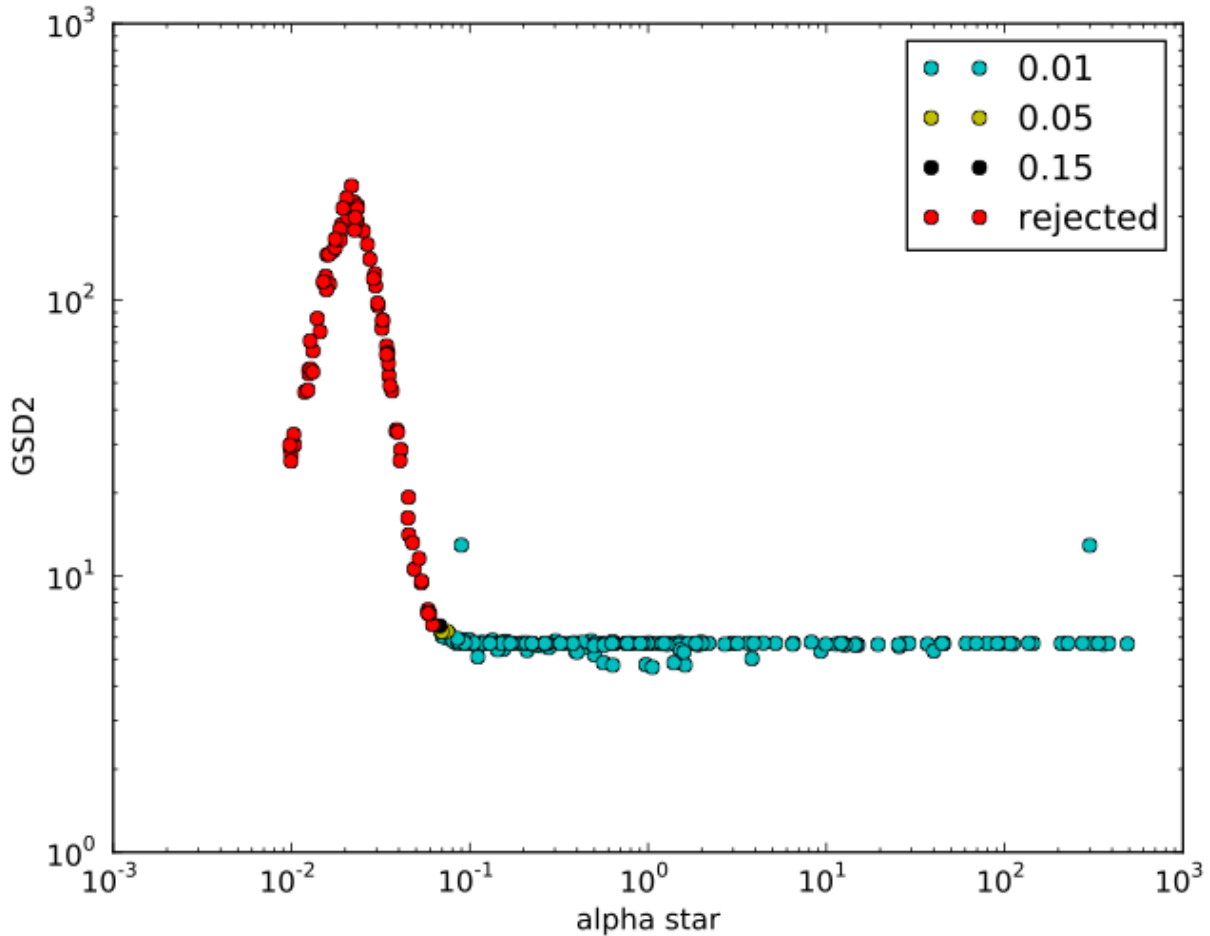
183 for 2d, also used in fisheries. The uncertainty of the endpoint CF is therefore controlled by a  
184 narrow possibility of user distribution.



185

186 **Figure 5: spread VS deterministic  $\alpha^*_i$  values for S2b water quality. More diverse user**  
187 **distribution yields a larger range of GSD2**





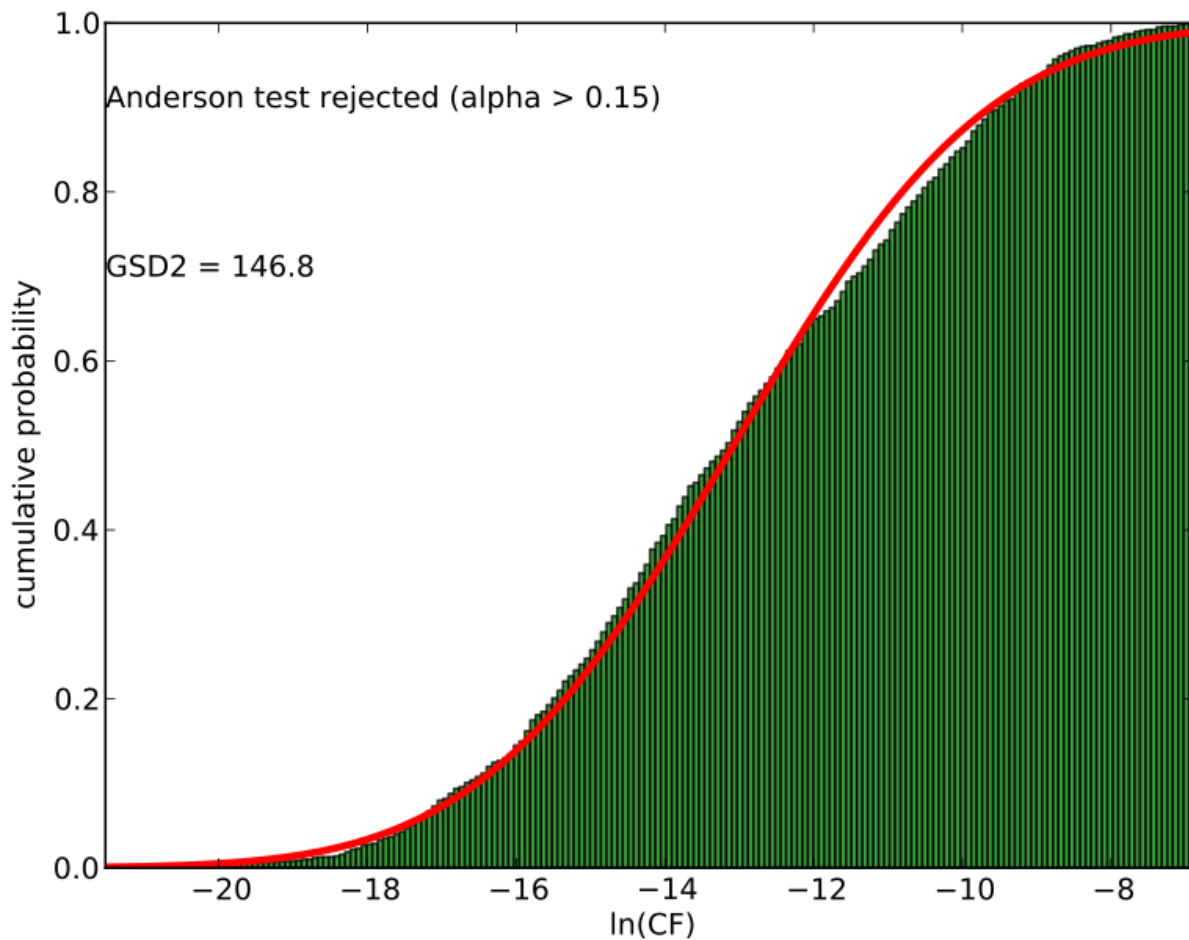
188

189

**Figure 6: spread VS deterministic  $\alpha^*_i$  values for S2c water quality**

190 Visual inspection of rejected Anderson-Darling tests revealed that despite the test result,  
 191 the lognormal best fit was not a drastically bad approximation, as shown by Figure 7. Since  
 192 the  $\alpha^*_i$  to  $\alpha_i$  correlation distorts distributions in an unpredictable way, that user distribution  
 193 is beta shaped, and that effect factors are lognormal, we do not expect the CF in the  
 194 transition zone to comfortably fit any classically used distributions. In the interest of  
 195 simplicity, uniformity and comparability of results across the model, we decided to keep the  
 196 lognormal distribution for the reaggregation procedure.

197



198

199

200

**Figure 7: cumulative probability curve of best lognormal fit and empirical distribution function for G2a water in cell 18 (Bolivia)**

## 201 5 References

- 202 1. Boulay, A.-M.; Bulle, C.; Bayard, J.-B.; Deschênes, L.; Margni, M., Regional  
 203 Characterization of Freshwater Use in LCA: Modeling Direct Impacts on Human Health.  
 204 *Environ. Sci. Technol* **2011**, *45* (20), 8948–8957.
- 205 2. Alcamo, J.; Döll, P.; Henrichs, T.; Kaspar, F.; Lehner, B.; Rösch, T.; Siebert, S.,  
 206 Development and testing of the WaterGAP 2 global model of water use and availability.  
 207 *Hydrological Sciences Journal* **2003**, *48* (3), 317–337.
- 208 3. (a) Howard, G.; Bartram, J. *Domestic Water Quantity, Service Level and Health*; World  
 209 Health Organization: 2003, 2003; p 33; (b) Aguilar-Manjarrez, J., WRI Major watersheds of  
 210 the world delineation. **2006**; (c) FAO, Aquastat. 2009.
- 211 4. *The United Nations World Water Development Report 3: Water in a Changing World*;  
 212 World Water Assessment Program: 2009.

- 213 5. (a) *Mortality and burden of disease estimates for WHO member states in 2004*; World  
214 Health Organization: Geneva, Switzerland, 2009; (b) *FAO Depth of Hunger*; 2009; (c) *FAO*  
215 *Prevalence of undernourishment in total population*; 2008.
- 216 6. Nellesmann, C.; MacDevette, M.; Marders, T.; Eickhout, B.; Svihus, B.; Prins, A. G.;  
217 Kaltenborn, B. P. *The environmental food crisis – The environment's role in averting future*  
218 *food crisis*; United Nation Environmental programme, GRID-Arendal: 2009.
- 219 7. Collomb, P. *Une voie étroite pour la sécurité alimentaire d'ici à 2050*; FAO: 1999.

220

221

**ANNEXE 2 – Article 2 : « Statistical hypothesis testing tool applied to LCA: Judging the significance of differences»**

## **Statistical hypothesis testing tool applied to LCA: Judging the significance of differences**

Guillaume Bourgault, Pascal Lesage, Réjean Samson

Keywords: uncertainty, decision making, model complexity, interpretation, statistical testing

*Summary* In LCA, it is often necessary to determine whether certain values can be considered significantly distinct given their uncertainty. While the most common case determines whether the impact scores of two product systems can be considered sufficiently different to influence decision making, other examples abound in life cycle inventory modeling (e.g. whether two production routes are sufficiently different to create two datasets) and impact assessment (e.g. whether the impact of a pollutant released to different sub-compartments is sufficiently different to justify creating more than one characterization factor (CF)). The concept of statistical power (SP), which expresses confidence in the assertion that two values are, in fact, different, can prove useful.

This paper describes how to calculate SP and two examples of the use of SP in LCA. The first example pertains to the most frequent situation: distinguishing two product systems based on their respective impact scores. Based on a survey of experienced LCA analysts, only impact scores with an SP of at least 0.9 can be comfortably deemed significantly different. The second example demonstrates how SP can be used to determine whether CFs for a fine spatial scale are significantly different given their respective uncertainty. The watershed-delineated CFs for the direct human health impact of water consumption, an impact category included in Impact World+, were evaluated. For most countries, despite significant spatial variations in deterministic CF values, the uncertainty only makes it possible to distinguish a few archetypes per country, limiting the number of new CFs.

### **1 Introduction**

Uncertainty in LCA has received increasing attention in the past few years. As shown in the reviews of published LCA results in scientific papers (Ross, Evans et al. 2002; Lloyd and Ries 2007), the frequency of uncertainty reporting, the types of uncertainty and the diversity of assessment tools have all increased. Since these reviews, more studies containing more or less in depth analyses of uncertainty were published (Ibáñez-Forés,

Bovea et al. 2011; Röö, Sundberg et al. 2011; Venkatesh, Jaramillo et al. 2011; Mattila, Leskinen et al. 2012; Sills, Paramita et al. 2012; Weber 2012; Gregory, Montalbo et al. 2013).

In the life cycle inventory (LCI) phase, the most widely used method to evaluate the uncertainty of the input is the pedigree matrix (Weidema and Wesnæs 1996). Possibility theory has also been discussed (Benetto, Dujet et al. 2006; André and Lopes 2012; Clavreul, Guyonnet et al. 2012; Weber 2012; Clavreul, Guyonnet et al. 2013), and strategies to deal with data shortage have been explored (Olivetti, Patanavanich et al. 2013). Uncertainty evaluation in the life cycle impact assessment (LCIA) phase is not yet standard practice (van Zelm and Huijbregts 2013) due to difficulties in quantifying the uncertainty of the input parameters (Björklund 2002), but new LCIA models are beginning to report uncertainty in their characterization factors (CFs) (Posch, Seppälä et al. 2008; Manneh, Margni et al. 2010; Boulay, Bulle et al. 2011; Saad, Margni et al. 2011; Helmes, Huijbregts et al. 2012). The importance of uncertainty evaluation has, however, been recognized enough to be included as an evaluation criterion in a systematic review of state-of-the-art LCIA practices (Hauschild, Goedkoop et al. 2013). Finally, certain contributions recommend different tools to propagate uncertainty and identify major contributors (Huijbregts, Gilijamse et al. 2003; von Bahr and Steen 2004; Venkatesh, Jaramillo et al. 2011; Mattila, Leskinen et al. 2012; Weber 2012; Imbeault-Tétreault, Jolliet et al. 2013; Mutel, de Baan et al. 2013).

In general, all the contributions advance useful tools to quantify input uncertainty, propagate it to the impact scores and understand which uncertainty sources dominate the total uncertainty. However, very few contributions tackle the fundamental issue of evaluating the significance of the results in the presence of uncertainty. For example, the authors of a study comparing the greenhouse gas emissions of the production of different sources of fuel note that, after a Monte Carlo, the scores of certain options exhibit “significant overlap”, without further quantification (Venkatesh, Jaramillo et al. 2011). In order to take into account some of the unquantified uncertainties in LCIA, a baby food packaging alternative assessment proposes rules of thumb on the minimum ratio of scores between scenarios to determine a meaningful difference (Humbert, Rossi et al. 2009). In a beer and wine comparison, a risk level of 10% is assumed without further discussion (Mattila, Leskinen et al. 2012). In a hand-drier comparison (Gregory, Montalbo et al. 2013),

a more sophisticated Kruskal-Wallis test with significance levels of 90% and 95% was carried out without justification.

When looking only at the deterministic impact scores of two compared systems, it is easy to conclude that one is higher than the other. Introducing uncertainty assessment is akin to framing LCA as a hypothesis test to discover whether the difference is significant or if there is even a possibility that the opposite conclusion could be true. Hypothesis testing comes with a multitude of tools that could be introduced in uncertainty assessment.

The objective of this paper is to formally introduce the concepts of the null hypothesis and statistical power (SP) in LCA. The null hypothesis normally states that there is no relationship between two measured phenomena or that a potential medical treatment has no effect. By analogy, the null hypothesis in LCA implies that the impact scores of two compared systems are not significantly different. The SP, which is also known as “significance”, is the likelihood that no error is made when the null hypothesis is rejected (i.e. that two values deemed different are indeed different).

A small group of experienced practitioners were surveyed to compare their current uncertainty assessment practices and evaluate how a tool like SP would be an added value at this stage of the interpretation of an LCA. The tool has the potential to limit the unnecessary multiplication of characterization factors (CFs) for different archetypes of spatial delineation and inventory datasets for different production routes and regions, provided they are not significantly different. In a detailed example, SP is applied to the model of the direct human health impact of freshwater use (Boulay, Bulle et al. 2011) to explore the possibilities of complexity reduction.

## **2 Method**

### **2.1 Calculating statistical power**

We aim to provide information on the distinctiveness of two random variables (RV1 and RV2), considering their respective central tendencies and spreads. In the interest of simplicity, we first describe the approach for RVs that are represented by discrete probability distribution functions (PDFs), also known as empirical distribution functions (EDFs) (see Figure 1).

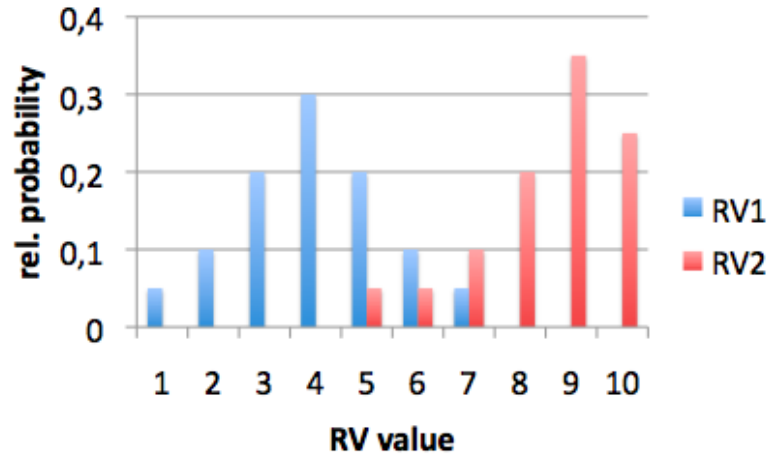


Figure 1: Two hypothetical empirical distribution functions

Table 1: SP calculation details based on Figure 1

RV value	Likelihood of drawing this value for RV2	Likelihood of drawing a lower value for RV1	Likelihood of drawing this value for RV2 and a lower value for RV1
5	0.05	0.65	0.0325
6	0.05	0.85	0.0425
7	0.1	0.95	0.095
8	0.2	1	0.2
9	0.35	1	0.35
10	0.25	1	0.25
<b>Total</b>			SP = 0.97

The means of RV1 and RV2 are 4.0 and 8.5, respectively. From a deterministic point of view, RV2 is therefore greater than RV1. However, as illustrated in Figure 1, there is a non-null chance that RV1 is, in fact, larger than RV2. The null hypothesis in this case is that there is no significant difference between RV1 and RV2. The alternative hypothesis is  $RV2 > RV1$ . The SP corresponds to the probability that it is correct to reject the null hypothesis (i.e. to assume that RV1 and RV2 are indeed different).

Table 1 shows how to calculate the probability that RV2 is larger than RV1. There is a 5% likelihood that the value of RV2 will be 5, and a 65% likelihood that the value of RV1 will be lower than 5. The likelihood that both are true simultaneously is  $5\% * 65\% = 3.25\%$ . The operation is repeated over the entire domain. Summing the likelihood of the events yields



the SP of the test, which, in this case, is 97%. There is therefore a 97% chance that we would not be wrong in stating that RV2 is indeed larger than RV1.

If continuous PDFs are available, the formula becomes:

$$SP = \int_{-\infty}^{+\infty} CDF_{RV1} \cdot PDF_{RV2} dx \quad \text{Eq. (1)}$$

where CDF is the cumulative distributive function. If there is no difference between RV2 and RV1,  $SP = 0.5$ ; when sampling twice from the same distribution, there is a 50% chance that sample 1 < sample2, and a 50% chance that sample 2 < sample 1.

The calculation of the SP from available EDFs or continuous PDFs is referred to as an *a posteriori* application of the SP concept (Park 2008). An *a priori* application is often used to determine required sample sizes: a desired SP is set along with the minimum significant difference that should be detected by the study from which the minimal sample size is calculated.

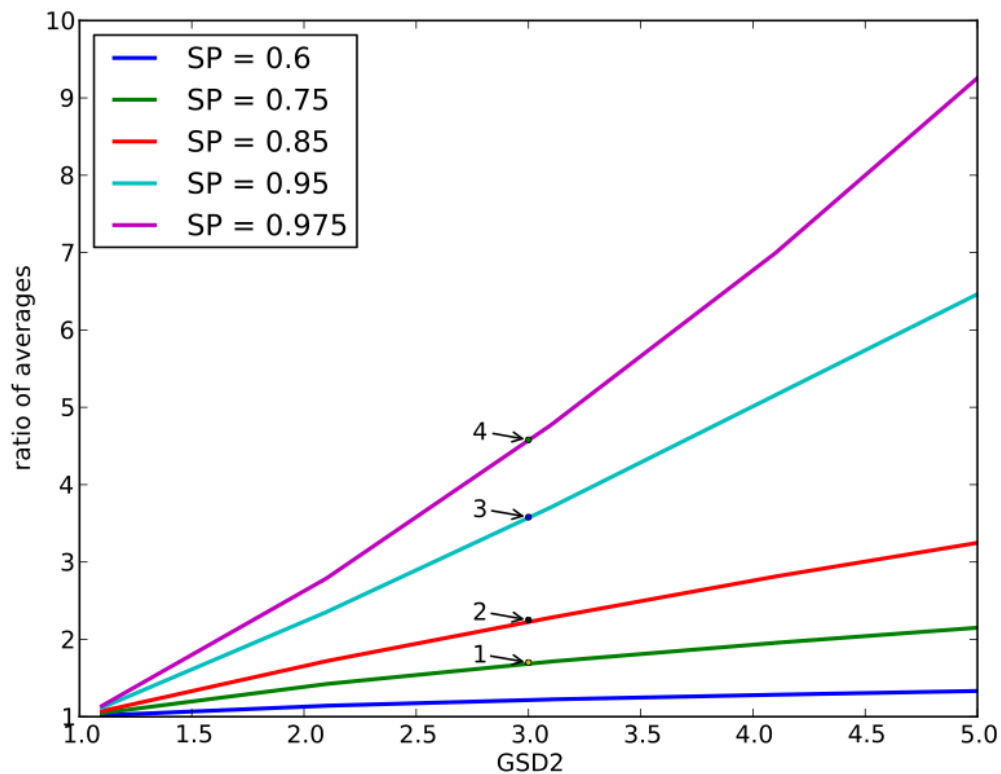
This raises the question of sample size. In statistical testing, thresholds for hypothesis acceptance and rejection depend on the sample size when the sample size is small. For example, in parametric tests, the variance of the Student's t-distribution diminishes and its shape tends towards the normal distribution as the sample size (degrees of freedom) tends towards infinity (Hines, Montgomery et al. 2003). In non-parametric tests like Kruskal-Wallis, lookup tables also depend on sample size (Spurrer 2003). However, in the presence of large samples like those generated by a Monte Carlo or of infinite samples like a PDF, SP can be calculated with a simple formula like Eq. (1) and Table 1.

## 2.2 Hypothetical cases

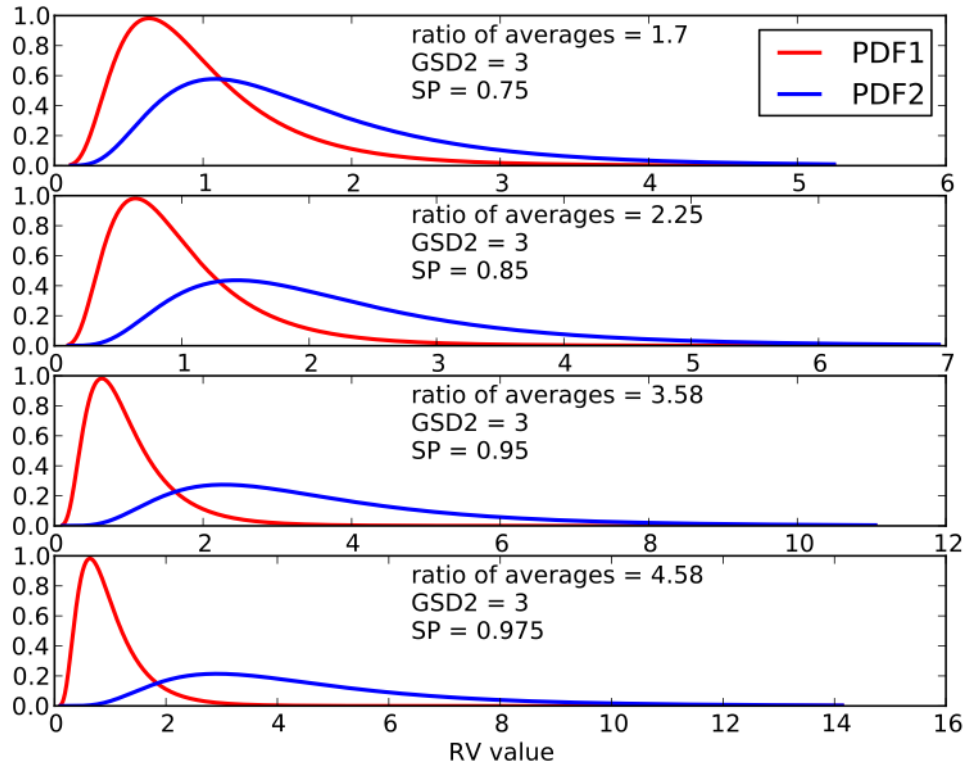
The aim of this section is to show what different SP values look like for RVs expressed as lognormals, one of the most commonly used distribution in LCI and LCIA. The typical measure of spread for this distribution is the geometric standard deviation (GSD2) (Heijungs and Frischknecht 2005). To reduce the number of variables, we check that SP is only a function of the  $\mu_2/\mu_1$  ratio when the GSD2s of both distributions are kept constant. It would be possible to formally prove this property if the CDF of a lognormal had an analytical expression, but this is not the case for many distributions, including the normal

and lognormal distributions. It is therefore numerically checked that when  $GSD2_1 = GSD2_2 = 2$ , if  $\mu_2/\mu_1$  is kept at 2, the SP always gives 0.9218 with several pairs of  $\mu_1$  and  $\mu_2$  values. This is true because the GSD2 is a dimensionless and relative expression of variance. To discard another variable, we also keep  $GSD2_1 = GSD2_2$ . Therefore, for the purpose of this study, SP is only a function of GSD2 and  $\mu_2/\mu_1$ .

