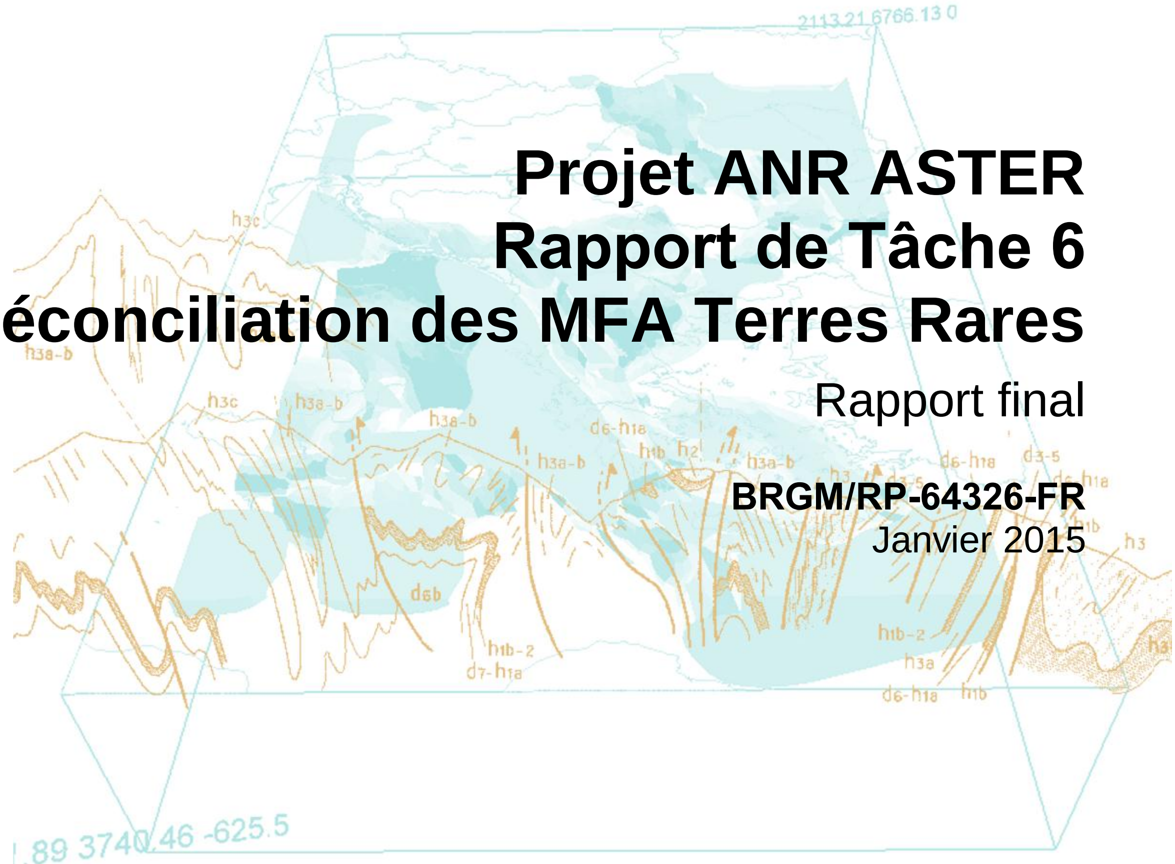
Document public



Projet ANR ASTER
Rapport de Tâche 6
Réconciliation des MFA Terres Rares

Rapport final

BRGM/RP-64326-FR

Janvier 2015

PROJET FINANCÉ PAR L'ANR - PROJECT FUNDED BY THE ANR

ANR

IRIT cnrs • inp • ups institut de recherche en informatique de toulouse

bio Intelligence Service

SOLVAY

Géosciences pour une Terre durable brgm

Projet ANR ASTER
Rapport de Tâche 6
Réconciliation des MFA Terres Rares

Rapport final

BRGM/RP-64326-FR
Janvier 2015

Étude réalisée dans le cadre du projet
ANR-11- ECOT-002

D. Dubois, H. Fargier, D. Guyonnet, R. Leroux, M. Planchon, A. Rollat, J. Tuduri

Vérificateur :	**Approbateur :**
Nom : Stéphane VAXELAIRE	Nom : Jean-Claude GUILLANEAU
Date : 03/02/2015	Date : 3/02/2015
Signature :	Signature :

Géosciences pour une Terre durable
brgm

Mots-clés : Terres Rares, MFA, Flux et Stocks, Economie Circulaire

En bibliographie, ce rapport sera cité de la façon suivante :

Dubois, D., Fargier, H., Guyonnet, D., Leroux, R., Planchon, M., Rollat, A., Tuduri, J. (2015) – Projet ANR ASTER. Rapport de Tâche 6. Réconciliation des MFA Terres Rares. Rapport final. BRGM/RP-64326-FR, 49p., 8 fig., 8 tabl., 4 ann.

Synthèse

Le