Figure 2 illustrates the ratio of averages required to obtain a specific SP given the GSD2 of the RVs. The situation is limited by the assumptions and should only be used to get a first estimate of the ratio of averages. Figure 3 provides a better feeling of what SPs of 0.75, 0.85, 0.95 and 0.975 look like.



**Figure 2:  $\mu_2/\mu_1$  VS GSD2 for lognormal distributions. Points 1-4 refer to subplots of Figure 3**



**Figure 3: Four combinations of  $\mu_2/\mu_1$  and GSD2 with their respective SPs**

### 3 Sample applications

#### 3.1 Analysts survey

##### 3.1.1 Goal and scope of the survey

The objective of the survey was to compare the SP with current uncertainty assessment practices of experienced LCA practitioners. Analysts were asked to grade how comfortable they would be in advising a decision maker to act on the premise of comparative LCA results when confronted with impact scores with different amounts of overlap. We wanted to know whether the specialists would be more at ease with an increased SP, whether it was influenced by other variables such as the width of the distributions, and the importance of the decision that the information was meant to inform. Practitioners also answered open questions to better understand how they rate their comfort level.

### 3.1.2 Description of the protocol

LCA analysts were presented with 40 graphs (see Figure 4 for an example; a complete set of graphs and related information is available in Supplementary Material). For each graph, the question was the same: on a scale of 1 to 5 (“not comfortable at all” to “very comfortable”), how comfortable are you to state in an LCA report that the impact score is significantly lower for system 1? Further context was provided. Half of the graphs were from a fictitious LCA comparing the impacts of cups made of different materials and sponsored by chain of coffee shops to inform their supplier choice. The other half were from a fictitious LCA comparing the impacts of two energy supply policies to be presented to legislators to support a national ten-year development plan. The two series of graphs were actually identical in order to test whether the study consequences that were perceived as more significant would make analysts more conservative in their assessment.

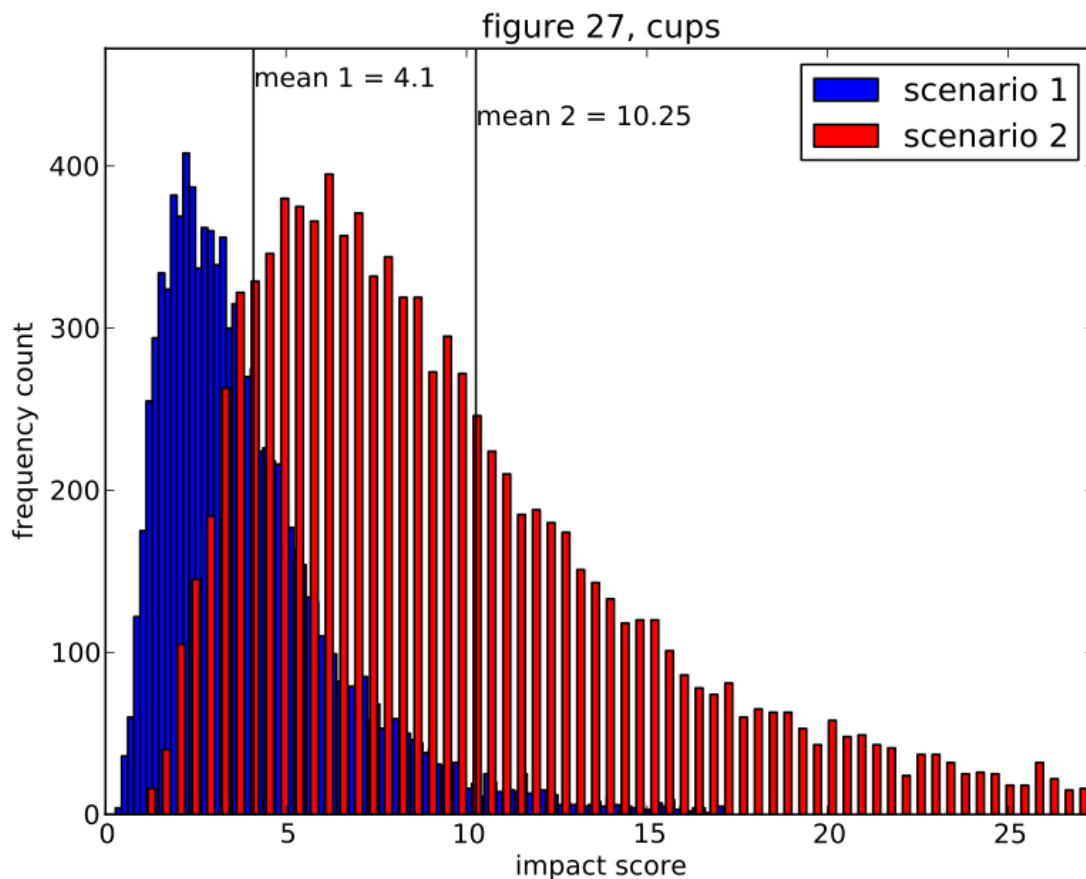
No impact category was specified, but the analysts were told that all the uncertainty had been factored in, including the uncertainty of the characterization factors, and that correlation had been taken into account. This additional information was provided because analysts sometimes rely on rules of thumb on meaningful differences between impact scores for categories such as human health and ecotoxicity, since the uncertainty of the CFs was not provided (Humbert, Rossi et al. 2009). We did not want them to use rules of thumb for this assessment and required that they focus only on the information provided in the survey.

SP levels were set at 0.75, 0.85, 0.90, 0.95 and 0.975, and GSD levels for both lognormals were set to 1.5, 3.5, 5.5 and 7.5. Every combination of SP/GSD was generated for a total of 20 graphs for each of the two fictional LCA. Presenting the same SP with different GSDs aimed to test whether the analysts would grade their level of comfort with the same SP differently: whether they, in other words, would be distracted by information that is irrelevant to the assessment. Graphs were generated for each case and randomly shuffled. The order of presentation of the graphs was the same for every analyst.

The analysts were shown the graphs one by one, without the possibility of seeing them a second time, changing their answers or accessing their previous answers (compiled by the interviewer). The analysts were consulted one-on-one and asked not to discuss their experience with other analysts.

The graphs displayed a histogram of the Monte Carlo, a vertical line indicating the means of each scenario and the numerical value of the means. The SP was not displayed, and the analysts were not prompted on the concept. After answering the question for every graph, the analysts were asked about their experience with open-ended questions: what kind of visual cues were they looking for? Why were they more comfortable with one graph as compared to the other? What kind of assumptions or previous experiences were they thinking about when answering?

In total, seven analysts with two to nine years' experience were surveyed. Six work for the CIRAIG and the other for Quantis Montréal.



**Figure 4: Example of graph presented to the analysts with SP = 0.85 and GSD2 = 3.5**

### 3.1.3 Quantitative survey results

Figure 5 shows that, as expected, the comfort level of the participants increased with increasing SPs. On one hand, most answers were more confident than not for an SP of 0.975. On the other hand, most answers reflected low comfort or total discomfort for SPs of 0.75

and 0.85. The results should be interpreted with a grain of salt because the scale of 1 to 5 is obviously subjective. Without a clear demarcation, two thresholds could be recommended:  $SP_{high} = 0.975$ , for which there is good consensus on the meaningfulness of the difference, and  $SP_{low} = 0.9$ , under which the difference is not considered meaningful.

Figure 6 shows that analysts were relatively unaffected by the change in spread. Presented with strictly identical SPs, analysts appeared, on average, slightly less comfortable when the GSD2 was higher. However, given the very small sample size, we do not believe that the difference is indicative of a significant trend.

Finally, Figure 7 indicates that, most of the time, the perceived consequences of the LCA had little to no influence on the comfort assessment. For the 20 cases, the difference in comfort level was calculated for each analyst. The graphs were strictly identical, except for the the title (“cup” or “policy”). For the most part, the experts did not feel that the information affected their assessment. Unless they were hesitating between two answers, the analysts would choose a lesser comfort level with the policy case.

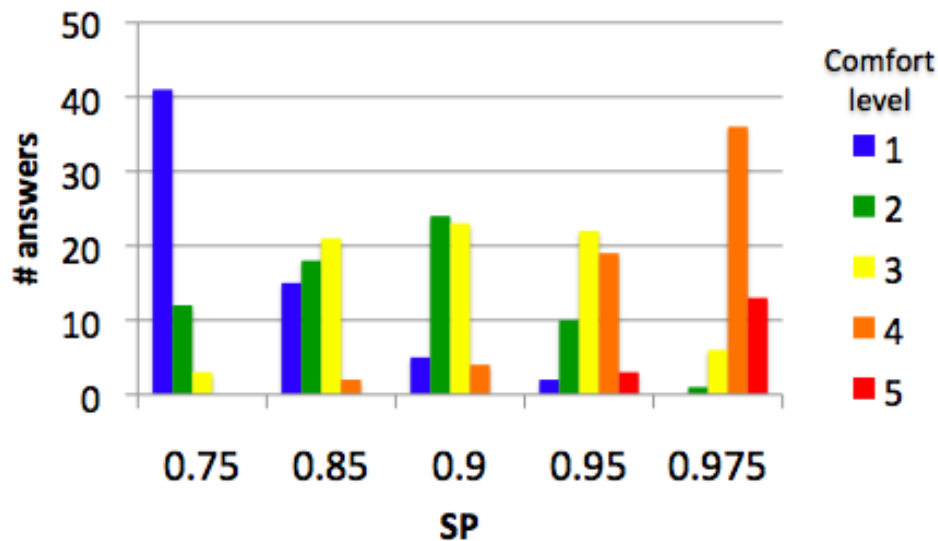


Figure 5: General survey results

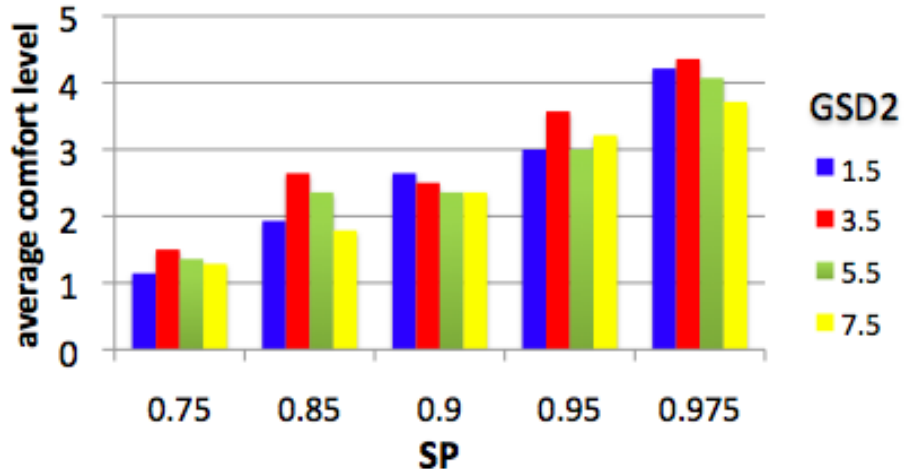


Figure 6: Average comfort level as a function of SP and GSD2

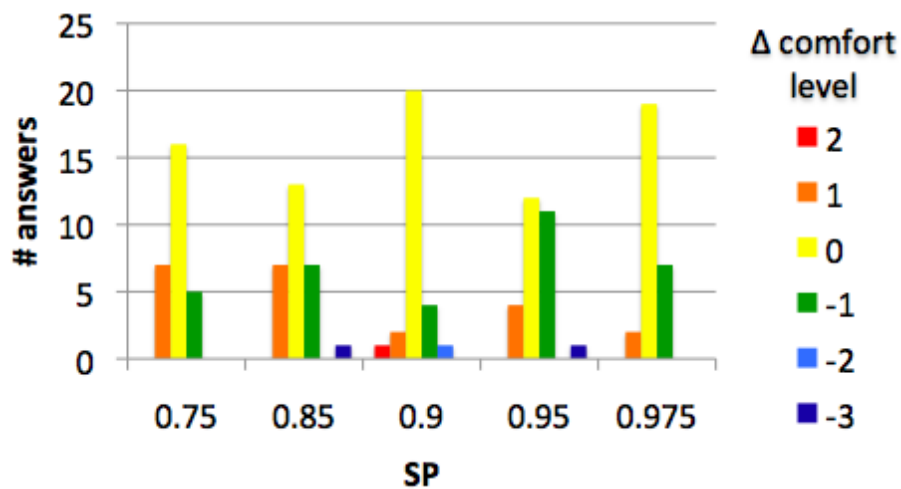


Figure 7: Change of answers between the “cup” and “policy” LCA. Negative numbers indicate lower comfort levels for the “policy” LCA.

### 3.1.4 Qualitative survey results

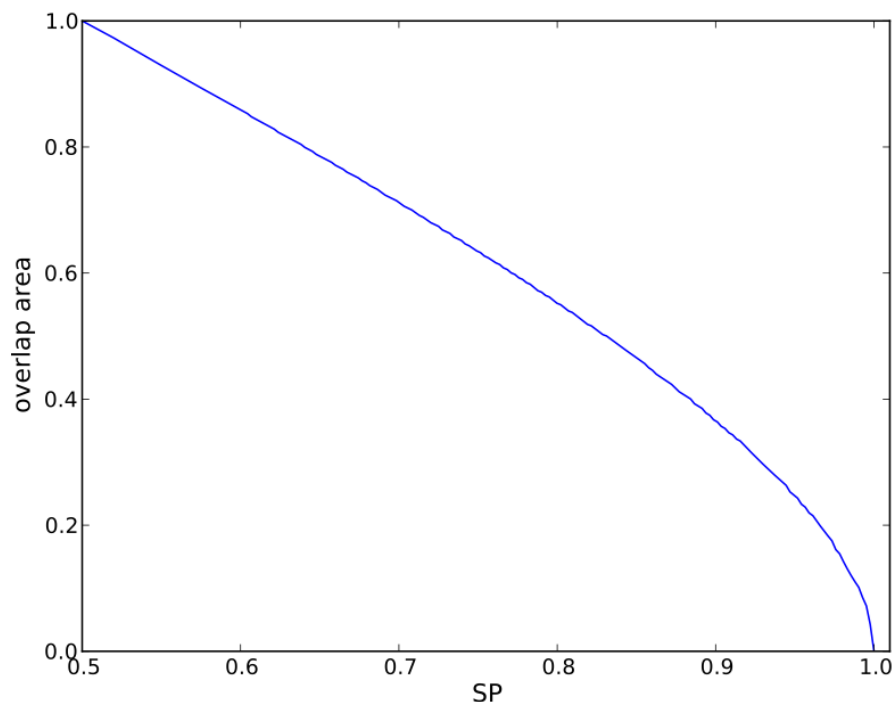
Even though rigorous interview standards were not applied, it quickly felt like more interviews would not provide new information. We believe that saturation was achieved (Bowen 2008). However, it should be stated that all the analysts interviewed were either employed by or graduates of the CIRAIG, potentially creating an organizational bias. The survey should be repeated with analysts who gained their skills and experiences in another organization.

In general, the analysts seemed to instinctively perceive high and low SP graphs with high and low degrees of comfort, respectively. The standard deviation of comfort level is higher

for intermediate SP values—a sign that the analysts relied on heterogeneous methods to dispel ambiguity they were not feeling when the SP was low or high.

Responses differed when the analysts were asked what they were looking for on the graphs to rate their comfort level. Most often, they looked at the area of overlap of the two curves. It is true that for continuous RVs, the overlap area is negatively correlated with the SP (Figure 8), and both are functions of the same set of parameters. But the situation is not as simple for a Monte Carlo result presented in histogram form. The area of overlap in a histogram is a function of an additional parameter: bin width. By reducing the bin width, it is possible to represent the same results with a greater number of smaller bins, artificially reducing the overlap area. In other words, one could manipulate perception. The overlap area should therefore be avoided as an SP proxy.

Despite their many years of experience, most analysts admitted that they lacked knowledge and tools to justify their comfort level. One participant mentioned that, under such circumstances, she preferred having clear quantitative guidelines to write recommendations. Decision makers should, of course, play a role in determining acceptable SPs for a given study.



**Figure 8: Overlap area vs. SP for continuous RVs**



### **3.1.5 Integrating SP in uncertainty assessment**

On one hand, LCA experts are capable of fuzzy assessment, weighing many aspects all at once, including qualitative information, to formulate more nuanced and complex judgments and discuss limitations. They conduct the study, interview specialists on the product/service studied, collect data and gain a deeper understanding of where the biases and uncertainties lie, even if they cannot translate these elements into a quantitative assessment. On the other hand, the SP is only a number computed with a formula but it remains interesting in its objectivity. It should be included as an assessment tool, but should not, by any means, replace expert judgment.

Two analysts underscored the somehow artificial setting of the survey since they had not carried out the LCAs leading to the results themselves. In their opinion, it made them more cautious than they may have been in a more realistic situation. We agree that a greater number of analysts should be consulted in more realistic and diverse contexts in order to establish guidelines.

We also doubt whether a definitive guideline is desirable. Maybe lower SPs are acceptable for smaller, cheaper or quicker LCAs. Perhaps decision-makers should be asked how risk-averse they are, what the consequences of the decisions may be and what SPs would be acceptable for them. There is an irreducible subjectivity in setting a threshold between significant and insignificant SP values.

### **3.2 Case study: reduction in the number of cells in the direct human health impact of freshwater use**

Two methods to reduce the number of cells in a spatially differentiated impact assessment model were recently proposed (Mutel, Pfister et al. 2012; Mutel and Hellweg 2013). They relied on the minimization of spatial auto-correlation and Ward hierarchical clustering, respectively. In both cases, the end result is delineation with a smaller number of cells, more clearly revealing the gradient of impact and the landscape parameters that drive it. However, the methods treat CFs as deterministic values. In this paper, we will neglect uncertainty in spatial localization and focus on CF uncertainty. Further work could explore how the two approaches could be combined in the same method (only aggregating

contiguous cells that are not significantly different) or sequentially (checking whether the cells created are significantly different from one another).

The Impact World+ category “direct human health impact of freshwater use” is used as a case study (Boulay, Bulle et al. 2011). Water is categorized according to chemical and biological characteristics and location (ground or surface), creating 16 different categories of water quality. The impact of the water use depends on the consumption to availability ratio (a measure of competition for water among users to determine which users will be affected by the lack of domestic, agricultural and aquaculture water) and the capacity to pay to compensate for the lack of water as a function of the gross national income. The result is a map of one native CF (nCF) per country-watershed for a total of 808 CFs per water category.

The uncertainty of 13 model input parameters was quantified either by data or expert judgment and then propagated to the endpoint CFs with 1 000 iterations from a stratified sampling Monte Carlo.

The top subplot of Figure 9 displays the 2.5 and 97.5 percentiles, quartiles, average and outliers of the Monte Carlo for each of the 14 watersheds in India for S3 water quality. The nCFs were ordered according to arithmetic average. The bottom subplot shows the SP for each pair of nCFs inside India. For a country with  $n$  watersheds, the total number of pair comparisons is  $(n^2-n)/2$  (represented by dots), and the number of pair comparisons between average-ordered neighbours is  $n-1$  (represented by stars). The x-axis represents the ratio of the compared averages. The horizontal red line represents an arbitrary  $SP_{\min}$  threshold of 0.95.

The number of average-ordered neighbour SPs above the thresholds represents clear demarcations inside a country. For S3 water quality in India, they appear between cells 555 and 375, and 518 and 515. Therefore, the 14 watersheds could be grouped into three different archetypes: low, medium and high impact of water consumption, respectively, at averages  $\sim 5 \times 10^{-10}$ ,  $\sim 1 \times 10^{-4}$  and  $\sim 5 \times 10^{-3}$  DALY/m<sup>3</sup>.

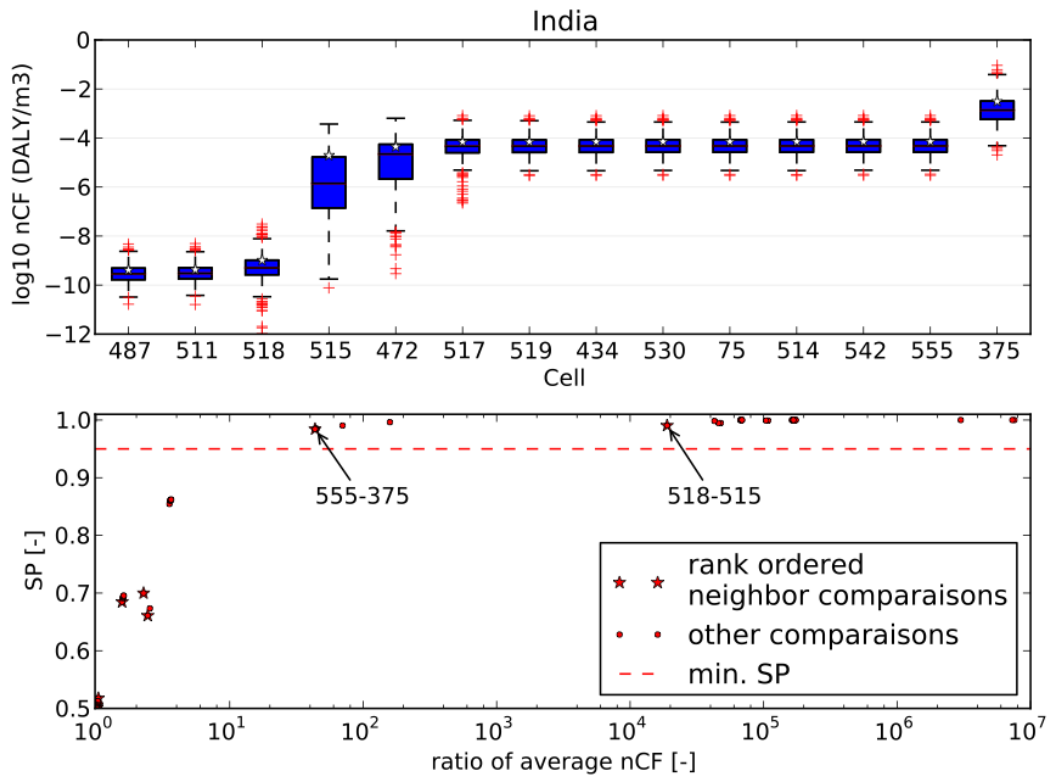


Figure 9: Monte Carlo results and SPs for the use of S3 water quality in India

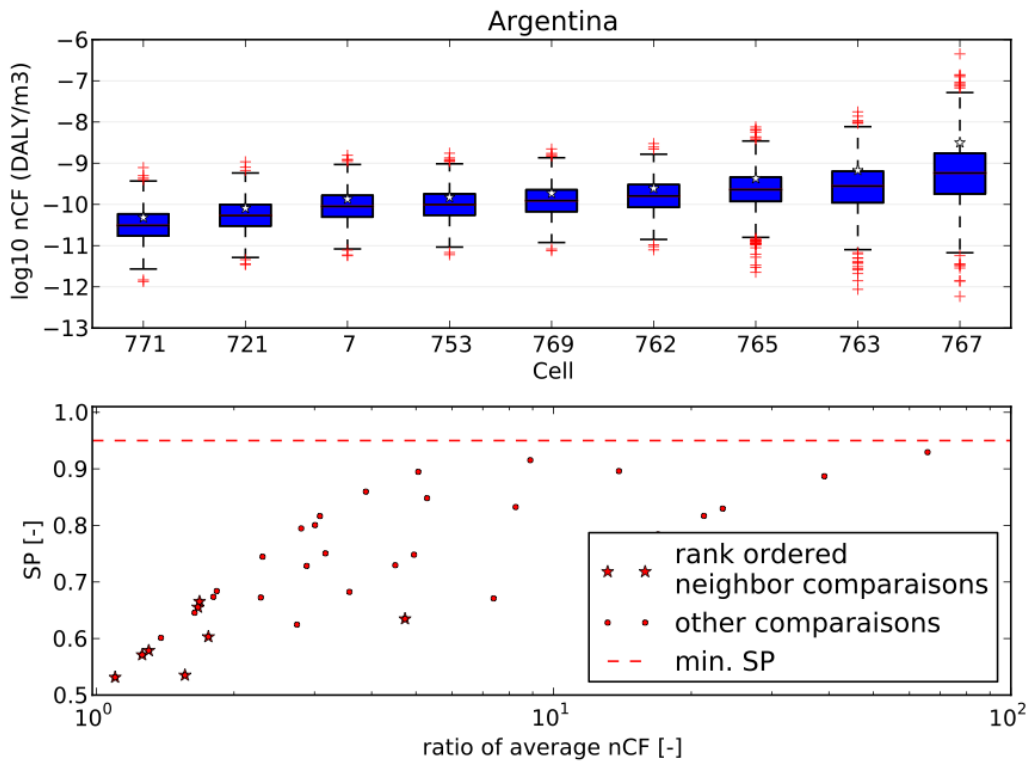
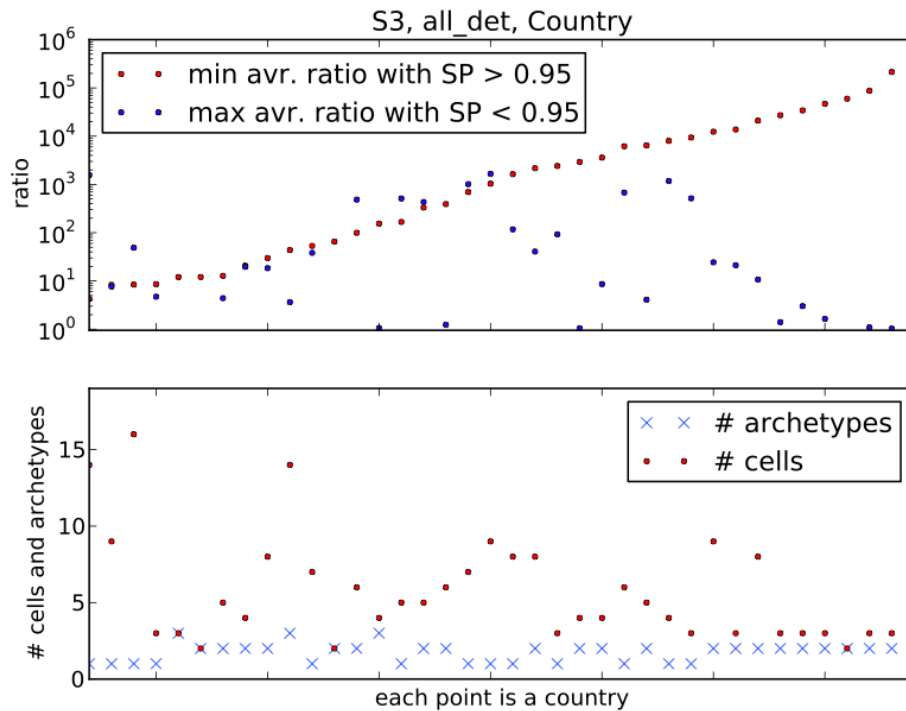


Figure 10: Monte Carlo results and SPs for the use of S3 water quality in Argentina

The situation is not as clear-cut for S3 water quality in the nine Argentinian watersheds (Figure 10). Even if no average-ordered neighbour SPs are above the threshold, the nCF of cell 767 is almost significantly greater than many of the lower impact watersheds. Category developers may want to create archetypes based on parameters other than the endpoint nCF (for example, the midpoint nCF) or lower the SP threshold. Any watershed close to the limit would be arbitrarily put in one or the other category, and careful judgment is required. It is impractical to look at the graphs separately for every country in order to gain a better perspective of the entire category for a particular water quality. Each graph was summarized as four numbers, represented in Figure 11 for every country with non-zero characterization factors. The top subplot shows the lowest ratio of average comparison inside a country producing an SP above the threshold and the largest ratio of average comparison producing an SP below the threshold. The numbers illustrate how far apart averages have to lie in order to be considered distinct and how powerful the uncertainty is at rendering nCF undistinguishable, respectively. It is confusing to see that, for some countries, the lowest ratio average above the threshold is higher than the highest ratio average below the threshold. This is possible because the uncertainty is not uniform within the category. The strong non-linearity of the model makes certain regions of the domain more or less sensitive to the uncertainty of the input (Bourgault, Lesage et al. 2013). The bottom subplot displays how many watersheds are present in each country and how many significantly different regions (“archetypes”) result for the selected threshold. Most countries have two archetypes. The maximum is three, even for a country like China, covered by watersheds. It should also be noted that the high/medium/low levels of impacts differ from country to country.



**Figure 11: Overview of the impact category for S3 water quality**

## 4 Discussion

### 4.1 Applications by analysts

The SP can be used to compare the impact scores of two systems. LCA practitioners and sponsors could agree on an SP objective, which would justify the resources spent on data collection. If funds run out before all potential uncertainty reduction sources are exploited, more funds could be allocated or the sponsors could accept that the model does not distinguish between the options considered.

The SP could also be used as a tool to assist analysts when discussing impact score uncertainty in the interpretation phase of an LCA. The survey determined that the analysts were comfortable in affirming that two impact scores were different when the SP was above 0.975 and uncomfortable affirming that two impact scores were different when the SP was below 0.9.

There is an unavoidable amount of subjectivity in setting the limit between “significantly different” and “not significantly different”. This involves knowledge of the risk aversion of the decision makers. One may also argue that setting a threshold should be avoided in order

to avoid blind reliance on a guideline and favour more nuanced case-by-case discussion. Perception of risk by decision-makers is typically a management research field and LCA would gain at letting it inform uncertainty assessment practices to ensure smooth integration in the decision-making process.

The SP can also be used in an LCA scenario analysis. Parameters whose realistic worst and best case values make the SP move from one side of the threshold to the other should be flagged as sensitive, and high quality data should be collected for the parameter.

The SP may also come in handy when inventory or primary data collection characterizes an elementary flow as “unspecified” (e.g. not specifying the release sub-compartment). The analyst can choose among a range of possible CFs. If the variability among the possible choices is comparable to the uncertainty, the choice does not have to be submitted to a scenario analysis for a potential change in the assessment conclusion.

#### **4.2 Application by LCI and LCIA developers**

Another purpose of the SP is to make sure that the time and resources allocated to the complexification of the LCA will result in impact score differences that will remain significant when compared to the uncertainty of the scores. It is only relevant to create archetypes based on geographical location, compartment or sub-compartment of emission, stack height, soil type or population density for a particular impact category if it enhances the discriminating power of the LCA.

As shown by its application to the Impact World+ direct human health impact of water consumption indicator, the SP is a relevant tool to simplify the implementation of regionalized impact assessment. The motto of model developers is “as complex as necessary, as simple as possible”. The introduction of multiple CF values for the same EF, leading to virtually undistinguishable results, does not follow this guiding principle. Impact assessment developers should strive to rigorously quantify the uncertainty of their model and use the SP to minimize the number of significantly different archetypes, thus increasing the discriminating power of LCA with minimal complexification.

The creation of aggregated CFs (aCFs) entails the choice of a weighting set to calculate the weighted averages of the nCFs: population density, agriculture production density,

emission proxy, etc. The choice of weighting set changes the value of the aCF, but the difference in the final result may be slight as compared to the uncertainty of the aCF.

In general, any change in assumption, model or data quality or any other improvement may result in a change in uncertainty. In this situation, the SP can be used as an indicator of the change in the discriminating power of an LCA.

The same assessment should be carried out when considering the creation of a new LCI dataset for a given product. Only when the flows associated with activities are significantly different across geographical locations, technological options or temporal horizons is the creation of the new dataset justified.

### **4.3 Limitation**

The use of SP in LCA unfortunately suffers from a serious limitation. So far, there are only two ways of dealing with inter-system correlation in a comparative LCA assessment: assuming no correlation at all or assuming 100% correlation. The correlation issue has yet to be properly established (Bourgault, Lesage et al. 2012). We already know that assuming no correlation is the most conservative assumption, but this overestimates the potential overlap between the compared quantities, possibly resulting in a lower SP and the false conclusion that two quantities are not significantly different.

Taking into account the correlation in a meaningful and realistic way requires many assumptions, which come with pitfalls and additional uncertainties. The task should only be done in cases in which the difference in conclusion between 0% and 100% correlation justifies the additional effort. An incentive to work on the correlation issue is to check the sensitivity of the SP to the 0% and 100% correlation assumptions.

## **5 Acknowledgements**

The authors acknowledge the industrial partners of the International Chair in Life Cycle Assessment (a research unit of CIRAIG [Interuniversity Research Centre for the Life Cycle of Products, Processes and Services]) for their financial support: Alcan, Arcelor-Mittal, Bell Canada, Cascades, Eco-Entreprises-Québec/Recyc-Québec, Groupe EDF/GDF-SUEZ, Hydro-Québec, Johnson & Johnson, Mouvement des caisses Desjardins, RONA, Total and Veolia

Environnement. We also wish to thank Vahid Partovi Nia and Anne-Sophie Julien for their input on the statistical aspects of the paper, the analyst involved in the survey.

## 6 References

- André, J. S. and D. Lopes (2012). "On the use of possibility theory in uncertainty analysis of life cycle inventory." The International Journal of Life Cycle Assessment **17**(3): 350-361.
- Benetto, E., C. Dujet, et al. (2006). "Possibility Theory: A New Approach to Uncertainty Analysis?" The International Journal of Life Cycle Assessment **11**(2): 114-116.
- Björklund, A. (2002). "Survey of approaches to improve reliability in LCA." The International Journal of Life Cycle Assessment **7**(2): 64-72.
- Boulay, A.-M., C. Bulle, et al. (2011). "Regional Characterization of Freshwater Use in LCA: Modeling Direct Impacts on Human Health." Environ. Sci. Technol **45**(20): 8948-8957.
- Bourgault, G., P. Lesage, et al. (2012). Uncertainty and correlation: Thinking ahead. LCA XII. Tacoma.
- Bourgault, G., P. Lesage, et al. (2013). Opening Pandora's box: uncertainty propagation in life cycle impact assessment. SETAC Europe. Glasgow.
- Bowen, G. (2008). "Naturalistic inquiry and the saturation concept: a research note." Qualitative Research **8**(1): 137-152.
- Clavreul, J., D. Guyonnet, et al. (2012). "Quantifying uncertainty in LCA-modelling of waste management systems." Waste Management **32**(12): 2482-2495.
- Clavreul, J., D. Guyonnet, et al. (2013). "Stochastic and epistemic uncertainty propagation in LCA." The International Journal of Life Cycle Assessment **18**(7): 1393-1403.
- Gregory, J., T. Montalbo, et al. (2013). "Analyzing uncertainty in a comparative life cycle assessment of hand drying systems." The International Journal of Life Cycle Assessment **18**(8): 1605-1617.
- Hauschild, M., M. Goedkoop, et al. (2013). "Identifying best existing practice for characterization modeling in life cycle impact assessment." The International Journal of Life Cycle Assessment **18**(3): 683-697.
- Heijungs, R. and R. Frischknecht (2005). "Representing Statistical Distributions for Uncertain Parameters in LCA. Relationships between mathematical forms, their representation in EcoSpold, and their representation in CMLCA." The International Journal of Life Cycle Assessment **10**(4): 248-254.
- Helmes, R., M. Huijbregts, et al. (2012). "Spatially explicit fate factors of phosphorous emissions to freshwater at the global scale." International journal of Life Cycle Assessment **17**(5): 646-654.
- Hines, W., D. Montgomery, et al. (2003). Probability and Statistics in Engineering, Hoboken, NJ: Wiley.
- Huijbregts, M., W. Gilijamse, et al. (2003). "Evaluating Uncertainty in Environmental Life-Cycle Assessment. A Case Study Comparing Two Insulation Options for a Dutch One-Family Dwelling." Environ. Sci. Technol **37**(11): 2600-2608.



- Humbert, S., V. Rossi, et al. (2009). "Life cycle assessment of two baby food packaging alternatives: glass jars vs. plastic pots." The International Journal of Life Cycle Assessment **14**(2): 95-106.
- Ibáñez-Forés, V., M.-D. Bovea, et al. (2011). "Life cycle assessment of ceramic tiles. Environmental and statistical analysis." The International Journal of Life Cycle Assessment **16**(9): 916-928.
- Imbeault-Tétreault, H., O. Jolliet, et al. (2013). "Analytical Propagation of Uncertainty in Life Cycle Assessment Using Matrix Formulation." Journal of Industrial Ecology **17**(4): 485-492.
- Lloyd, S. and R. Ries (2007). "Characterizing, Propagating, and Analyzing Uncertainty in Life-Cycle Assessment: A Survey of Quantitative Approaches." Journal of Industrial Ecology **11**(1): 161-179.
- Manneh, R., M. Margni, et al. (2010). "Spatial Variability of Intake Fractions for Canadian Emission Scenarios: A Comparison between Three Resolution Scales." Environ. Sci. Technol **44**(11): 4217-4224.
- Mattila, T., P. Leskinen, et al. (2012). "Uncertainty in environmentally conscious decision making: beer or wine?" The International Journal of Life Cycle Assessment **17**(6): 696-705.
- Mutel, C. and S. Hellweg (2013). Improving the Spatial Scale of Impact Assessment: Analyzing and Applying the Regionalized LC-IMPACT Methods. SETAC Europe. Glasgow.
- Mutel, C., S. Pfister, et al. (2012). "GIS-Based Regionalized Life Cycle Assessment: How Big Is Small Enough? Methodology and Case Study of Electricity Generation." Environ. Sci. Technol **46**(2): 1096-1103.
- Mutel, C. L., L. de Baan, et al. (2013). "Two-Step Sensitivity Testing of Parametrized and Regionalized Life Cycle Assessments: Methodology and Case Study." Environmental Science & Technology **47**(11): 5660-5667.
- Olivetti, E., S. Patanavanich, et al. (2013). "Exploring the Viability of Probabilistic Under-Specification To Streamline Life Cycle Assessment." Environmental Science & Technology **47**(10): 5208-5216.
- Park, H. M. (2008) "Hypothesis Testing and Statistical Power of a Test." The University Information Technology Services (UITS) Center for Statistical and Mathematical Computing, Indiana University.
- Posch, M., J. Seppälä, et al. (2008). "The role of atmospheric dispersion models and ecosystem sensitivity in the determination of characterisation factors for acidifying and eutrophying emissions in LCIA." The International Journal of Life Cycle Assessment **13**(6): 477-486.
- Röös, E., C. Sundberg, et al. (2011). "Uncertainties in the carbon footprint of refined wheat products: a case study on Swedish pasta." The International Journal of Life Cycle Assessment **16**(4): 338-350.
- Ross, S., D. Evans, et al. (2002). "How LCA studies deal with uncertainty." The International Journal of Life Cycle Assessment **7**(1): 47-52.
- Saad, R., M. Margni, et al. (2011). "Assessment of land use impacts on soil ecological functions: development of spatially differentiated characterization factors within a Canadian context." International journal of Life Cycle Assessment **16**(3): 198-211.

- Sills, D. L., V. Paramita, et al. (2012). "Quantitative Uncertainty Analysis of Life Cycle Assessment for Algal Biofuel Production." Environmental Science & Technology **47**(2): 687-694.
- Spurrer, J. D. (2003). "On the null distribution of the Kruskal–Wallis statistic." Journal of Nonparametric Statistics **15**(6): 685-691.
- van Zelm, R. and M. A. J. Huijbregts (2013). "Quantifying the Trade-off between Parameter and Model Structure Uncertainty in Life Cycle Impact Assessment." Environmental Science & Technology **47**(16): 9274-9280.
- Venkatesh, A., P. Jaramillo, et al. (2011). "Uncertainty in Life Cycle Greenhouse Gas Emissions from United States Natural Gas End-Uses and its Effects on Policy." Environmental Science & Technology **45**(19): 8182-8189.
- von Bahr, B. and B. Steen (2004). "Reducing epistemological uncertainty in life cycle inventory." Journal of Cleaner Production **12**(4): 369-388.
- Weber, C. L. (2012). "Uncertainty and Variability in Product Carbon Footprinting." Journal of Industrial Ecology **16**(2): 203-211.

### Supplementary material for “Statistical hypothesis testing tool applied to LCA: Judging the significance of differences”

This supplementary material supplies all the information presented to the analysts during the survey. The 40 figures are provided, as they were shown to the analysts, and in the same order. Table 1 provides the information not appearing on the figures but necessary for analysis of the results.

**Table 1 : Information about the survey figures**

figure	LCA	DI	GSD
1	cups	0.85	5.5
2	cups	0.975	1.5
3	cups	0.975	5.5
4	policy	0.85	7.5
5	policy	0.75	3.5
6	cups	0.975	7.5
7	policy	0.975	7.5
8	cups	0.95	7.5
9	policy	0.95	3.5
10	policy	0.975	5.5
11	policy	0.95	1.5
12	cups	0.95	1.5
13	policy	0.9	5.5
14	cups	0.9	1.5
15	cups	0.85	7.5
16	policy	0.95	5.5
17	cups	0.95	3.5
18	policy	0.9	3.5
19	cups	0.75	1.5
20	cups	0.9	3.5
21	policy	0.75	5.5
22	policy	0.85	1.5
23	cups	0.85	1.5
24	policy	0.75	1.5
25	cups	0.9	7.5
26	policy	0.975	3.5
27	cups	0.85	3.5
28	policy	0.975	1.5
29	policy	0.9	7.5
30	cups	0.95	5.5
31	cups	0.75	7.5
32	cups	0.9	5.5
33	policy	0.95	7.5
34	policy	0.75	7.5
35	cups	0.75	5.5
36	cups	0.75	3.5
37	policy	0.85	3.5
38	policy	0.85	5.5
39	policy	0.9	1.5
40	cups	0.975	3.5

**ANNEXE 3 – Article 3 : « Relevance and feasibility of sector-specific country-level characterization factors»**

## **Relevance and feasibility of sector-specific country-level characterization factors**

Guillaume Bourgault, Pascal Lesage, Réjean Samson

### **Abstract**

Many regional life cycle impact assessment (LCIA) categories provide characterization factors (CFs) with a finer geographical delineation than the typically national scale of life cycle inventory. One method to deal with the incongruence between spatial scales is to create a weighted average of the CFs included in a country. Weighting sets based on population count have been used, assuming correlation between population and emissions. However, we know this correlation will only be true for some industry sectors. Industry-generic weighting sets could lead to over- or under-estimated CFs for sectors that are not located in populated areas.

Data obtained from government agencies on facility locations, emissions and economic activity reveal that, for terrestrial acidification and freshwater eutrophication, population and industry-generic weighting lead to comparable CFs. However, industry-specific CFs can differ from industry-generic CFs by more than one order of magnitude. The increase in size of the linear system of equations implied by the creation of industry-specific CFs was estimated, and the associated calculation times were measured. Using an average laptop computer and software packages optimized for sparse matrix operations, the calculation time remained around one minute, indicating that the large-scale implementation of industry-specific CFs would not create computation problems.

### **1 Introduction**

Recent developments in life cycle impact assessment (LCIA) models have shown the relevance of delineating territories in fine air grids, watersheds or subwatersheds in order to capture significant spatial variability for regional impact categories<sup>1-4</sup>. However, the life cycle inventory (LCI) phase typically yields an inventory with a coarser delineation (country, continent)<sup>5-6</sup>. To ensure the applicability of the finely delineated LCIA methods as required by the latest evaluation criteria<sup>7</sup>, an aggregation to a delineation consistent with the inventory is necessary. The weighted average solution has already been retained on

certain occasions<sup>2</sup>. The procedure calls for a weighting set that represents the likelihood of emission in each cell.

Spatially resolved population counts are readily available and were used as an emission proxy to build weighting sets in the previously cited cases, thus assuming a good correlation between the population counts and rates of emission of different chemicals. A superficial examination reveals that this may not be true for all economic sectors. While, for example, sewerage-related activities correlate very well with population count, natural resources extraction (mining, forestry, agriculture and off-shore petroleum extraction) is often located in less populated regions. However, it is not clear whether these differences would result in significant differences in the characterization factors (CFs) for regionalized impact categories.

The first objective of this paper is to demonstrate the relevance and feasibility of the creation of industry-specific CFs through the use of industry-specific geolocalized emission rates. For this paper, three government databases containing the required information were selected (E-PRTR for Europe, NEI for the United States and NPRI for Canada): geolocalized facilities are categorized with an economic activity classification number (ISIC, NAIC or NACE) and report quantities emitted for a list of pollutants. Other datasets could complement or validate the findings of this paper.

The second objective is to measure the agreement between industry-generic, industry-specific and population count weighting sets. We show the relevant calculation and interpretation methods to evaluate the differences caused by the choice of weighting set. These methods are not specific to the data sources but will remain the same for every impact category and source of proxy data. The results show that the creation of industry-specific CFs is justified by the range of values they cover as compared to the industry-generic CFs. It is also demonstrated how, when integrating sector-specific and spatially-differentiated CFs, the increase in model size does not pose significant computation problems.

## **2 Material and methods**

### **2.1 Sources of proxy data for emission proxies**

In the LCA framework, the fate of pollutants is calculated within the LCIA phase. It was therefore essential to find data on the location of the emission *sources* without taking fate into account. Many measurements and detection methods report the location where the pollution is found after being displaced by the elements. As such, these data sources were not considered.

#### **2.1.1 European Pollutant Release and Transfer Register (E-PRTR)**

The European Pollutant Release and Transfer Register is managed by the European Environment Agency (EEA) and gathers the data reported by EU member states, Iceland, Liechtenstein, Norway, Serbia and Switzerland<sup>8</sup>. The facility report includes information on the latitude, longitude and NACE number of 135 546 facilities. The pollutant release report contains the amount of pollutants released by the facilities for 91 substances. Emissions are identified as measured, calculated or estimated. Only the point sources were used in this analysis.

Some facilities are exempt from the report<sup>9</sup>. The EEA lists the threshold of production capacities under which a facility does not have to report emissions. The threshold depends on the activity sector. Also, facilities do not have to report emissions when they are under a certain release threshold.

The EEA has implemented several validation mechanisms, including checks against data received from licensing and permit procedures, community eco-management and audit schemes like ISO14001, the Convention of Long Range Transboundary Air Pollution (CLRTAP) and the Water Information System of Europe (WISE).

#### **2.1.2 National Emissions Inventory**

The National Emissions Inventory (NEI) is managed by the US Environmental Protection Agency (USEPA) and gathers data on criteria and hazardous air pollutants about air emission sources. Its scope is very similar to the E-PRTR<sup>10</sup>. Facilities are associated with their North America Industry Classification System (NAICS) number. The database also

includes a list of quality assurance criterion (stability of emissions for each facility over the year and checks for outliers). Since data comes from a variety of programs (for example, Toxic Release Inventory and Acid Rain Program), cut-off rules are not uniform for each substances. In Toxic Release Inventory, a facility is required to report if it has 10 or more full-time employee equivalents and manufactures, processes or otherwise uses EPCRA Section 313 chemical over an established threshold.

### **2.1.3 National Pollutant Release Inventory**

The National Pollutant Release Inventory (NPRI) is managed by Environment Canada and gathers data similar to the NEI and E-PRTR for 8 096 facilities on its territory<sup>11</sup>. Facilities are geolocalized and associated with their NAICS number.

Barring certain exceptions, facilities with less than the equivalent of 10 full-time employees are not required to report their emissions. Emission thresholds per substance is ten manufactured, processes or otherwise used (MPO) tonnes per year, except for PM, VOC, heavy metals, PAH, dioxin and furans. Along with the NEI, the NPRI has already been used to build datasets for normalization references in North America<sup>12-13</sup>.

## **2.2 Data sources for CFs**

The following sections use two regional impact categories with a fine delineation: freshwater eutrophication<sup>1</sup> and terrestrial acidification<sup>3</sup>. The methods are briefly described here, and more detailed information can be found in the original papers.

The eutrophication method uses a 0.5°x0.5° grid cell delineation. Within each cell, mass balance is achieved between soil, retention, small rivers and the main water system. Advection between cells is also calculated through topological networks, making it possible to calculate fate factors. Effect factors relate phosphorous concentrations to species effects in each cell and take into account existing phosphorous concentrations.

The acidification method uses a coarser 2°x2.5° grid. A grid-to-grid atmospheric model was used to predict the fate of emissions. Cells also display a varying sensitivity factor, representing the change in pH for a unit amount of acidifying substance. Finally, the effect



factor considers the sensitivity of the ecosystems to pH change. The CFs were calculated for SO<sub>2</sub>, SO<sub>4</sub>, NO<sub>x</sub>, NH<sub>3</sub> and HNO<sub>3</sub> emissions.

### **2.3 Population count data source**

The population count for 2005 was obtained at the 0.5°x0.5° resolution<sup>14</sup>. The total population was calculated via GIS software for each cell of the delineation at hand. When a cell was divided into two or more countries, the population of the cell was allocated to each country according to the proportion of surface it occupied on the cell.

### **2.4 Algorithm to create weighting sets and CFs**

The population count of each cell of a given country was divided by the country's population count. A similar procedure was applied to emission proxies: the amount emitted in each cell was divided by the total emitted in the whole country. The procedure was repeated for each ISIC number emitting the substance at hand. In the case of cells split by national borders, allocation based on surface was not necessary since each plant was tagged with its country.

For each country, the vector containing a set of weights was multiplied element-by-element with the vector containing the CFs of each cell. The resulting vector represents the contribution of each cell to the national CF, and the sum of those contributions led to the ISIC-generic or ISIC-specific value of the national CF.

The routine carrying the operations was coded in Python 2.7 using the NumPy and SciPy libraries.

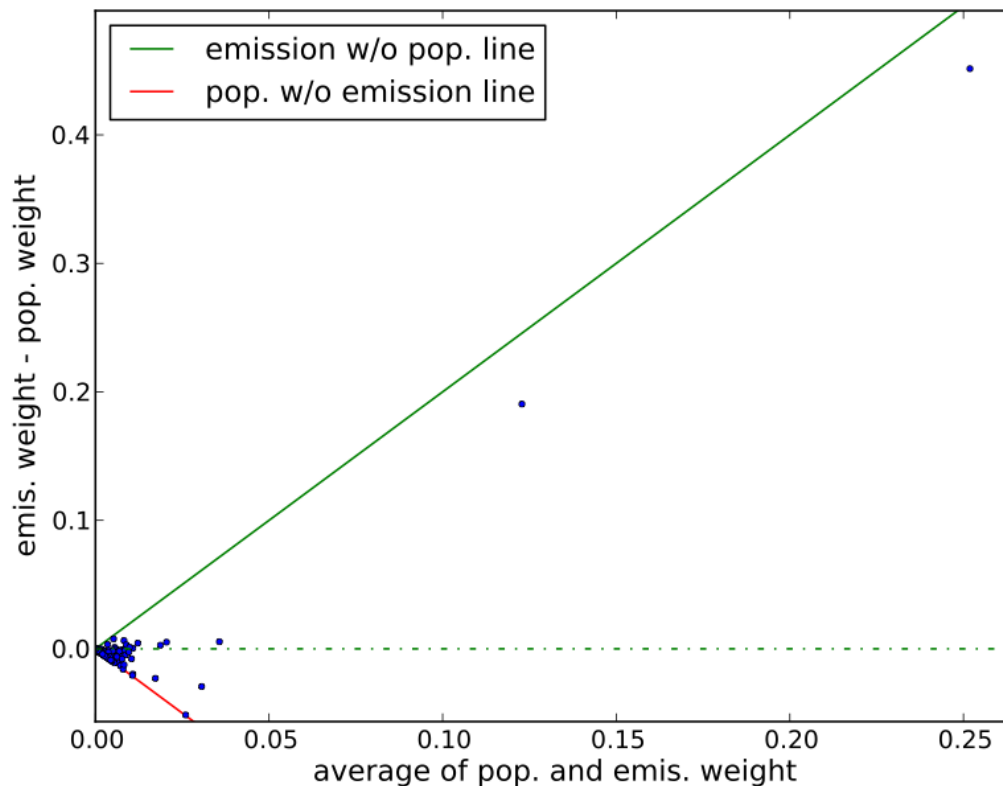
## **3 Results**

### **3.1 Population versus ISIC-generic weighting sets and characterization factors**

The Bland-Altman plot, also known as difference plot or Tukey mean-difference plot, was popularized in 1983 for medical statistics<sup>15</sup>. It aims to assess model agreement.

In a more traditional application comparing two models or measures of the same quantity, a maximum acceptable discrepancy would be set. In the present case, the agreement between the two weighting sets is only an intermediate calculation step: even if a maximum

acceptable discrepancy could be fixed, doing so would be answering the wrong question. In Figure 1, many cells have populations but no reported phosphorus emissions. It is difficult to assess whether this discrepancy is of any concern since each of the weights will be multiplied by their own CFs, which could potentially be many orders of magnitudes apart. Disagreement between weights in cells with low CFs might not produce a significant difference in the aggregated CF if the weights in cells with large CFs agree. The original presentation of Bland-Altman is only of limited interest in this context.

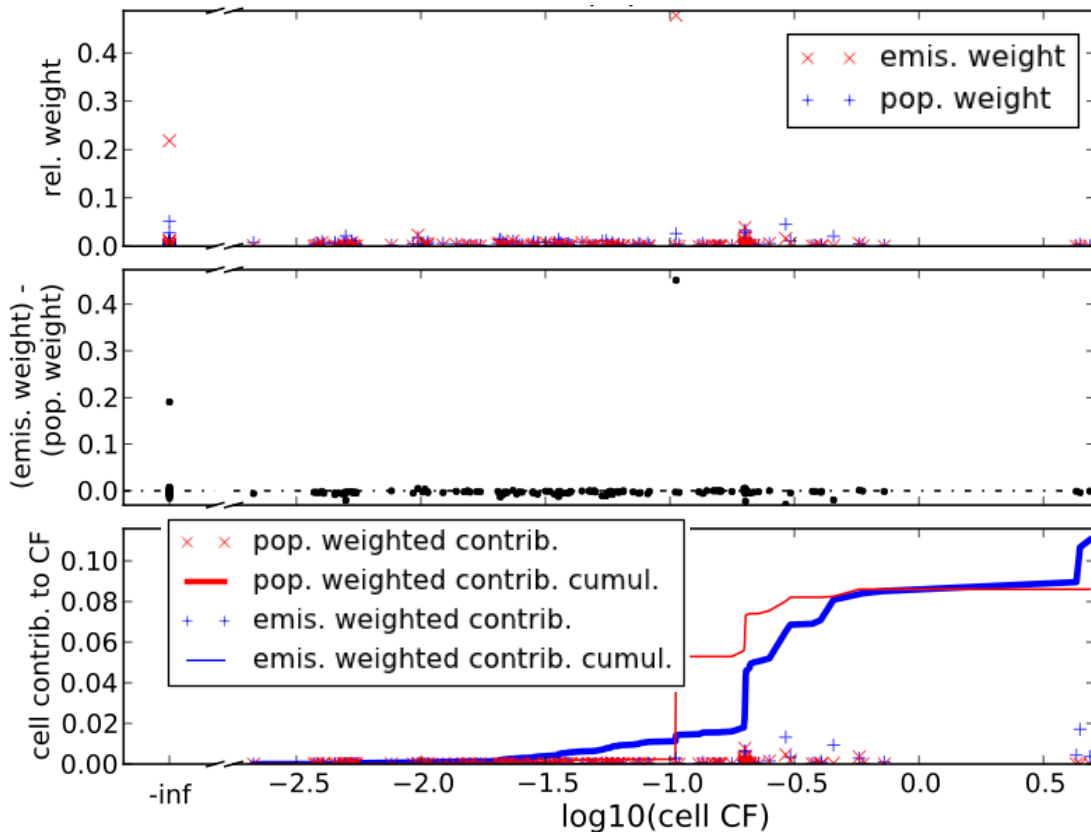


**Figure 1: Bland-Altman plot comparing population and phosphorus emission weights in Italy. The position on the y-axis represents the difference in the emission and population weights for each cell. A cell without inhabitants that reports an emission would fall along the green line. A cell with inhabitants but without an emission would fall along the red line.**

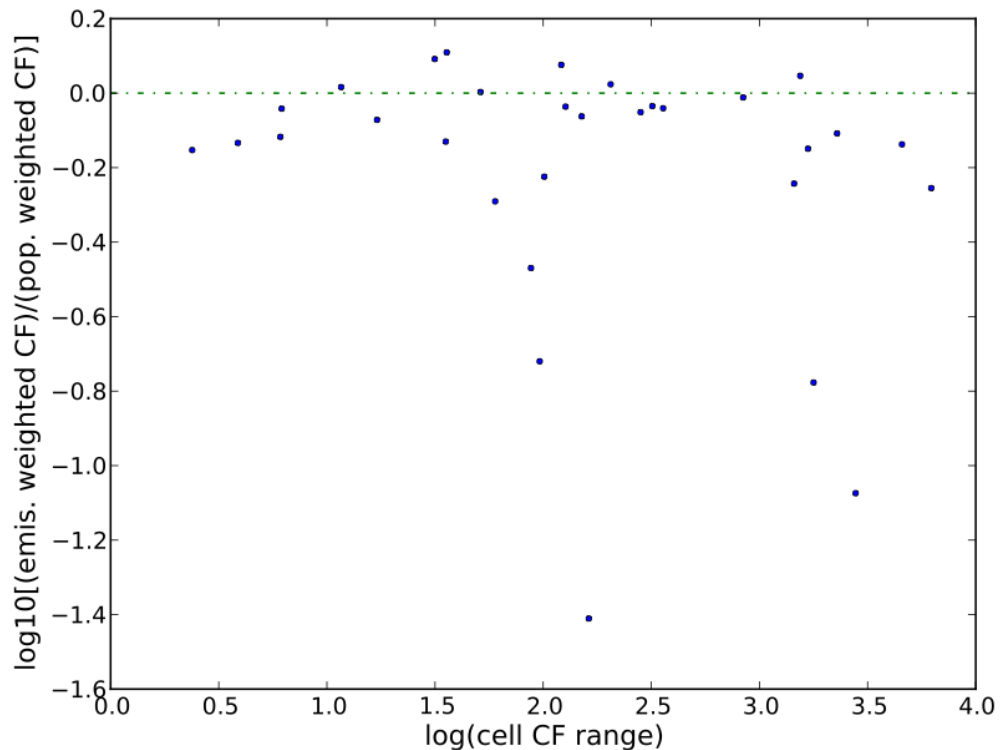
The Bland-Altman plot inspired the presentation in Figure 2. All three subplots share the same x-axis, representing the logarithm of the CF for each cell. On subplot 1, the weights for both sets are plotted. Subplot 2 shows the disagreement of weights: negative values mean that the population weight is greater than the emission weight. On subplot 3, the y-axis

value of each point is the weight multiplied by the CF of the cell—the cell contribution to the aggregated CF.

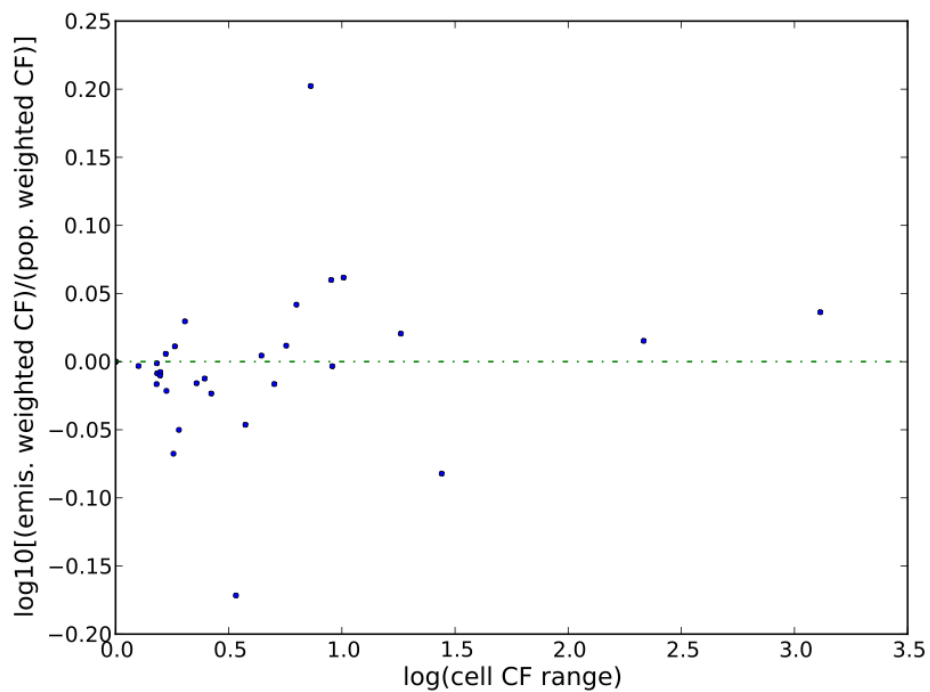
The case of phosphorus emissions in Italy is interesting. Despite one large disagreement (the cell containing the coastal region around Rome) and many small ones, the aggregated CF is almost the same, up until the three cells with the largest CF. These cells are situated in the Alps, where the CF is almost one order of magnitude greater than anywhere else in Italy. No emission is reported for the cells, and only 0.56% of the population resides there. Still, this small difference in weights almost entirely explains the 25% difference in the final value of the aggregated CF.



**Figure 2: Weights and CFs for phosphorus emission and population weighting sets, freshwater eutrophication in Italy**



**Figure 3: Comparison of CFs aggregated with population and phosphorus emission weights, freshwater eutrophication, as a function of the range of cell CFs**

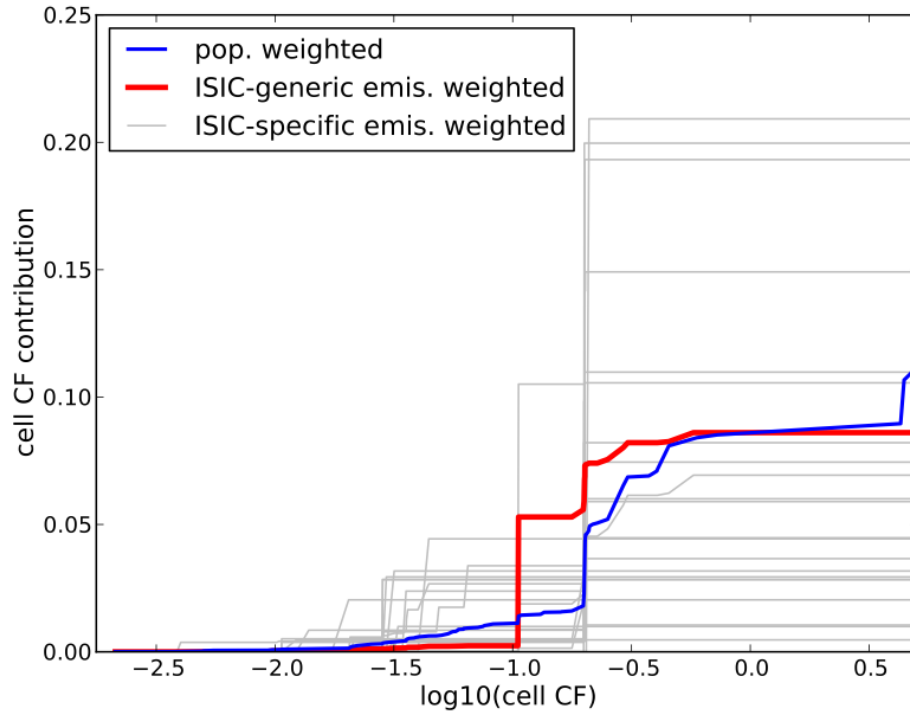


**Figure 4: Comparison of CFs aggregated with population and NH<sub>3</sub> emission weights, terrestrial acidification, as a function of the range of cell CFs**

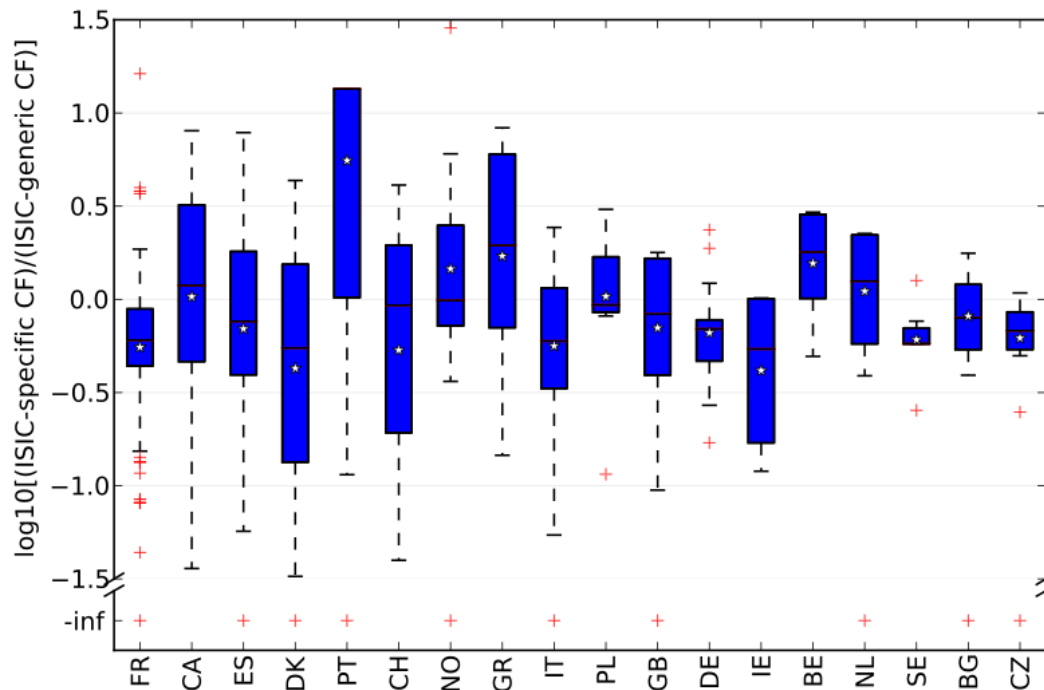
Depending on the country and substance, disagreements between the weighting sets led to small to moderate disagreements between population-weighted and emission-weighted aggregated CFs (more examples are available in the Supporting information). The disagreements between aggregated CFs were summarized in Figure 3 and Figure 4. For the eutrophication impact category, countries covered by a larger range of CFs are more likely to bear a greater disagreement between population- and emission-weighted aggregated CFs. The relationship does not seem to hold for the terrestrial acidification impact category.

### **3.2 ISIC-specific versus ISIC-generic weighting sets and characterization factors**

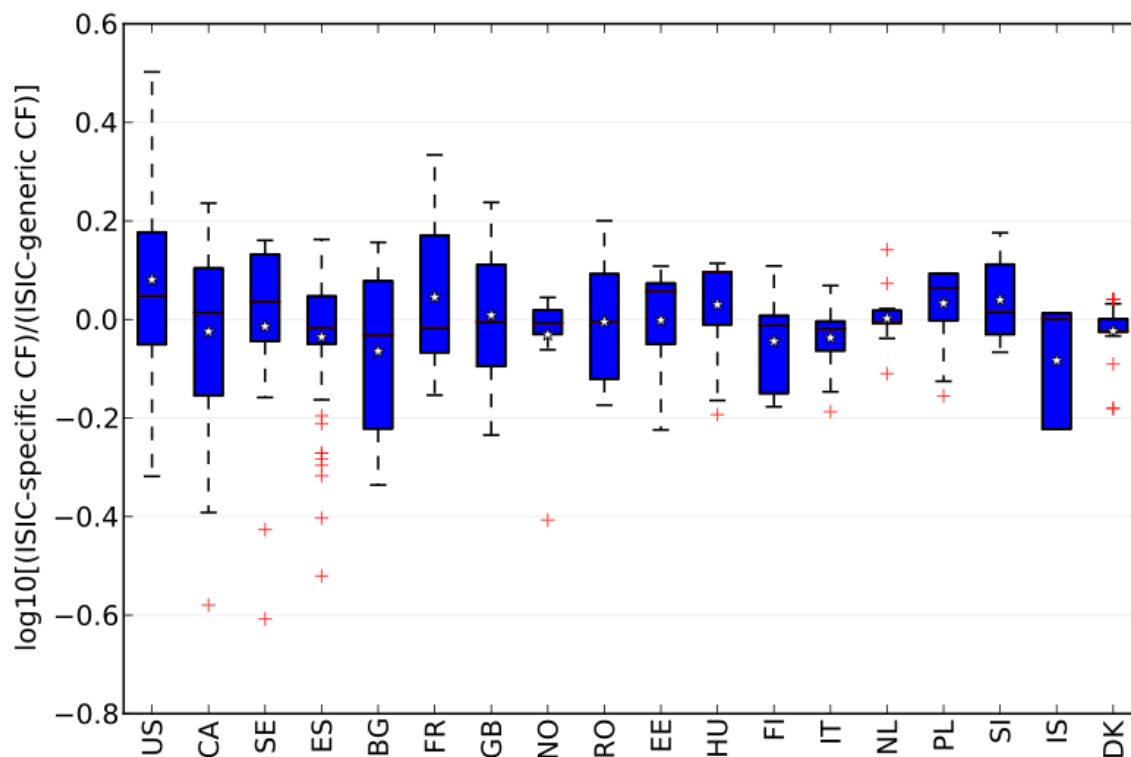
The previous section's plot format was adapted in Figure 5, where only the cumulative curves of the population- and ISIC-generic-weighted CF contributions were kept. The light grey curves represent the cumulative curves of ISIC-specific-weighted CFs. Their shape is very different from the two previous curves: they tend to contain only a handful of large contributors. Industries seem concentrated in certain cells and almost or completely absent in others, leading to ISIC-specific CFs up to one order of magnitude larger or smaller than the ISIC-generic CFs. In the case of Italy, the ISIC-specific CFs above the ISIC-generic CFs include sectors such as sugar, dairy, coke oven production, water treatment, treatment and disposal of hazardous wastes and pharmaceutical manufacturing. The activities are concentrated where the CFs are relatively high. More examples are presented in the supporting information and summed up in Figure 6 and Figure 7.



**Figure 5: Population weighted, ISIC-generic weighted and ISIC-specific weighted CFs for phosphorus emissions in Italy, freshwater eutrophication**



**Figure 6: Range of ratios between ISIC-specific and ISIC-generic CFs for phosphorus emissions, freshwater eutrophication**



**Figure 7: Range of ratios between ISIC-specific and ISIC-generic CFs for NH<sub>3</sub> emissions, terrestrial acidification**

### 3.3 Computational aspects

The established computational structure of LCA<sup>16</sup> would require modification to implement ISIC-specific CFs. This section presents the limitations of the current setup and the necessary modifications and addresses previously voiced concerns on computation time when the system size is increased as much as required by the proposed modifications<sup>17</sup>.

#### 3.3.1 Mathematical formalism

When using the original matrix approach, multiplying the scaling vector with the intervention matrix leads to an inventory vector in which the contribution of all unit processes to each emission are summed. This summing operation prior to characterization violates the assumption of equivalency<sup>18</sup>: emissions from different unit processes have different impact scores because of the spatial variation of the CFs.

**Table 1: Matrices of the computational structure**

Name	Symbol	Content	Dimension	Formula
Technology matrix	<b>A</b>	Technosphere inputs by unit processes	$n \times n$	Supplied by the LCI database
Intervention matrix	<b>B</b>	Emissions and natural resource consumption of unit processes	$m \times n$	Supplied by the LCI database
Demand vector	<b>f</b>	System reference flow	$n \times 1$	Set up by LCA practitioner
Scaling matrix	<b>S</b>	Total amount of intermediary products	$n \times n$	<b>S = diag(A<sup>-1</sup>f)</b>
Inventory matrix	<b>G</b>	Total emissions and natural resource consumption	$m \times n$	<b>G = BS</b>
Unit process to ISIC mask	<b>U</b>	If UP $i$ is part of ISIC $j$ : $U_{ij} = 1$ ; zero otherwise	$n \times p$	Supplied by the LCI database
ISIC-aggregated inventory	<b>G_ISIC</b>	Sum of each elementary flow for each ISIC	$m \times p$	<b>G_ISIC = GU</b>
Characterization matrix	<b>Q</b>	Characterization factors	$m \times p$	Supplied by impact method
Impact matrix	<b>H</b>	Contribution of each elementary flow to each ISIC	$m \times p$	<b>H = element-by-element multiplication of G_ISIC and Q</b>

Alternatives were proposed to avoid the use of the rather impractical sequential method <sup>16</sup>: a hybrid of the matrix and sequential methods <sup>19</sup> and the diagonalization of the scaling vector <sup>20</sup>. We will focus on the latter for the rest of this section. Table 1 summarizes the matrix notation and formulas.

Diagonalization leads to a two-dimensional inventory matrix in which a coefficient  $g_{i,j}$  represents the amount of substance  $i$  emitted by unit process  $j$ . This assumes that unit processes appear as many times as there are regions considered in the system (see section 3.3.2). The intervention matrix has rows for each elementary flow in each region.



A unit process to ISIC mask (**U**) is also required. Coefficients  $U_{i,j}$  are equal to 1 when a unit process  $i$  is tagged as part of the ISIC  $j$  and equal to zero otherwise. The dot multiplication of the mask with the inventory matrix leads to an ISIC-aggregated inventory (**G\_ISIC**).

Finally, a CF matrix is created per impact category (**Q**). Impact score is calculated by the element-by-element multiplication of **Q** with the ISIC-aggregated inventory, followed by summing rows and columns into a single scalar.

### 3.3.2 Computation time testing

Large matrices were created to test the time required to solve the linear equation system. Time testing requires only the presence of non-zero coefficients in the matrices. For the technology matrix, the value is relevant only to avoid positive feedback loops in the supply chain. Detailed assumptions for the construction of the matrices and test parameters can be found in the supporting information.

The simulated system represents a worst-case scenario similar to the one described by Heijungs<sup>17</sup> with a 51 000 x 51 000 technology matrix and a 60 000 000 row intervention matrix (**B**). The use of a linear equation solver package, optimized for sparse matrices, avoids the costly matrix inversion operation. Depending on the final demand vector (**f**), computation times to obtain the scaling matrix (**S**) varied between 51 and 65 seconds. Dot multiplication of the **S** matrix by the monstrous **B** matrix took, on average, 1.2 seconds. Calculations were performed on an average personal laptop computer. Given the fast calculation time for dot multiplication, the next operations were not tested since they also involved dot multiplication between matrices of comparable sizes.

This simple experiment demonstrates that computation time should not be a limitation in the implementation of industry and country specific CFs involving trivial modifications to the matrix approach.

## **4 Discussion**

### **4.1 Case for industry-specific CFs**

As shown in Figure 6 and Figure 7, the relative spatial distribution of emissions of many ISIC sectors can result in large differences between the ISIC-specific and the ISIC-generic CFs. For the phosphorus emission impact on freshwater eutrophication where CFs cover up to four orders of magnitude in a country, ISIC-specific CF values range by up to three orders of magnitude. For impacts on terrestrial acidification, the maximum national variation is of three orders of magnitude, and ISIC-specific CF values range up to one order of magnitude—variations that are far from trivial.

This effect is solely due to the fact that the ISIC-specific CFs are dominated by a handful of cells that can be situated where the cell's CF are above or below the ISIC-generic CF. The absolute quantity that an industry is emitting is not relevant in the calculation. One could even argue in favour of the creation of unit process-specific CFs. This has the potential to exacerbate the geographic concentration effect and create even greater discrepancies.

In the context of a full LCA, ISIC-specific CFs will significantly change the impact score of a system and, potentially, the conclusion of a study only in cases in which emissions originate from industries with large ISIC-specific to ISIC-generic discrepancies.

### **4.2 Prioritizing the creation of industry-specific CFs**

The creation of industry-specific CFs is a data-intensive endeavour. In order to minimize the resources spent on the process, industries and substances could be classified according to their discrepancies relative to ISIC-generic CFs. The idea is to start with the countries and industries that are the most likely to make a difference in an LCA, test whether the differences really occur and proceed to lower priority industries and countries if data and resources are available. Normalization references already contain the total impacts for Europe and North America, and, since they are partially based on the same source of information as this paper, they have the potential to be broken down according to economic sectors for certain substances<sup>12-13</sup>. Normalization references by industry are different from CFs by industry: they are scaled to the economy and point towards industries that, on average, are more likely to cause larger impacts.

### 4.3 Limitations

The algorithms and graphical analysis presented here are valid for any source of emission data. However, the conclusions depend on the data that is selected. The data sources presented here are of good quality, but are known to be incomplete and possibly contain mistakes. Some one percent of the facilities listed in E-PRTR had lacking or obviously mistaken latitude/longitude coordinates (the plant was tagged as part of a country but the coordinates placed it in a neighbouring or remote nation). Phosphorus emissions are non-existent in NEI. In E-PRTR, there are phosphorus emissions for intensive livestock production, animal and vegetable production from the food and beverage sector but only for certain activities, such as installations for intensive poultry or hog rearing, intensive aquaculture, slaughterhouses, treatment and processing intended for the production of food and beverage products from raw animal materials (milk and others), aquaculture, agriculture and livestock. The releases from diffuse sources to water was based on the river basin district scale, 185 large zones dividing the European territory and provided by the CORINE land cover database. Since around 50% of phosphorus emissions to Europe's rivers, lakes and coastal zones are the result of agricultural practices <sup>21</sup>, the lack of data imposes limits on the conclusions that can be drawn.

The other concern is the unknown amount of unreported emissions by facilities whose activity or emission levels are below the legislature-defined thresholds. These thresholds were designed to avoid burdening the database with a large number of data points whose emissions were relatively insignificant. Since the amount of missing information is unknown, it is difficult to assess the magnitude and even the directionality of the bias introduced by their absence. The thresholds were not designed with spatialized CFs in mind, and, as shown above, small changes in emission weights in the most sensitive areas can make a significant difference in the aggregated CFs.

Off-site and waste treatment transfer were identified as such in the three databases used. They were not included in the assessment since the LCI datasets are assumed to report waste treatments as such.

#### **4.4 Complementary sources of information and alternative proxy generation methods**

Databases focused on specific economic activities could generate more reliable weighting sets. For example, in the United States, eGRID provides information on many relevant pollutants at the facility level and has already been used in a spatialization study <sup>22</sup>. Nutrient loss from agriculture could be estimated by connecting land cover databases (for example, CropScale by the National Agricultural Statistics Service in the US) and average fertilizer use for each crop.

The geographic locations of facilities and the nature of their outputs do not constitute sensitive information and could be obtained from producer associations. If there is no available information on the emission rate, the figure can be estimated via LCI databases that contain data on facilities with similar outputs.

At this point, it is also unclear whether the difference observed between ISIC-specific and ISIC-generic CFs is meaningful as compared to the uncertainty of the CF themselves. A statistical analysis must be carried out to test whether the discriminating power of LCA is increased by this complexification. If the difference between the ISIC-specific and the ISIC-generic CFs is not significant for certain sectors, the ISIC-generic could be kept to avoid unnecessary complexification.

## **5 Acknowledgements**

The authors acknowledge the industrial partners of the International Chair in Life Cycle Assessment (a research unit of CIRAIG [Interuniversity Research Centre for the Life Cycle of Products, Processes and Services]) for their financial support: Alcan, Arcelor-Mittal, Bell Canada, Cascades, Eco-Entreprises-Québec/Recyc-Québec, Groupe EDF/GDF-SUEZ, Hydro-Québec, Johnson & Johnson, Mouvement des caisses Desjardins, RONA, Total and Veolia Environnement. We also wish to thank Pierre-Olivier Roy, Andrew Henderson and Christopher Mutel for their manipulation of GIS data.

## 6 References

1. Helmes, R.; Huijbregts, M.; Henderson, A.; Jolliet, O., Spatially explicit fate factors of phosphorous emissions to freshwater at the global scale. *International journal of Life Cycle Assessment* **2012**, *17* (5), 646-654.
2. Manneh, R.; Margni, M.; Deschênes, L., Spatial Variability of Intake Fractions for Canadian Emission Scenarios: A Comparison between Three Resolution Scales. *Environmental Science & Technology* **2010**, *44* (11), 4217-4224.
3. Roy, P.-O.; Deschênes, L.; Margni, M., Life Cycle Impact Assessment of Terrestrial Acidification: Modeling Spatially Explicit Soil Sensitivity at the Global Scale. *Environmental Science & Technology* **2012**, *46* (15), 8270-8278.
4. Boulay, A.-M.; Bulle, C.; Bayard, J.-B.; Deschênes, L.; Margni, M., Regional characterization of freshwater use in LCA: Modeling direct impacts on human health *Environmental science and technology policy analysis* **2011**, *45* (20), 8948-8957.
5. Lenzen, M.; Kanemoto, K.; Moran, D.; Geschke, A., Mapping the Structure of the World Economy. *Environmental Science & Technology* **2012**, *46* (15), 8374-8381.
6. Weidema, B.; Bauer, C.; Hischier, R.; Mutel, C. L.; Nemecek, T.; Vadenbo, C.; Wernet, G., Overview and Methodology: Data Quality Guidelines for the Ecoinvent Database Version 3. Ecoinvent Center: Dübendorf, Switzerland, 2013.
7. Hauschild, M.; Goedkoop, M.; Guinée, J.; Heijungs, R.; Huijbregts, M.; Jolliet, O.; Margni, M.; Schryver, A.; Humbert, S.; Laurent, A.; Sala, S.; Pant, R., Identifying best existing practice for characterization modeling in life cycle impact assessment. *The International Journal of Life Cycle Assessment* **2013**, *18* (3), 683-697.
8. European Pollutant Release and Transfer Register, version 4.2. <http://www.eea.europa.eu/data-and-maps/data/member-states-reporting-art-7-under-the-european-pollutant-release-and-transfer-register-e-prtr-regulation-7>.
9. Regulation (EC) No 166/2006 of the European Parliament and of the Council of 18 January 2006 concerning the establishment of a European Pollutant Release and Transfer Register and amending Council Directives 91/689/EEC and 96/61/EC. European Parliament and Council, 2006; <http://rod.eionet.europa.eu/instruments/615>
10. National Emission Inventory. <http://www.epa.gov/ttn/chief/eiinformation.html>.
11. National Pollutant Release Inventory. <http://www.ec.gc.ca/inrp-npri/default.asp?lang=En&n=4A577BB9-1>.
12. Laurent, A.; Lautier, A.; Rosenbaum, R.; Olsen, S.; Hauschild, M., Normalization references for Europe and North America for application with USEtox™ characterization factors. *The International Journal of Life Cycle Assessment* **2011**, *16* (8), 728-738.
13. Lautier, A.; Rosenbaum, R. K.; Margni, M.; Bare, J.; Roy, P.-O.; Deschênes, L., Development of normalization factors for Canada and the United States and comparison with European factors. *Science of The Total Environment* **2010**, *409* (1), 33-42.
14. Gridded Population of the World Version 3 (GPWv3): Population Grids. Center for International Earth Science Information Network (CIESIN), Socioeconomic Data and Applications Center (SEDAC), Columbia University; Centro Internacional de Agricultura Tropical (CIAT): Palisades, NY, United States of America, 2005.
15. Altman, D.; Bland, J., Measurement in Medicine: The Analysis of Method Comparison Studies. *Journal of the Royal Statistical Society. Series D (The Statistician)* **1983**, *32* (3), 307-317.

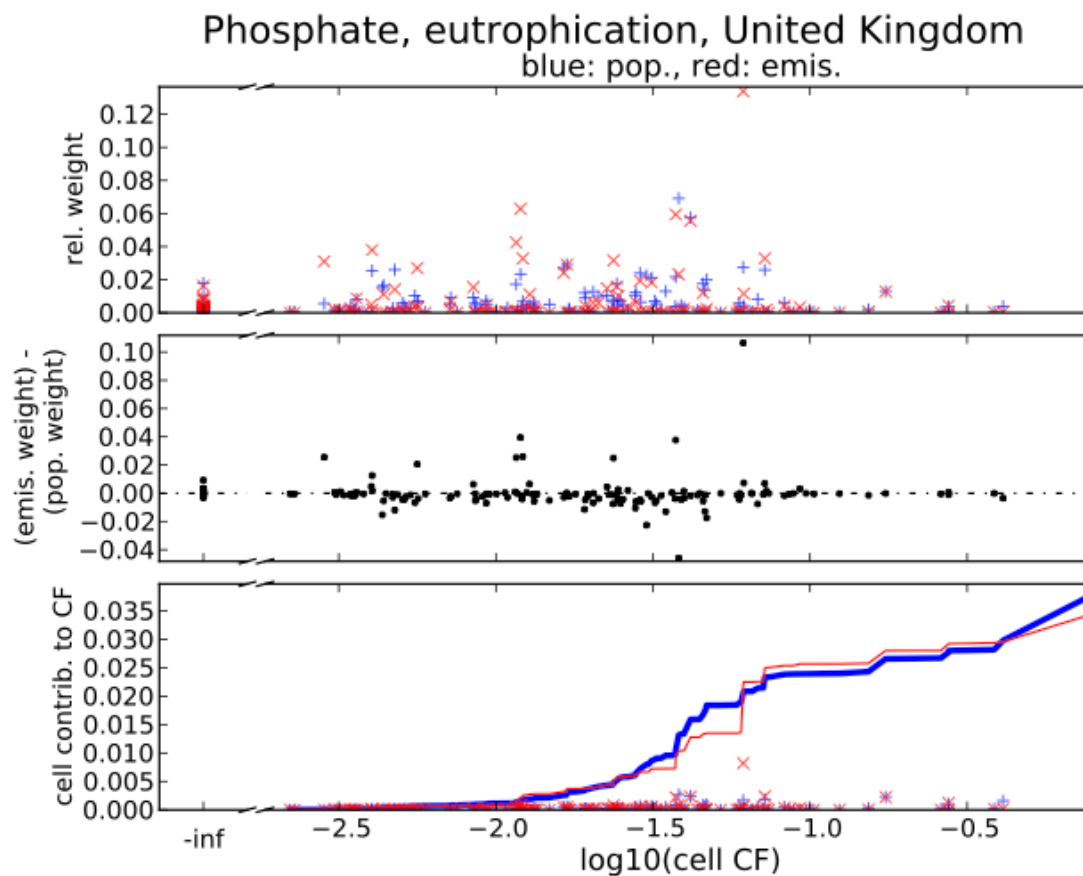
16. Heijungs, R.; Suh, S., *The Computational Structure of Life Cycle Assessment*. Kluwer Academic Publishers: Leiden, 2002.
17. Heijungs, R., Spatial differentiation, GIS-based regionalization, hyperregionalization, and the boundaries of LCA. In *Environment and Energy*, FrancoAngeli, Ed. 2012; pp 167-175.
18. Owens, J. W., Life-Cycle Assessment: Constraints on Moving from Inventory to Impact Assessment. *Journal of Industrial Ecology* **1997**, 1 (1), 37-49.
19. Bourgault, G.; Lesage, P.; Samson, R., Systematic disaggregation: a hybrid LCI computation algorithm enhancing interpretation phase in LCA. *International journal of Life Cycle Assessment* **2012**, 17 (6), 774-786.
20. Mutel, C. L.; Hellweg, S., Regionalized Life Cycle Assessment: Computational Methodology and Application to Inventory Databases. *Environmental Science & Technology* **2009**, 43 (15), 5797-5803.
21. Styczen, M.; Poulsen, R. N.; Panagopoulos, Y. *Diffuse emissions of N and P from agriculture: Data collection and analysis for WISE maps*; European Environment Agency, European Topic Center on Water: 2008.
22. Mutel, C.; Pfister, S.; Hellweg, S., GIS-Based Regionalized Life Cycle Assessment: How Big Is Small Enough? Methodology and Case Study of Electricity Generation. *Environmental Science & Technology* **2012**, 46 (2), 1096-1103.

## Supplementary information for ISIC-specific CFs paper.

### 1 Additional plots

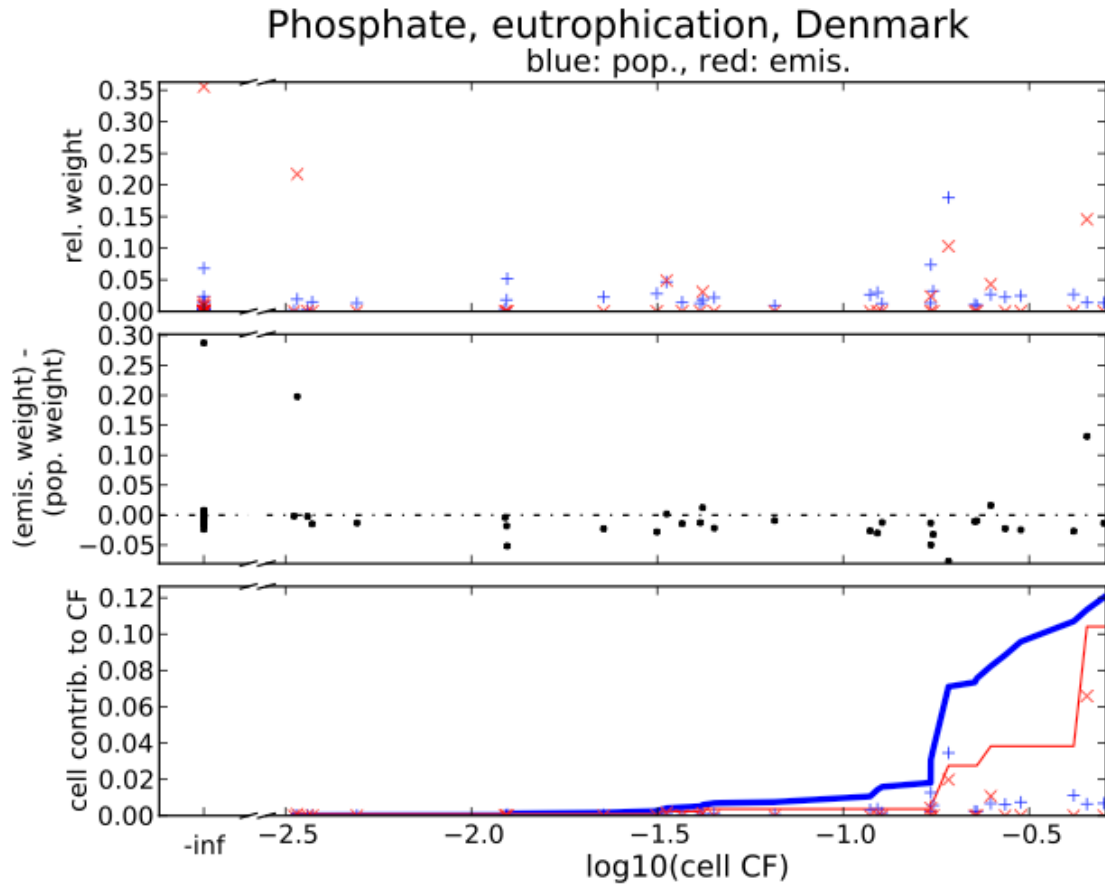
Because of space limitation, only one detailed example (phosphate, freshwater eutrophication, Italy) was shown in sections 3.1 and 3.2 of the main paper. The different countries were exhibiting a limited range of patterns. In this section, one plot from each typical behavior is presented and briefly commented.

#### 1.1 Emission-weighted CF VS population-weighted CF



**Figure 1**

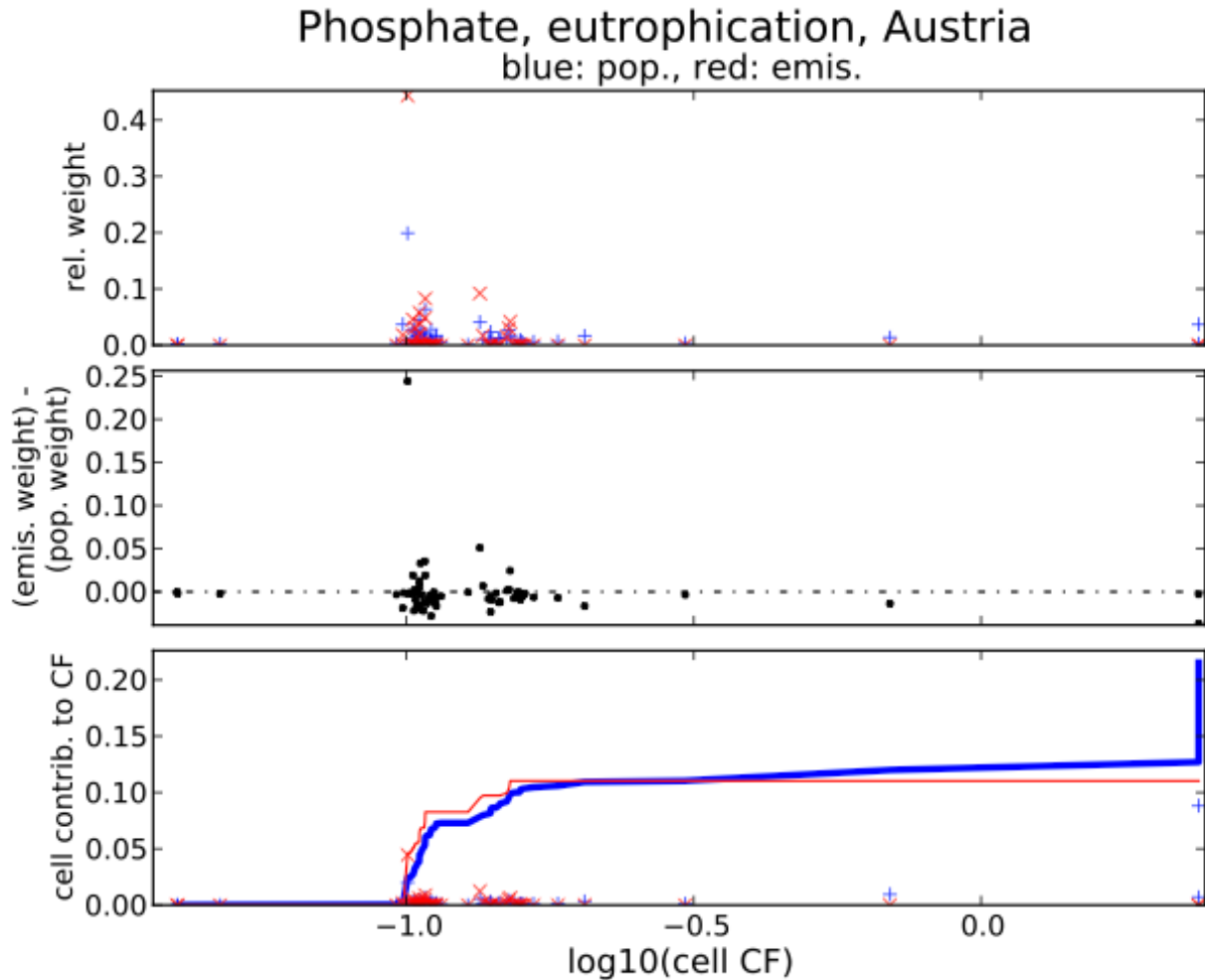
United Kingdom is an example of country where the population and the emissions weights agrees relatively well. The largest disagreement, around 10%, happen in a cell with a CF 1.5 orders of magnitude smaller than the largest cell CF. The result is very small difference between the population and emission weighted CFs.



**Figure 2**

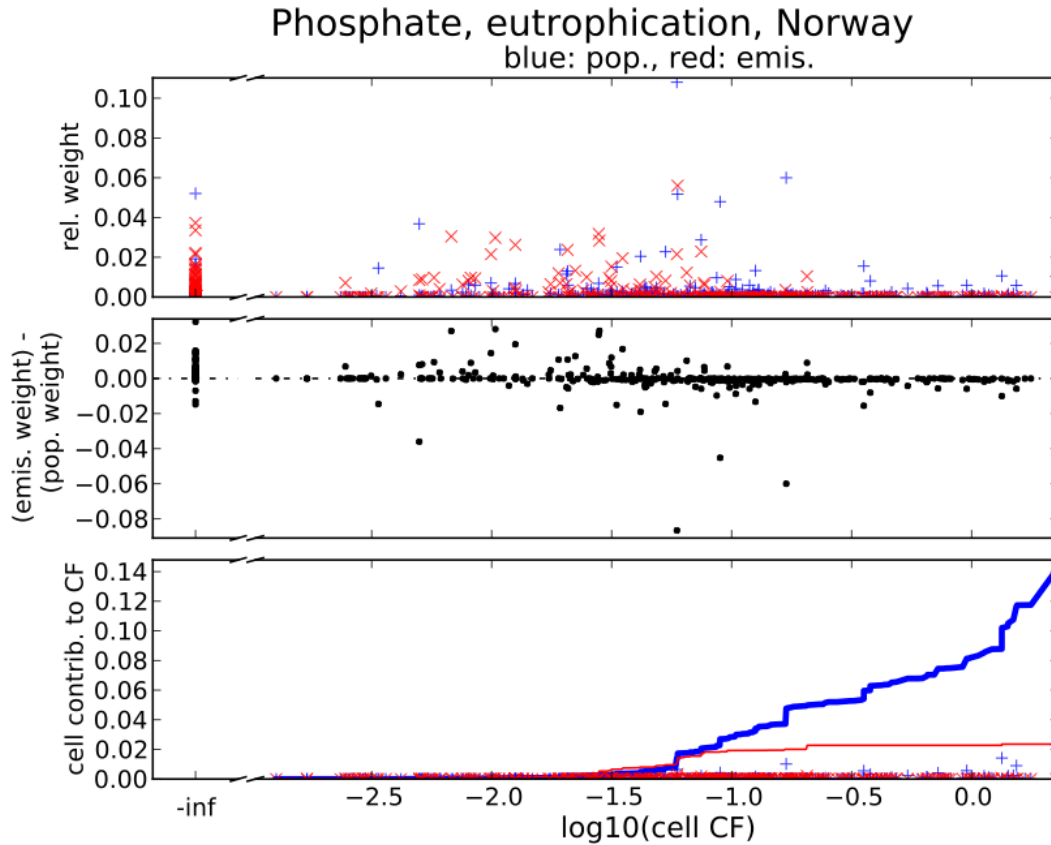
In Denmark, 57% of the emissions happen in the two cells containing Copenhagen and its suburbs, where the CF is equal to zero or very small. Those two cells contain only 8.7% of the population. That could have led to a large difference in the population and emission weighted CF. However, the cell that receives the third largest amount of phosphorus emission (14.5%) also happens to be among the most sensitive to phosphorus emissions. This cell only contains 1.4% of the population. This surprising symmetry brings the two aggregated CFs to a very similar values. This pattern is present in many countries for phosphorus emissions, including Ireland, Finland, Sweden, Switzerland, Bulgaria and Serbia.





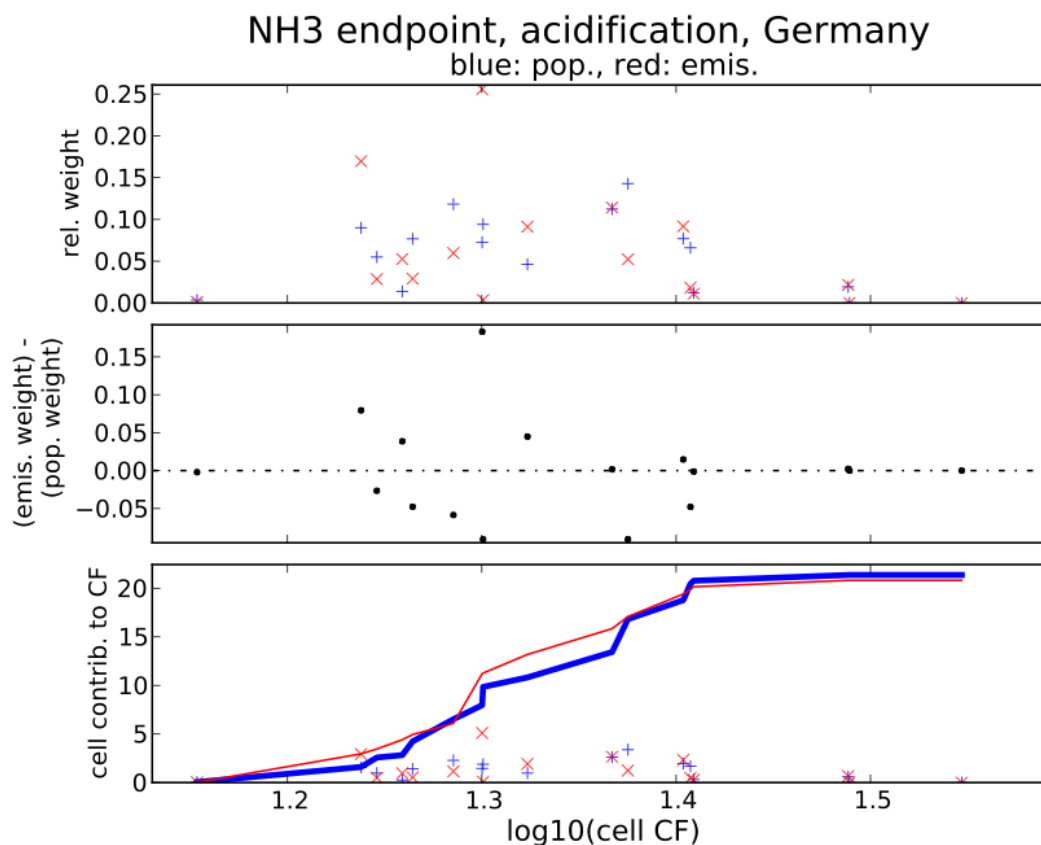
**Figure 3**

Austria is a good example of the importance of correlation of emission and population in cells where the CF is large compared to the rest of the country. The cell containing part of the Constance Lake and the Alps region has no phosphorus emission reported, but is home to 3.7% of the population. This single disagreement, happening in the most sensitive cell of the country, is sufficient to increase the difference between aggregated CFs from about 10%, to about 100%.



**Figure 4**

In Norway, there is no radical disagreement between the weighting sets, and the largest difference do not even occur in the most sensitive cells. However, small disagreements in most cells compound to a large disagreement between the aggregated CFs.



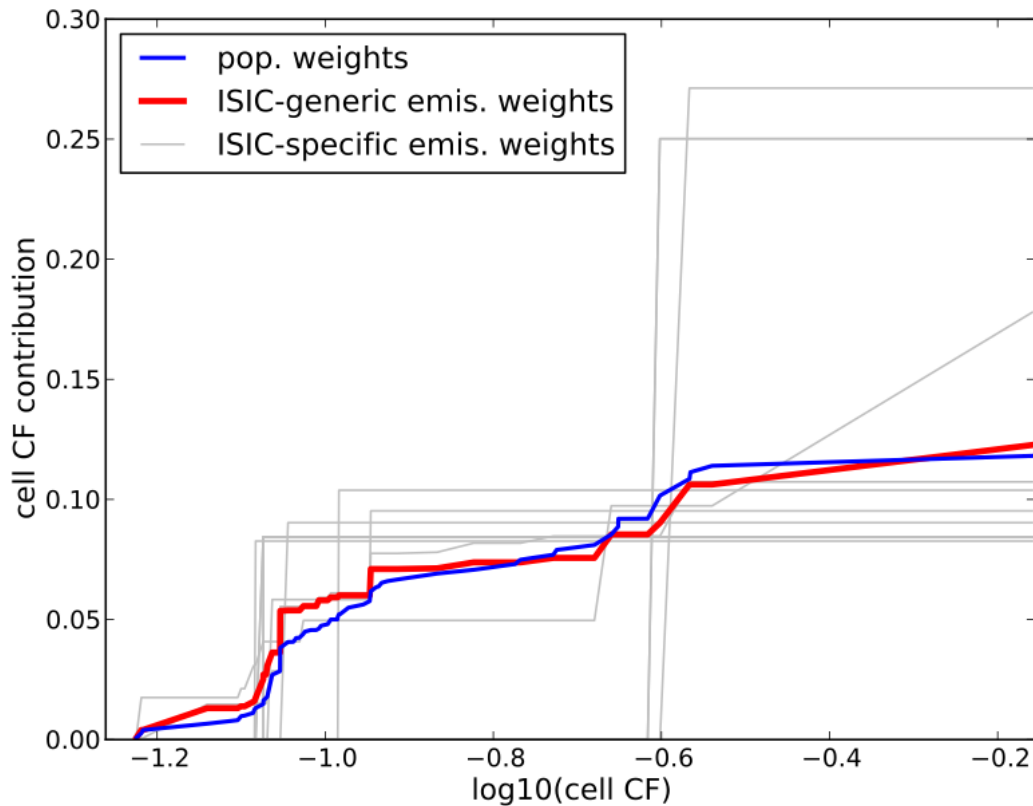
**Figure 5**

Agreement between weighting sets and aggregated CFs are much better in Germany for NH<sub>3</sub> than for phosphorus. This better agreement is not necessarily a sign of structural differences between the two pollutants and impact categories. The range of values covered in the eutrophication category is larger, increasing the chances for exacerbating differences. Cells are six times larger for acidification, potentially averaging larger spatial variations in emissions or population into a single, less extreme value.

Other acidifying substances display similar patterns than for NH<sub>3</sub>.

## 1.2 ISIC-specific VS ISIC-generic CFs

### Phosphate, eutrophication, Hungary



**Figure 6**

Less crowded graph Hungary allows to show more clearly that some ISIC are concentrated in the cells where the population is. (Less interesting stuff to say, those graph are crowded, not much to say that has not already been said. Ideas?)

## 2 Computation time testing

### 2.1 Computing equipment and software

Calculation was carried on a MacBook laptop equipped with a 2.4 GHz Intel Core 2 Duo processor, 2 Go of RAM (an average power personal computer in 2008 standards). The calculation routine was written in Python 2.7, and the matrices were built and solved with the Scipy sparse matrix and linear algebra for sparse matrix packages. More precisely, the Compressed Sparse Row (CSR) matrix format was used, which is optimized for the CSR\*CSR operation (Scipy Reference Guide).

### 2.2 Technology matrix

The number of products/services in the technology matrix (**A**) was set to 1 000 based on the number found in the ecoinvent version 3 (slightly above 900). The number of regions was arbitrarily set to 50. It was also supposed that each products/services was produced in each region, leading to 50 000 unit processes (UP).

UP are not connected directly, but through markets, imitating the structure of ecoinvent v3 (Weidema, Bauer et al. 2013). We assume that each UP is supplied by 10 different markets, which is close to the 10.6 average of economic flow per UP in the ecoinvent v2.2 database (Bourgault, Lesage et al. 2012). The 10 markets were chosen randomly among the 1 000 available, and the demand for each market is set to 0.1 units. Then, we assume that each market was dispatching 0.02 unit of the demand to the 50 regions.

Each element of the diagonal was set to 1. In order to avoid positive feedback loops in the supply chain, off-diagonal elements were chosen in the  $] -1, 1[$  interval. It can be easily demonstrated that this is equivalent a geometric series (infinite sum of  $1/r^n$ ), which has been established to converge for  $|r| < 1$  (Hazewinkel 2001). This is in fact a more stringent condition than the ones shown to be necessary and sufficient in this context (Suh and Heijungs 2007), but it was easier to implement.

**Table 1 : the number of coefficients in the technology matrix**

Formula	Number of coefficients
50 000 UP x 1 diagonal element/UP	50 000
1 000 markets x 1 diagonal element/market	1 000
10 market inputs/UP x 50 000 UP	500 000
50 UP inputs/market x 1 000 markets	50 000
<b>Total</b>	<b>601 000</b>

The result is a very sparse 51 000 x 51 000 technology matrix, containing 601 000 coefficients. The density is  $601\,000 / 51\,000^2 = 2.31 \times 10^{-4}$ .

One hundred final demand vectors (**f**) were randomly generated, each containing a single unit demand for one of the UP. The system was solved with the `spsolve` function, part of the SciPy package.

### 2.3 Intervention matrix

For the intervention matrix (**B**), exaggerated assumptions were chosen, similar to the catastrophic scenario described by Heijungs (Heijungs 2012): 250 chemicals, 12 subcompartments, 50 regions. If we assume that each chemical is potentially emitted in all subcompartments, all regions and all 400 economic sectors of the ISIC classification, the intervention matrix should contain  $250 \times 12 \times 50 \times 400 = 60\,000\,000$  rows. Since markets do not emit or consume resources, the number of columns of the intervention matrix can be equal to the number of unit processes (50 000). Each column is populated with coefficients equal to 1. In the 2.2 version of ecoinvent, the number of coefficients per column is 26 on average. We increase this number to 100, and their rows were randomly selected.

The diagonalized scaling vector (**S**) was the sparse identity matrix, assuming each UP would have a non-zero scaling coefficient. The multiplication was achieved with the `dot` method included in the `csr_matrix` subclass: `B.dot(S)`. The operation was completed in 1.2 seconds.

- SciPy Reference Guide, Release 0.11.0, The Scipy Community, 2012. Available at: <http://docs.scipy.org/doc/scipy/reference/>.
- Bourgault, G., P. Lesage, et al. (2012). "Systematic disaggregation: a hybrid LCI computation algorithm enhancing interpretation phase in LCA." International journal of Life Cycle Assessment **17**(6): 774-786.
- Hazewinkel, M. (2001). Geometric progression. Encyclopedia of Mathematics, Springer.
- Heijungs, R. (2012). Spatial differentiation, GIS-based regionalization, hyperregionalization, and the boundaries of LCA. Environment and Energy. FrancoAngeli: 167-175.
- Suh, S. and R. Heijungs (2007). "Power series expansion and structural analysis for life cycle assessment." The International Journal of Life Cycle Assessment **12**(6): 381-390.
- Weidema, B., C. Bauer, et al. (2013). Overview and Methodology: Data Quality Guidelines for the Ecoinvent Database Version 3. Dübendorf, Switzerland, Ecoinvent Center.

**ANNEXE 4 – Article 4 : « Systematic disaggregation: a hybrid  
LCI computation algorithm enhancing interpretation phase  
in LCA »**

# Systematic disaggregation: a hybrid LCI computation algorithm enhancing interpretation phase in LCA

Guillaume Bourgault\*, Pascal Lesage, Réjean Samson

CIRAIG, Chemical Engineering Department, École Polytechnique de Montréal, P.O. Box 6079, Montreal, Quebec, H3C 3A7

\*Corresponding author: [guillaume.bourgault@polymtl.ca](mailto:guillaume.bourgault@polymtl.ca), 1.514.340.4711 #3129. Fax : +1.514.340.5913

## Abstract

*Purpose* Before the advent of large databases, practitioners often lacked data for calculating life-cycle inventories (LCI), but the actual computation was a straightforward task. Now that databases represent supply chains including feedback loops and several thousands unit processes and emissions, more formalized calculation methods are necessary. Two methods are widely used: sequential method and matrix inversion. They both suffer from different advantages and drawbacks. The present paper proposes a hybrid algorithm combining the advantages of both methods while minimizing their inconveniences.

*Methods* Sequential algorithm requires a form of cut-off criteria, as the supply chains are of infinite length in the presence of feedback loops. The proposed implementation allows the detailing of individual paths until their upstream contribution falls below a user-defined disaggregation criteria, while also allowing the total impact scores of all paths to be stored and considered. The output is then structured to facilitate consultation and re-aggregation, enhancing the work of practitioners in the interpretation phase of LCA. The algorithm is a variation on Structural Path Analysis (SPA) and Accumulative Structural Path Analysis (ASPA). It is computationally efficient and uses a reporting threshold criterion based on multiple impact categories.

*Results* Although the algorithm leads to a more voluminous inventory than matrix inversion, it produces detailed, useful information on the particular instances of processes responsible for the impacts. An average laptop can compute the results within seconds. This algorithm has the potential to improve the interpretation phase of LCA. More specifically, selective replacement of values (characterization factors, input from technosphere or emission intensities) in parts of the process tree can be applied without affecting the rest of the system.

*Conclusions* LCA software would benefit from the inclusion of the algorithm presented in this paper. It produces additional information on the structure of the supply chain and the impacts of its constituents, which would be available for a more in depth interpretation by practitioners. Its potential for understanding the propagation of uncertainty and acceleration of Monte Carlo assessment should also be investigated.

**Keywords:** Structural Path Analysis, LCI algorithm, interpretation phase, disaggregation, matrix computation



## 1 Introduction

Life cycle assessment requires the modeling of a large number of intermediary product exchanges between clients and suppliers. A small team cannot gather the large body of information for each LCA. In practice, generic databases representing the background web of exchanges are used, as the LCA practitioners focus on data collection of foreground processes. Databases that contain information on processes that are at the unit process level (i.e. are not pre-aggregated) usually contain feedback loops, reflecting the fact that supply chains are, in theory, infinite. Special methods are required to compute the impacts of an infinite supply chain. Otherwise, the number of mathematical operations grows exponentially and the amount of information on emissions rapidly becomes unmanageable.

Two types of algorithm are available to solve the linear system to reach the inventory of the whole supply chain of a reference flow: those derived from the matrix inversion, and those derived from the sequential method (described in section 2.1 and 2.2). The latter gives access to a disaggregated inventory. However, it can lead to truncation errors, whereby supply chains upstream of selected nodes are cut-off to reduce computation time, or, conversely, to unacceptably long computation time if one seeks to minimize truncation errors. Matrix inversion, on the other hand, is a fairly straightforward manner to provide an exact inventory for infinite supply chains (i.e. to avoid truncation errors), as those found in our highly intricate economy. However, this approach produces a partially aggregated inventory, which impedes an in-depth analysis of contributions.

A number of papers have presented structural path analysis (SPA, related to the sequential method) as a means to answer specific research questions in the context of Input-Output tables (Defourny and Thorbecke 1984, Lenzen 2002, Peters and Hertwich 2006). However, we will show in Section 2.2 that they are not completely adapted to process LCA. The description of the algorithm used in those papers is usually formalistic and focused on a specific research questions.

We believe algorithms based on the sequential approach are under-utilized in LCA, both in the day-to-day practice and to answer research questions. They have been presented in a very formalistic way that do not make the approach user-friendly, and do not reveal their full potential.

This paper presents a hybrid of the sequential and matrix inversion algorithms. The new algorithm, which we name systematic disaggregation, is described with full detail, and designed in the most general manner to meet the changing needs of different LCA studies, as well as a research tool. The algorithm is inspired by the previously published papers on the subject, particularly Suh and Heijungs (2007) and Suh (2004), and works on the same basic principles. Its originality lies in the structure of the output, allowing greater flexibility in the

interpretation phase. The performance of the algorithm is analyzed. Finally, specific examples are shown to demonstrate how the disaggregated output of the algorithm enhances the interpretation phase of LCA.

## 2 Methodology

### 2.1 Matrix inversion

A very fast and easily implemented way to calculate the total impact of a system is the matrix inversion approach (Heijungs and Suh 2002). An outline of essential variables of this algorithm can be found at Table 1.

**Table 1: Variables in the matrix representation**

Name	Symbol	Content	Dimensions	Formula
Technology matrix	<b>A</b>	Consumption of intermediary products by unit processes	$n \times n$	Normally supplied by the database
Intervention matrix	<b>B</b>	Emissions and natural resource consumption of unit processes	$m \times n$	Normally supplied by the database
Demand vector	<b>f</b>	Reference flow for the system	$n \times 1$	Set up by LCA practitioner
Scaling vector	<b>s</b>	Total quantity of intermediary products	$n \times 1$	$\mathbf{f} = \mathbf{A}\mathbf{s} \Leftrightarrow \mathbf{s} = \mathbf{A}^{-1}\mathbf{f}$
Inventory	<b>g</b>	Total emissions and natural resource consumption	$m \times 1$	$\mathbf{g} = \mathbf{B}\mathbf{s}$
Characterization matrix	<b>Q</b>	Characterization factors	$k \times m$	Normally supplied by impact method
Impact vector	<b>h</b>	Total score for each impact category	$k \times 1$	$\mathbf{h} = \mathbf{Q}\mathbf{g}$

How inverting a matrix account an infinite supply chain can best be demonstrated using power series expansion.

It has been shown that under certain conditions (Suh and Heijungs 2007), the following relation stands:

$$\mathbf{A}^{-1} = \mathbf{I} + (\mathbf{I}-\mathbf{A}) + (\mathbf{I}-\mathbf{A})^2 + (\mathbf{I}-\mathbf{A})^3 + \dots = \mathbf{I} + \mathbf{Z} + \mathbf{Z}^2 + \mathbf{Z}^3 + \dots$$

The basic  $\mathbf{Z}$  matrix is defined as  $\mathbf{Z} = \mathbf{I} - \mathbf{A}$ . Each  $\mathbf{Z}^n$  matrix represents the contribution of each tier (other equivalent expressions found in the literature: “production layer”, “round of the supply chain”) to the scaling vector. Since the infinite sum converges to the inverse of the matrix, using the matrix inversion is analogous to integrating over a supply chain of infinite length.

Matrix inversion produces an aggregated scaling vector. It represents the total of the direct and indirect output each unit process has to produce to supply the reference flows, i.e. the sum of the production of each individual instance, or occurrence, of each unit process. The contributions of each of those instances are automatically summed together: individual contributions are not accessible. This degree of aggregation is carried on along the rest of the calculation, and leads to the total impact scores of the system, without further detail. If only this information is needed, this algorithm fulfills all expectations.

Some more details can be obtained by transforming the scaling vector into a square matrix, where the diagonal corresponds to the scaling vector:

$$S(i,j) = s(i) \text{ if } i = j, \text{ zero otherwise}$$

By substituting this matrix to the scaling vector in the algorithm described in Table 1, a  $m \times n$  inventory matrix and a  $k \times n$  impact score matrix are obtained. The scores have been disaggregated around  $n$  poles, corresponding to the unit processes defined in the database. This allows discriminating which unit processes of the life cycle contribute most to impact categories, but without revealing the position of each instances of the unit processes in the supply chain. This approach has been successfully applied to allow the application of site-dependant characterization factors (Mutel and Hellweg 2009).

## 2.2 Sequential method

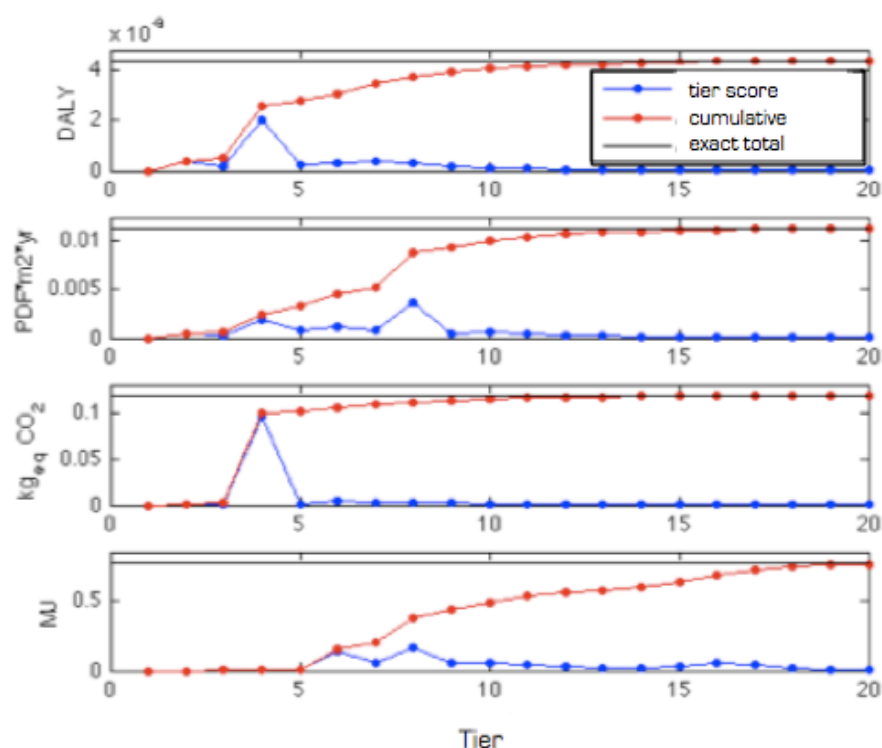
### 2.2.1 Basic implementation

The sequential method is described as an alternative to solve the system of linear equations inherent to LCA, not simultaneously, but sequentially, hence its name (Heijungs and Suh 2002). It is a very instinctive way to present the global scope of life cycle assessment. It goes backwards in the supply chain of a product or service, starting from the reference flows. When a final product is analyzed, the emissions of the site directly producing it will be quantified and stored. Then, the algorithm successively looks at its suppliers 1, 2 ...  $n_0$  and quantifies the output they need to produce to support the production of the reference flow. The emissions of these processes are quantified and stored, and those emissions are multiplied by characterization factors (CF) to calculate the impact score for all the categories considered. The same steps are repeated for supplier 1.1, 1.2 ... 1. $n_1$  ... 2.1, 2.2 ... 2. $n_2$ , etc. A more formal description of the sequential method (described in the context of Structural Path Analysis, see Section 2.2.2) can be found in Peters and Hertwich (2006), or Defourny and Thorbecke (1984).

Theoretically, the supply chain could be extended infinitely. In practice, the impacts of processes very far in the supply chain are often insignificant compared to processes closer to the reference flows. When this is the case, those processes can be safely omitted, although one must keep in mind that the calculated total will inevitably be inexact.

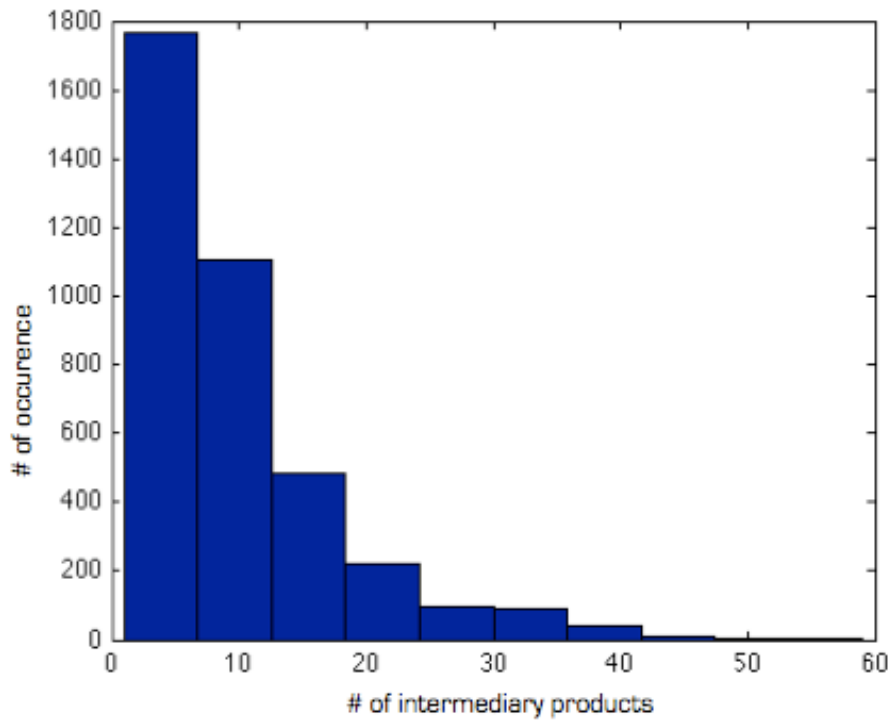
Rigorously defining a criterion for cut-off based on distance from the initial reference flows is a delicate task. One might think that impacts are monotonously decreasing along the supply chain: the further the intermediate supplier relative to the functional unit, the smaller the impact. Power series have demonstrated that this is not the case. Figure 1 shows the distribution of supply chain impacts associated with the production of a concrete block, (LCI data from the ecoinvent 2.0 database, impact scores evaluated with IMPACT 2002+). The stability

of potential impacts in a specific branch of the supply chain over several tiers is not a reliable cut-off criterion, nor is a fixed number of tiers. In this particular example, for the Resource indicator, the contribution of each tier decreases steadily starting tier 8, and reaching at tier 14 an insignificant contribution compared to the cumulative impacts of the previous tiers. Cutting off after observing such a trend seems safe, but the cumulative curve shows that roughly 20% of the impacts, calculated using matrix inversion, are still missing. The calculation has to be pushed further in the supply chain to capture the remaining impacts. The number of tiers necessary to capture a certain threshold of impacts (for example, 95%) varies widely depending on the structure of the supply chain of the reference flows.

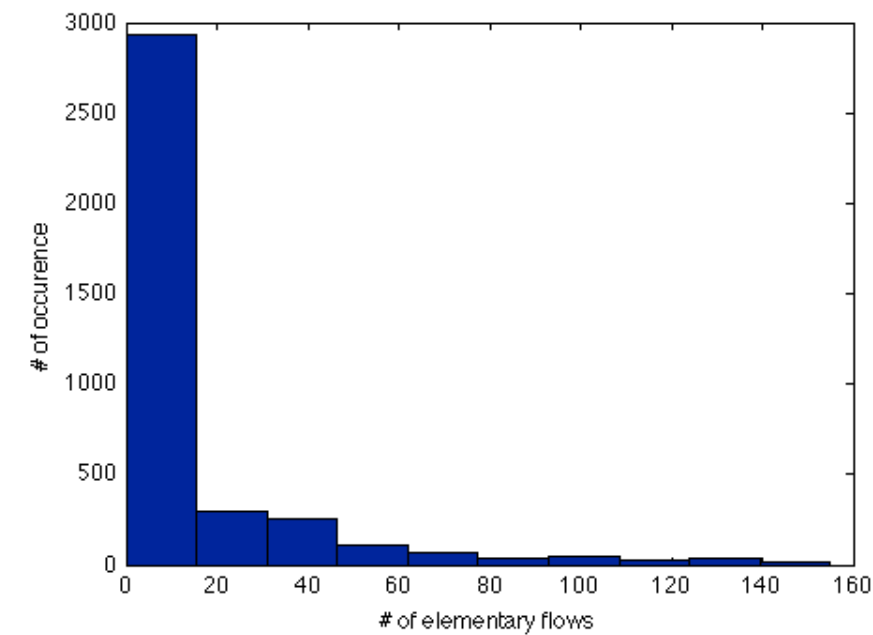


**Figure 1: Repartition of potential impacts along the supply chain for “Concrete block, at plant/DE”, ecoinvent v2.0**

If one aims to minimize cut-off errors, the sequential method suffers from the exponential growth of the number of suppliers. In ecoinvent v2.0 (Frischknecht and Jungbluth 2007), each unit process requires, on average, 9.1 intermediary product flows, each of which is the output of another unit process. Also, each unit process has, on average, 21.6 elementary flows. For MIET, based on the United States 1996 input/output table (Suh and Huppes 2002), those numbers are 205.0 and 318.5, respectively. Figure 2 and Figure 3 show a histogram of the distribution of those two numbers, noted  $z_{mean\_nnz}$  and  $B_{mean\_nnz}$ , for ecoinvent v2.0.



**Figure 2: Number of intermediary product flows per unit process in ecoinvent v2.0**



**Figure 3: Number of elementary flows per unit process in ecoinvent v2.0**

The cumulative number of instances of unit processes and inventory entries can be estimated for product systems truncated after a number  $t$  of tiers:

$$\text{cumulative number of instances of unit process entries in supply chain} \approx (z_{\text{mean\_nmz}})^t$$

$$\text{cumulative number of inventory entries} \approx B_{\text{mean\_nmz}} \times (z_{\text{mean\_nmz}})^t$$

Table 1 shows those two values, calculated with ecoinvent v2.0 and MIET, for several values of  $t$ .

**Table 2: Cumulative number of entries produced by the sequential algorithm with ecoinvent v2.0 and MIET**

$t$	Number of unit processes entries		Number of inventory entries	
	ecoinvent v2.0	MIET	ecoinvent v2.0	MIET
<b>5</b>	6.10e+04	3.62e+11	1.32e+06	1.15e+14
<b>10</b>	3.73e+09	1.31e+23	8.05e+10	4.18e+25
<b>15</b>	2.27e+14	4.75e+34	4.91e+15	1.52e+37
<b>20</b>	1.39e+19	1.72e+46	3.00e+20	5.49e+48

In a power series based analysis of the ecoinvent 2.0 and the MIET databases using the IMPACT 2002+ methodology, it was shown that respectively 16 and 10 tiers are approximately necessary to capture 95% of the impacts of a supply chain (Bourgault, Lesage et al. 2009). Using this information, one can calculate the time and memory required to calculate an inventory with the sequential method. The sequential algorithm was coded in a Matlab environment, which can compute roughly 10 000 unit process entries per second on a regular laptop computer. The information on those 10 000 entries was stored in a variable of type “structure” and occupied 7 megabytes. If calculation had to be carried on to the 15<sup>th</sup> tier with ecoinvent database, only for unit process entries, computation time would be around 720 years, and the result would occupy 10 000 petabytes (peta =  $10^{15}$ ).

The usefulness of such a large quantity of information can be doubted. Although some meaningful contributors to final impact scores can be found up to the 16<sup>th</sup> tier, most branches of the supply chain have a negligible contribution. Computing the whole system would lead to an overwhelming number of unit process instances adding no further information about the impacts of the life cycle.

Considering all these disadvantages, the sequential algorithm is not a viable way of computing the impacts of a product or a service with large databases if one is concerned with cut-off errors.

### 2.2.2 Structural Path Analysis and Accumulative Structural Path Analysis

The sequential method described in the previous section can be used to estimate life cycle inventories and associated impacts, and to identify individual significant contributors. A slight adaptation of the approach allows one to analyze sets of linked nodes, or “paths”, that are of specific interest. This approach, termed Structural Path Analysis, has been used with economic Input-Output (IO) models, which share many similarities with LCA regarding the structure of its quantitative information. Defourny, Thorbecke (1984) use SPA in the context of social accounting. Two poles are chosen and all the path of influence between them are identified. The contribution of each path to the total influence is then calculated, revealing the ones of higher interest. In

LCA terms, it is analogous to choosing two unit processes, listing all the supply chain links between them and calculating how the demand for one unit of the first unit process is transferred to the other unit processes via all the other unit processes of the database. Although of a certain practical interest, this is not typical of the type of information one aims to derive from LCA. Because the contribution of a path decreases with its length, Lenzen (2002), in his use of SPA to identify environmentally important input paths, fixes the maximum path length to 5 tiers. Figure 1 shows that this limit is somehow arbitrary, and that it is certainly not adapted in the context of LCA. Peters and Hertwich (2006) apply SPA to determine how pollution is embodied in Norway's imports. Their approach is more typical of LCA, although they apply their cut-off criteria on a single metric (CO<sub>2</sub> emissions) and a maximal path length (8 tiers). This could lead to ignore significant paths. Being concerned with energy efficiency, Treloar's method uses a "threshold value" on embodied energy to reject insignificant paths.

Suh and Heijungs (2007) have formally described the procedure of using the system score at the end of a path. They make the distinction between SPA and ASPA. In the former, the impacts of each individual occurrence of a unit process are revealed to find the most contributing one, using the unit score. In the later, an occurrence of a unit process is chosen in the supply chain; the impacts of this occurrence and all the other occurrences of unit processes between the chosen one and the reference flow are added, using the system score for the chosen one. The results are then displayed by path length. Suh (2004) uses SPA to show which occurrence of unit processes are the most contributing to the total impact. It enables him to draw conclusions about which unit process is an important direct and indirect source of pollution. This is an example of what would be called re-aggregation in the context of the systematic disaggregation. The proposed algorithm builds on the approach proposed by Suh by making it easy for the user to re-aggregate results around any poles deemed relevant in the context.

One of the key considerations is that the cut-off or relevance threshold should be based on multiple impact categories. Multiplying an inventory by characterization factors will drastically change the contribution of each elementary flow (and of each path), to the overall impact scores. The assumption that paths a certain lengths won't contribute significantly to the total impact is no longer safe in this context, especially, as seen below, for the Resources indicator.

### 2.3 Graphical representation allowed by both approaches

A fictional simplified system has been set up. The **Z** matrix of the system is shown at Figure 4. Units in the matrix are "kg of product/service bought per kg of produced product/service". Figure 5a) has been obtained using SimaPro, which uses matrix inversion, supposing an emission of 1 kg of CO<sub>2</sub> per kg of product supplied.

This representation, known as “network”, becomes hard to read in the presence of feedback loops, but it is the finest resolution that can be achieved when the linear system is solved with the matrix inversion algorithm. Figure 5b) is the “tree” representation of the same system. The arrows in a tree representation always point in the same direction, and instances of a given unit process will appear as many times as necessary. It is much easier to read than the intricate network representation, but it implies the truncation at some arbitrary point in the supply chain (dashed arrows).

It should be emphasized that the notation used in this paper is different than the one of Input/Output analysis. It is more convenient in process-based LCA to use a technology matrix with unit processes normalized to an output of 1 unit. Also, full allocation should be preferred, leading to a square matrix with positive coefficients in the main diagonal (this point is further discussed in section 3.1).



		Demanding process				
		1	2	3	4	5
Supplying process	1	0	0	0	0	0
	2	0.5	0	0	0	0.5
	3	0.5	0.5	0	0.5	0
	4	0	0.5	0.5	0	0
	5	0	0	0.5	0.5	0

		Demanding process				
		1	2	3	4	5
kg of CO <sub>2</sub> eq. emissions/kg of product	1	1	1	1	1	

Figure 4: Z matrix (top) and B matrix (bottom) of a fictional system

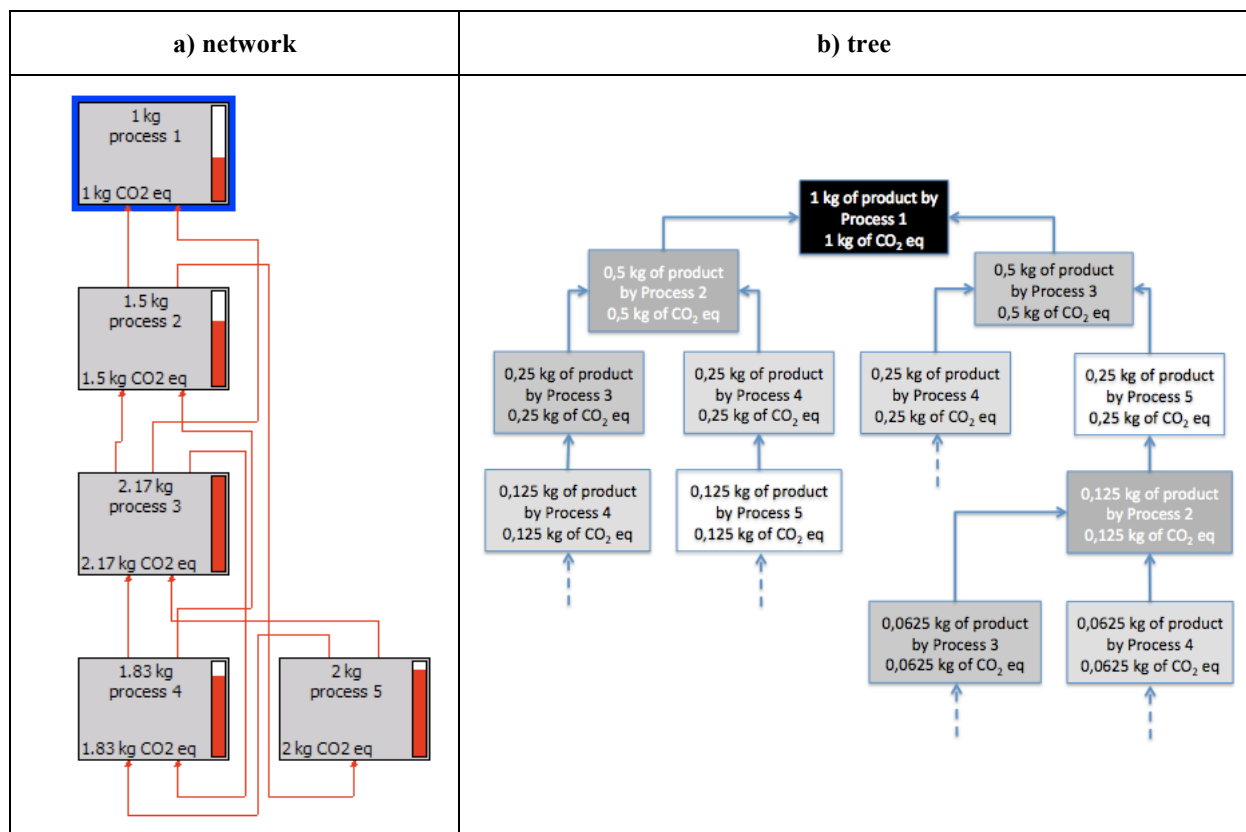


Figure 5: Network and tree representation of the impact of 1 kg of output of « process 1 »

#### 2.4 Adaptation of the sequential algorithm

The purpose of the calculation is to generate sufficient information to adequately interpret the results, relative to the objectives of the study. Regardless of the objectives, finding the main contributors to each impact category is almost always on the list of things to do.

On the one hand, as exposed in section 2.2, after an initial zone of apparent chaos, each new tier of the supply chain will bring an ever-decreasing contribution to impact scores. On the other hand, due to the exponential nature of the sequential method, those contributions will be spread over an ever-increasing number of contributors. Such a large body of insignificant impacts is not necessary to understand the key aspects of the life cycle of a product or a service. It is also not helpful because no action can be taken to tackle each of those sources of impacts: they are outside the sphere of influence of the sponsors of an LCA. Finally, the further down the supply chain a process is, the bigger is the uncertainty, because there is no means to assess the representativeness of the data.

An algorithm producing a large amount of unhelpful and uncertain data after a long computation time is counter-productive. Yet, the method has merit because it allows one to access information necessary to build a tree representation of the supply chain. Consequently, the proposed adaptation of the sequential algorithm focuses on reducing the amount of unnecessary calculations, while retaining all the useful information. We have called it “systematic disaggregation”. As will be seen, it shares many characteristics with SPA as described in e.g. Peters and Hertwich (2007), but with certain adaptations to make the approach generalizable to multiple research questions, allow the calculation of exact results and to use multiple criteria, based on impact category scores, to define the resolution of the results.

The proposed solution is a hybrid between sequential algorithm and matrix inversion. The scheme of the calculation remains identical to the sequential algorithm, but the cut-off problem, mentioned earlier, is solved by comparing cumulative (system) impacts of a path to the impact of the whole system calculated by matrix inversion. If it is smaller than a certain criterion, the calculation is stopped for the specific branch, and the cumulative impacts of the upstream supply chain are stored. Table 3 shows the structure of the result file.

To implement the algorithm, the following steps are applied:

- 1- A pre-calculation is performed, avoiding the repetition of the same operations later in algorithm, hence, saving computation time. The impact of every unit process is calculated for a unit demand, using a final demand matrix equal to the identity matrix of dimension  $n$ . This is done for both cumulative, or “system”, impact (using  $\mathbf{QBA}^{-1}$ ) and unit process specific, or “unit”, impact (using  $\mathbf{QB}$ ). This results in two  $k \times n$  matrices containing the system and unit scores for each of the  $k$  impact categories, for each of the  $n$  unit process. This pre-calculation is specific to the technology matrix used, but not to the product system analyzed, so it can be reused for several analysis as long as it can be described using the same technology matrix (e.g. using the same LCI database).

- 2- A system is built by a practitioner with the unit processes available in the technology matrix. This is done by listing in the final demand vector ( $\mathbf{f}$ ) the quantified reference flows.
- 3- Since this system is a linear combination of elements of the technology matrix, the system-wide score (i.e. the life cycle impacts) can be quantified using the pre-calculated system scores performed at step 1.
- 4- A disaggregation criterion is chosen. This is not the same as a cut-off criterion, the latter implying that the results do not cover 100% of the impacts. As the life cycle impacts are calculated using matrix inversion, the disaggregation criterion does not affect the final result. Rather, it only changes the amount of information that will be revealed in a disaggregated fashion. The smaller the disaggregation criterion, the more detail will be available. It is difficult to choose the appropriate disaggregation criterion before performing any calculation. If the algorithm does not reveal the necessary information, it might be necessary to choose a smaller criterion and perform the calculation again.

**Table 3: Structure of the output table**

Column of the result table	Comment
a) Instance ID number	Unique ID number for each line, allowing the differentiation of instances of the same unit process in the product system
b) Instance ID number of the parent branch	Useful in the construction of the tree
c) Demand	Scaling factor for each instance of a unit process, e.g. how much of the intermediate flow is required, along with the unit (kWh, kg, etc.)
d) Unit process ID	The name and number of the unit process, as described in the database
e.1) Category 1 unit process score	The scaled impact scores associated with each instance of the unit process
e.2) Category 2 unit process score	
e.3) Category $k$ unit process score	
f.1) Category 1 system process score	The scaled impact scores caused by each instance of a unit process and of its supply chain
f.2) Category 2 system process score	
f.3) Category $k$ system process score	
g) Aggregation flag	Flag indicating whether the system score of the instance divided by the total score is smaller than the disaggregation criterion, for all categories. If the score is larger than the disaggregation criterion for all impact categories, it takes the value “disaggregated”, and the instance ID number will appear in later lines in column b. If smaller, then it takes the value “aggregated”
h) Path length	Number indicating how many tiers are found between this instance and the reference flow
i.1) Level 1 unit process ID	Information allowing the visualization of the path without having to refer to a sequence of parent branch number scattered over the entire table
i.2) Level 2 unit process ID	
i.3) Level [level of supply chain] unit process ID	
j) Country tag	Complementary information about the nature of the unit process or the emission, allowing the creation of customized re-aggregation criteria in the interpretation phase.
k) Infrastructure process tag	
l) Process categories (i.e. economic sectors)	
m) Other complementary information tag, customized by the practitioner	

- 5- The algorithm starts with one of the intermediary product flows present in the final demand vector. A unique branch ID is created (Table 3, element a), the demand and the name of the unit process are stored (Table 3, element c and d). The unit and system scores have been calculated for a unit demand at step 1 (Table 3, element e and f). Calculating the unit and system scores for this branch is simply a matter of scaling them according to the demand.
- 6- For each impact category, the system score is divided by the total score of the whole reference flow. The result represents the relative contribution of the impacts of this unit process and its entire supply chain. If this contribution is smaller than the disaggregation criterion for each category, no further details should be computed for this branch and the aggregation flag is set to “aggregated” (Table 3, element g).
- 7- The path length and the path between the branch and the functional unit are stored (Table 3, element g and i). Since the parent branch number is stored, it is easy to retrieve this information. Additional tags can also be stored (Table 3, element j through m).
- 8- Steps 5 through 7 are repeated for each unit process necessary to supply level 1 (reference flows). The algorithm then scans through every branch of level 2. The branches that have been identified as aggregated will not be submitted to any further calculation. To speed up the calculation, the system score of the branch to be evaluated (previously calculated for an arbitrary unit demand in step 1) is scaled by the corresponding  $a_{ij}$  of the technology matrix and the demand of the parent branch. The algorithm presented by Peters and Hertwich (2006) recalculates the score from scratch. We believe this is why they report a calculation time of several minutes to two hours. The approach proposed here is avoiding the unnecessary duplication of calculation. The algorithm will scan each new level of the supply chain until every instance generated receives the tag “aggregated”. The calculation is then completed.

### 3 Results and discussion

#### 3.1 Condition for convergence and analysis

Algorithm based on the sequential method have similar conditions for convergence as those of the power series algorithm (Suh and Heijungs 2007). For the power series to converge, elements in the main diagonal of the technology matrix have to lie in the interval  $[1, 2[$ . For algorithm based on the sequential method, the condition is identical. This is because the  $\mathbf{Z}$  matrix, defined as  $\mathbf{I} - \mathbf{A}$ , is used to calculate the supply for each unit process. The ecoinvent database has a main diagonal equal to 1, so no unit process ever supplies itself. This is not the case in an input/output table. Due to intra-sector sales (for example, cattle food is sold to a bovine farm, both part of the same sector), the main diagonal of the  $\mathbf{Z}$  matrix might not be null. In the MIET database, the

diagonal contains value ranging from 1 to 1,3378. In the network representation, this situation translates as an arrow pointing back towards its sector of origin. In a tree representation, the process will be repeated over and over again with an ever-decreasing demand, as in a fractal. This explains why the  $\mathbf{Z}$  diagonal has to range between 0 and 1. If it would be equal or greater than 1, a stationary or ever-increasing demand of the same process would propagate through the tree and the algorithm would never converge. It is possible to apply the proposed approach to IO-LCA, but columns of the technology matrix have to be rescaled to fulfill the main diagonal condition. Since the computational structure of hybrid LCA is also similar, if not identical, to process LCA, systematic disaggregation can also be to a database constructed with this methodology.

For the power series to converge, it is not necessary that each element outside the main diagonal of the  $\mathbf{A}$  matrix be negative, i.e., that processes should be fully allocated (Peters 2007). Multi-product unit processes are allowed, as long as the technology matrix remains square. However, in the case of incomplete allocation, the term-by-term contribution will include the avoided (i.e. negative) production of other unit processes. This makes the interpretation of results of the power series approach more difficult to interpret. This is not the case for the algorithm based on the sequential method. Each instance of avoided impact can be traced back to the original multi-product unit process. Each possible reference flow of the ecoinvent 2.0 and the MIET database was tested with the systematic disaggregation algorithm and the power series algorithm. Neither database are fully allocated, and the calculation for every reference flow converged to the score calculated with the matrix inversion.

Finally, because of the multiplication of the  $\mathbf{Z}$  matrix by itself, an additional criterion is imposed on the power series algorithm. The spectral radius of the  $\mathbf{Z}$  matrix (the modulus of the largest eigenvalues) should be smaller than unity. This is not a necessary condition for algorithm based on a sequential approach to converge.

### **3.2 Characteristics and originality of the output**

The result of the algorithm does not look like the more familiar scaling vector, inventory and scores per category. The result is a table that allows access to the full details of the supply chain tree and can be consulted in its raw form. It can also be subjected to re-aggregation and filtering to produce easy-to-understand figures.

Systematic disaggregation reports both unit impact (caused by the direct emissions of the instance of the unit process) and system impact (caused by direct emissions of the unit process and the rest of its upstream supply chain). This characteristic has several advantages. First, it is easy to check if the calculation has been performed correctly: the system scores of the instances of unit process that were aggregated and the unit scores of the ones that were disaggregated should sum up to the initially calculated system score of the whole product system, for

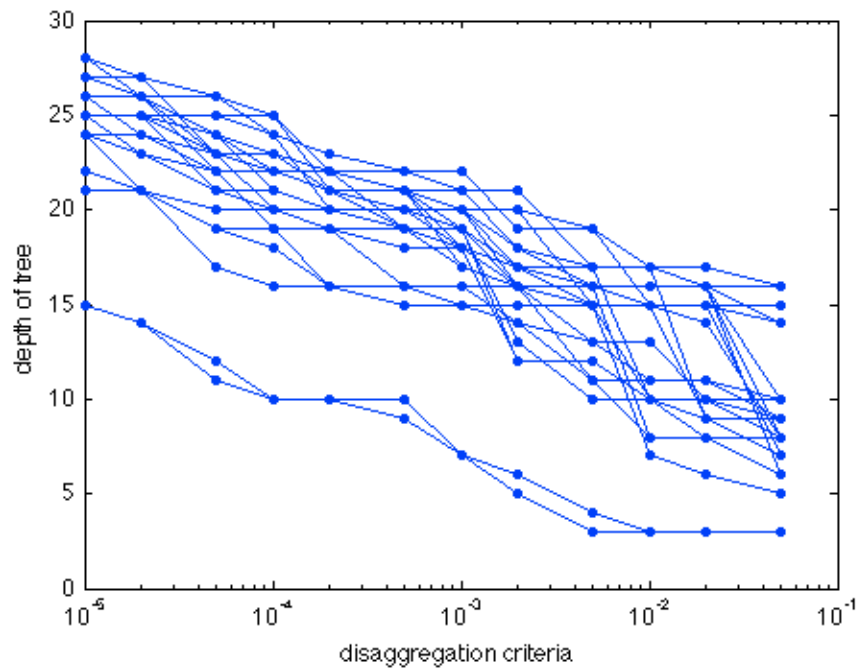
each category. This is why we talk about “disaggregation criterion” instead of “cut-off criterion”. The decision not to dig deeper in the supply chain leads not to a loss of impact, but only a loss of detail on which unit process is causing the impacts. Also, a practitioner might be interested in looking both unit and system scores. It might facilitate understanding, calculation or reporting. When a unit score is high, it points out to a potentially actionable item: it can guide the decision maker (supply chain management decision, eco-design, etc.) and the practitioner (i.e. in primary data collection, see example in section 3.3).

In the short version of the output table (called unit process output table), a line gathers the information on an occurrence of a unit process. This means that the same unit process can appear more than once. This will occur in the presence of feedback loops, or if the same unit process is part of the supply chain of more than one reference flow or intermediary unit processes. This is the main difference between the systematic disaggregation algorithm and the matrix inversion algorithm. In the latter, all the instances of a unit process are automatically aggregated in the matrix inversion operation. They cannot be dissociated. Details on which elementary flows are emitted and how they contribute to the score of the occurrence is not shown.

In its long version (called inventory output table), each line represents a single elementary flow with its associated impact scores. This output is obtained by breaking down the contribution of each elementary flow to the score of each occurrence of unit processes. Which occurrence of unit process is responsible for the elementary flow and the path between the elementary flow and the reference flow are also shown, in addition to all the information present in the unit process output.

The output of the algorithm takes more memory than a traditional inventory list, but is still easily managed by a personal computer. The size of the output depends on the disaggregation criterion. A smaller criterion will lead to a larger output, because the algorithm will have to push further in the supply chain. Figure 6 shows the number of lines of the unit process output as a function of the disaggregation criterion for sampleecoinvent processes. Figure 7 shows the “depth” of the tree, or how many rounds of the supply chain are covered, before the algorithm stops. Figure 8 displays the number of revealed unit processes for a particular round of the supply chain (the “width” of the tree) as a function of the tree depth for different disaggregation criterion. Finally, Figure 9 shows what proportion of the impacts are shown as disaggregated. This quantity is not equal to the disaggregation criterion. If the disaggregation criterion is 1%, each instance of a unit process contributing to less than 1% will be kept aggregated. If, for example, two unit processes contribute to 0,99% of the score, they will both be kept aggregated and their sum will be higher than 1%.

The difference between ASPA as presented in Suh and Heijung (2007) and the current approach is subtle. In ASPA, the impacts of every contributing occurrences of unit processes a path are summed, from the reference flow to the chosen “end of path” occurrence of unit process. This does not allow for a detailed analysis of every contributor on the path. The analysis has to be complemented by SPA and both output have to be interpreted in conjunction of each other. Systematic disaggregation keeps separated the contribution of each occurrence on paths, until the system score contributions become smaller than the disaggregation criteria. The contribution can be quickly re-aggregated to obtain the same results as ASPA using filters and summing the right columns of the output.



**Figure 6: Tree depth as a function of the disaggregation criterion for a random sample of 26 processes of the ecoinvent database**

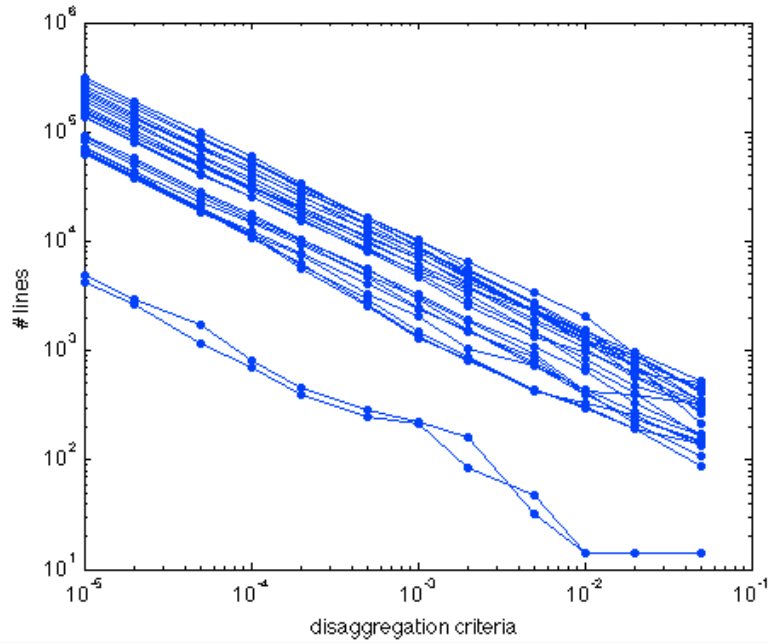


Figure 7: Length of the unit process output table as a function of the disaggregation criterion for a sample of 26 processes of the ecoinvent database

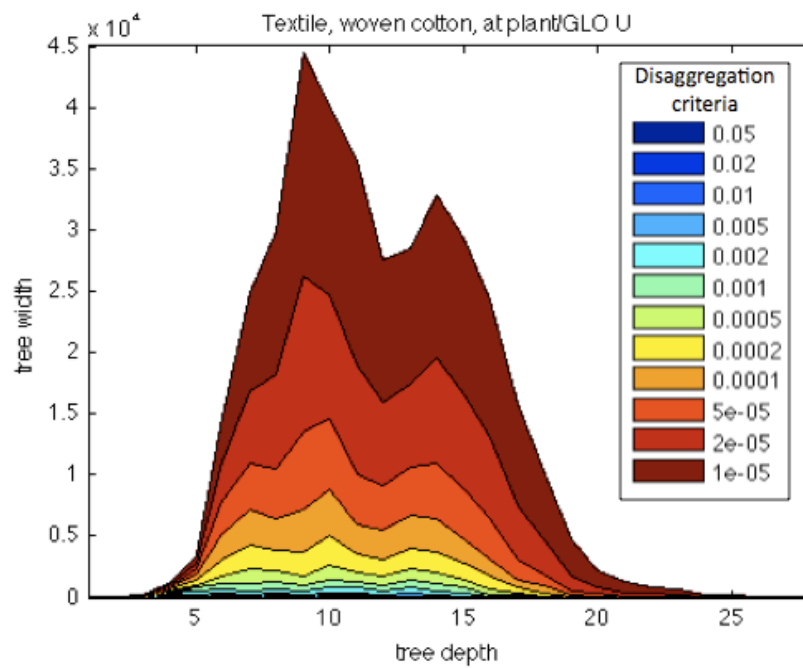
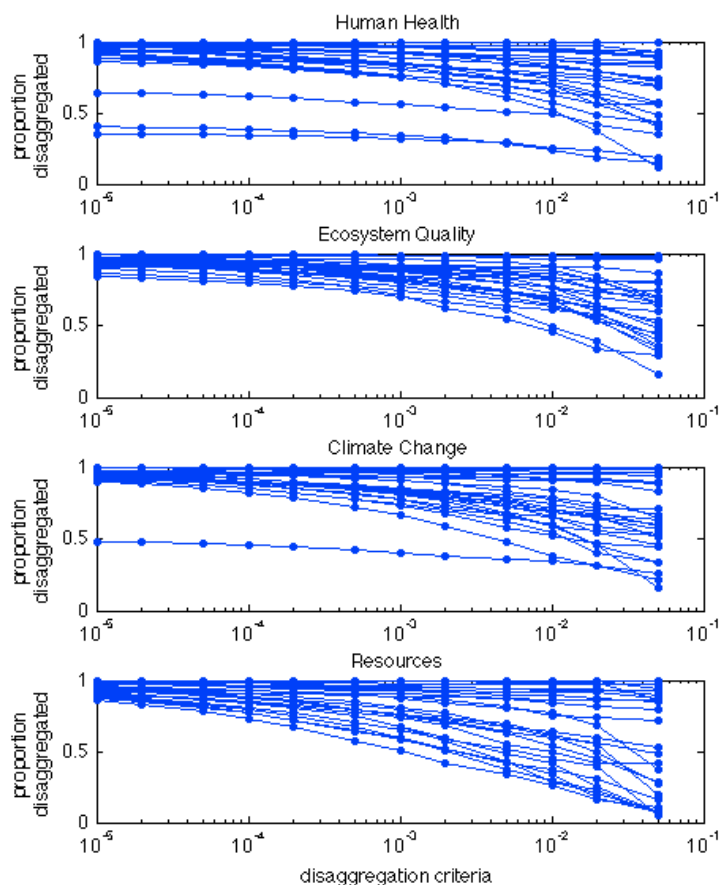


Figure 8: Width of the unit process output tree varying with tree depth and disaggregation criterion for a randomly chosen ecoinvent unit process





**Figure 9: Proportion of impacts disaggregated for a sample of 26 processes of the ecoinvent database**

### 3.3 Application: selective attribution of site-dependant characterization factors

Some impact categories display a high spatial variability. The major contributors to those impacts are often in the background of the model, i.e., they are represented by a generic unit process. If the matrix inversion is used, the major contributions are arbitrarily added to a very large number of contributions from other branches of the supply chain. Those contributions might individually be very small, but because of their very high number, they will add up to a meaningful value compared to the major contributor. It is impossible, and also not necessary, to gather spatial information on the large number of small contributors. It can however be assumed that the pressure will keep building to do so for major contributors, since more and more spatialized CF will be available in the near-future. It is a task that has been retrospectively done for the acidification impact category (Bellekom, 2006), and it is deemed a reasonable demand on practitioners.

After identifying that a unit process is indeed a major contributor, the first task is to disaggregate it in its different occurrences. The systematic disaggregation algorithm allows doing so. Inspection of the paths of each contributor is displayed, and the potential candidates for attribution of site-dependant CF are easily identified. At that point, a new column can be added to the output, where the practitioner will proceed to tag each candidate

for site-dependant CF attribution with a number. Using pivot table, the scores can be re-aggregated according to the tag in the newly filled column. Those left unidentified will be added together by default, and those identified with a specific site number will appear in separate entries. Replacing their site-generic scores by site-dependant scores becomes a straightforward task.

The previous example focuses on selective attribution of CF, but this is a specific application of a more general functionality of the systematic disaggregation algorithm: replacement of values (CF, emission rates, intermediary product demand) affecting only the portion of the supply chain that makes sense. Showing the disaggregated results allows the practitioner to prioritize the primary data collection. By making no assumption and including all potentially useful information in the output, the algorithm creates almost infinite flexibility of manipulation, adapting itself to the context of each study.

### **3.4 Application: reaggregation around Scope 1, 2 and 3**

Let us assume the goal and scope of an LCA include the separation of emissions according to Scope 1, 2 and 3 (as defined by the Greenhouse Gas Protocol) of the production of 1 kg of paper by an integrated paper mill. The company running these activities does not buy wood harvested by another company, but carries on all aspects of the production, from the management of the forest to the final product. According to the GHG Protocol (2004), Scope 1 emissions occur in facilities or devices (such as vehicles) owned or controlled by the company. Scope 2 emissions are the one caused by the production of electricity bought and consumed by the company. Scope 3 emissions occur in facilities that supply the company, but are not owned by it or under its control. Emissions from a wide variety of unit processes should be included in the Scope 1, and the electricity bought by all those activities should count as Scope 2. However, electricity (potentially from the same power plants) consumed by the company's suppliers would fall under Scope 3, and yet be represented by the same unit process in the LCI database. It becomes clear that highly aggregated results as produced by the matrix inversion are not suited to make the necessary distinctions.

Making a demand of 1 kg of "Paper, newsprint, at plant/CH" in ecoinvent will lead to impact scores of  $5.84 \times 10^{-7}$  DALY,  $1.757 \times 10^{-1}$  PDF\*m<sup>2</sup>\*yr,  $8.035 \times 10^{-1}$  kg<sub>eq</sub> CO<sub>2</sub> and  $2.077 \times 10^1$  MJ, when the impacts are evaluated with IMPACT 2002+ endpoint characterization factors. The unit process output generated with a disaggregation criterion of 0.5% leads to 4 269 entries. 30 of those belong to Scope 1, including "Paper, newsprint, at plant/CH" itself, and some occurrences of "Diesel, burned in building machine/GLO", "Operation, lorry >16t, fleet average/RER" and "Power sawing, without catalytic converter/RER". A column called Scope is added to the output and the mentioned occurrences are tagged with a "1" in this column. Although this identification has

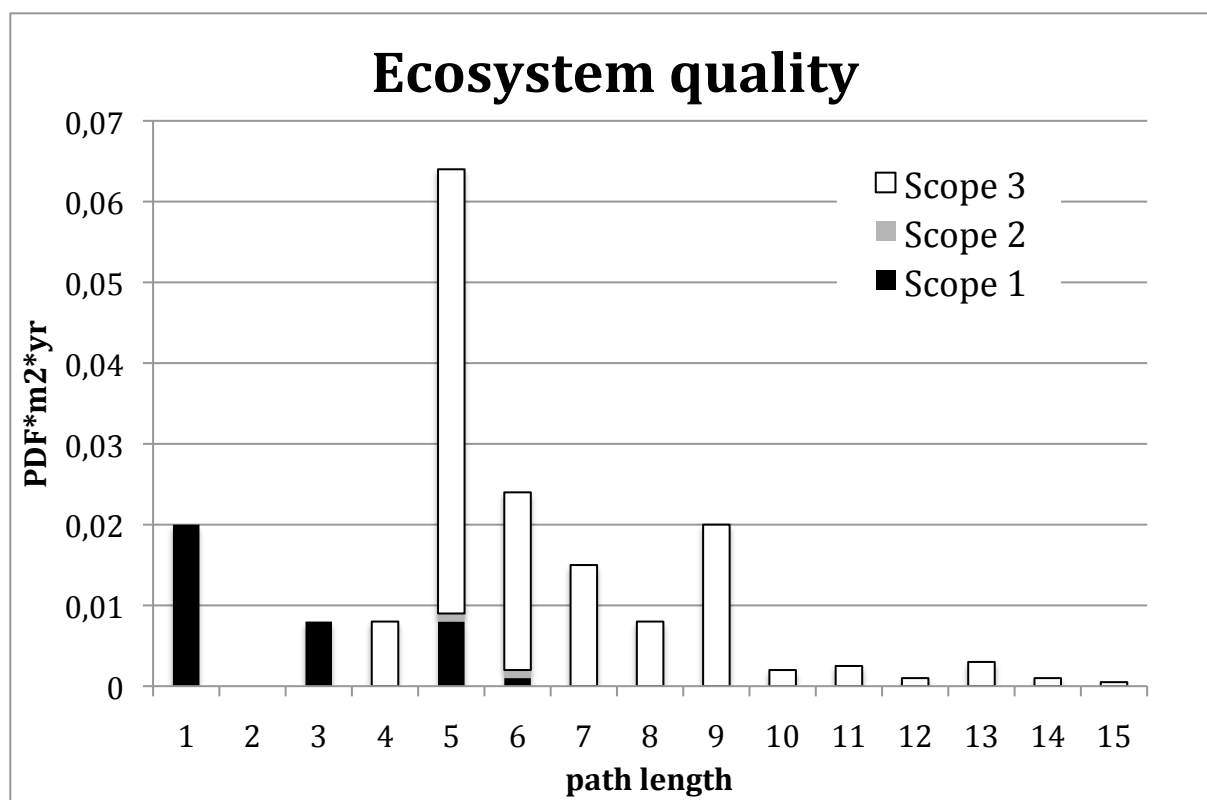
to be done manually, a clever and systematic use of the filters in Excel makes it relatively easy and quick process, with minimal risk of omission.

The scaling vector produced by the matrix inversion informs us that a total of  $2.198 \times 10^{-2}$  MJ of “Diesel, burned in building machine/GLO” is necessary in the whole supply chain. The disaggregated unit processes output reveals 33 instances of this unit process. Among them, 5 have been tagged as part of Scope 1. They amount to  $9.987 \times 10^{-3}$  MJ, or 45.4% of the total demand of this unit process. The rest of the demand occurs elsewhere in the supply chain, for example: “Paper, newsprint, at plant/CH” → “Sodium dithionite, anhydrous, at plant/RER” → “Zinc, primary, at regional storage/RER” → “Diesel, burned in building machine/GLO” which is clearly part of Scope 3.

Identifying instances of unit processes part of Scope 2 is just as straightforward. Filtering to show only energy production unit processes, the instances directly linked to an electricity demand originating from activities within Scope 1 will be attributed a “2” in their Scope column. 107 of them can be tagged as such in the present example. The remaining 4 132 instances of unit processes, neither part of Scope 1 nor 2, fall by default in the Scope 3.

Using Pivot Tables under Excel will aggregate the scores around the newly created tags Scope 1, 2 and 3. A second dimension can be added to create a histogram as in Figure 10. The path length can be used as a second re-aggregation criterion.

For the re-aggregation of the scores, it is necessary to take the unit score of the instances identified as part of Score 1 and 2. Otherwise, impacts of activities occurring outside those activities will be taken into account. Then, the impact of Scope 3 is the difference between the total and what is part of Scope 1 and 2.



**Figure 10: Distribution of impact scores by path length of the supply chain and scope for Ecosystem quality impact category**

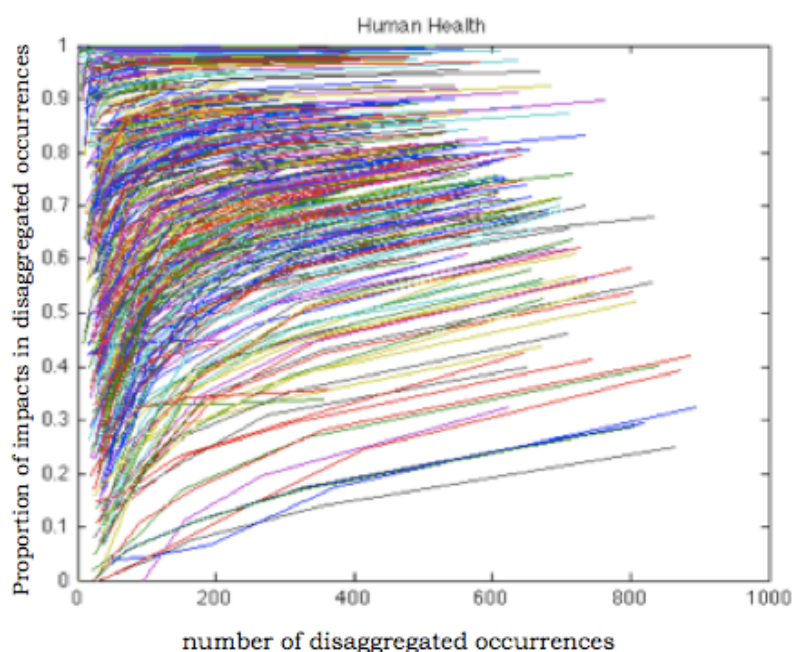
### 3.5 Application: estimation of the number of relevant production sites

As mentioned earlier, systematic disaggregation is also adapted to answer research questions. We show here how systematic disaggregation is used to approach a question about the number of production sites appearing in the model.

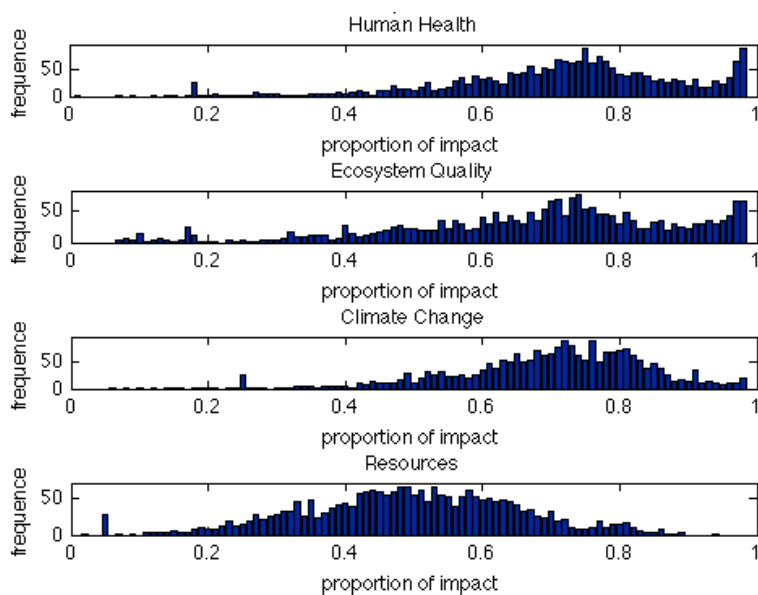
Application of spatialized characterization factors requires that the actual location of an emission is known. In the absence of this information, a site-generic characterization factor should be used (Udo de Haes, Jolliet et al. 1999). A relevant question has since been asked: from how many sites do life cycle emissions originate (Norris 2002)? According to Norris, with an input-output database, pushing the calculation to the 6<sup>th</sup> tier will capture 95% of the criteria air pollutants and toxic releases for 75% of the commodities, which means approximately  $2.8 \times 10^{15}$  instances of unit processes. The number of actual *sites* is potentially much smaller, because the same sites can occur at several instances in a supply chain. Coming back to Figure 5b), the different instances of the same unit process could be representing the same production site, or different ones with a similar output, but it is impossible to determine because the information comes from a generic database. As mentioned in section 3.3, this is an important question: knowing the precise coordinates is necessary to choose the right spatialized CF.

Not knowing means introducing either a bias or uncertainty in the result. Using a streamlined simplification of the database, Norris shows that the 1 000 process instances with the most important contribution will sum up to less than 50% of the emissions for a typical product.

The systematic disaggregation algorithm provides a more precise answer to the question. Specifying criteria of 5%, 2% 1%, 0.5% and 0.1%, the disaggregation is applied to every unit process of the ecoinvent database. For each criterion and process, the proportion of impact generated by disaggregated instances of unit processes is plotted against the number of those disaggregated instances. If the 4 000 curves are plotted at once, the resolution does not allow for a lot of observations. But plotting a random selection of 10% of the curves as in Figure 11 shows that some zones are denser than others. Taking a “cut” at 400 on the x-axis creates a histogram as in Figure 12. The former figure tells us that as a general rule, the effect of revealing more instances of processes in the supply chain is fairly independent of the initial reference flow. Also, after a certain point that appears to be around 300 instances, revealing more instances reveals less and less impacts because of the high dispersion of the impacts over a large number of instances in higher tiers. The histogram shows that for a specific number of disaggregated instances, the proportion of impacts revealed varies wildly: no general rule can be established. It is our opinion that specifying an average, in the presence of such a large variance, would be misleading.



**Figure 11: Impact in disaggregated instances of unit processes**



**Figure 12: Distribution of impact in disaggregated unit processes for 400 instances of unit processes**

## 4 Conclusion

The proposed algorithm is a hybrid between the sequential algorithm and matrix inversion algorithm. It allows the access to details of the supply chain provided by the sequential algorithm, and precision of the matrix inversion algorithm. Compared to previously published algorithm, it fixes a threshold (disaggregation criterion) for significant paths based on multiple impact scores, ensuring the display of all the paths significantly contributing to the total scores. The output of the algorithm can be easily interpreted by simple inspection, or with data mining tools like pivot table and graphs, or database software. With a regular personal computer, the computation is carried on within seconds with some of the largest disaggregated database available so far (ecoinvent and MIET) and for very small disaggregation criteria ( $10^{-5}$ )

The result structure facilitates the interpretation of the life cycle of a product or a service. With minimal adaptation and a bit of imagination, the algorithm can also be used to answer research questions, such as how the uncertainty propagates through the supply chain. For example, since the algorithm displays all the significant paths, which are a minority, Monte Carlo simulation could be applied only to those paths. This would greatly speed the computation.

## 5 Acknowledgement

The authors acknowledge the financial support of the industrial partners in the International Chair in Life Cycle Assessment (a research unit of CIRAIG): ArcelorMittal, Bell Canada, Cascades, Eco Entreprises Québec,

RECYC-QUÉBEC, Groupe EDF, Gaz de France, Hydro-Québec, Johnson & Johnson, Mouvement des caisses  
Desjardins, Rio Tinto Alcan, RONA, SAQ, Total and Veolia Environment

## 6 References

- Bellekom, S., J. Potting, et al. (2006). "Feasibility of Applying Site-dependent Impact Assessment of Acidification in LCA." The International Journal of Life Cycle Assessment **11**(6): 417-424.
- Bourgault, G., P. Lesage, et al. (2009). Enhancing LCA Interpretation with an Algorithm Preserving Total Disaggregation. SETAC Europe: 19th Annual Meeting, Göteborg, Sweden.
- Defourny, J., and E. Thorbecke (1984). "Structural path analysis and multiplier decomposition within a social accounting matrix framework". Economic Journal, Vol. 94, No. 373, pp. 111-136.
- Frischknecht, R. and N. Jungbluth (2007). "ecoinvent report No.1." 1-77.
- Heijungs, R. and S. Suh (2002). The Computational Structure of Life Cycle Assessment. Leiden, Kluwer Academic Publishers.
- Lenzen, M. (2002). "A guide for compiling inventories in hybrid life-cycle assessments: some Australian results". Journal of Cleaner Production, Vol. 10, No. 6, pp. 545-572
- Mutel, C. L. and S. Hellweg (2009). "Regionalized Life Cycle Assessment: Computational Methodology and Application to Inventory Databases." Environmental Science & Technology **43**(15): 5797-5803.
- Norris, G. (2002). "Life cycle emission distributions within the economy: Implications for life cycle impact assessment." Risk Analysis **22**(5): 919-930.
- Peters, G. P. (2007). "Efficient algorithms for life cycle assessment, input-output analysis, and Monte-Carlo analysis." The International Journal of Life Cycle Assessment **12**(6): 373-380.
- Peters, G.P., and E.G. Hertwich (2006). "Structural analysis of international trade: Environmental impacts of Norway". Economic Systems Research, Vol. 18, No. 2, pp. 155-181.
- Suh, S. (2004). "Materials and energy flows in industry and ecosystem networks." The International Journal of Life Cycle Assessment.
- Suh, S. and R. Heijungs (2007). "Power series expansion and structural analysis for life cycle assessment." The International Journal of Life Cycle Assessment **12**(6): 381-390.
- Suh, S. and G. Huppes (2002). "Missing inventory estimation tool using extended input-output analysis." The International Journal of Life Cycle Assessment **7**(3): 134-140.
- Treloar, G. J. (1997). "Extracting embodied energy paths from input-output tables: towards an input-output-based hybrid energy analysis method." Economic Systems Research.
- Udo de Haes, H., O. Jolliet, et al. (1999). "Best available practice regarding impact categories and category indicators in life cycle impact assessment." The International Journal of Life Cycle Assessment **4**(2): 66-74-74.
- World Business Council for Sustainable Development (2004). "The Greenhouse Gas Protocol." 1-116.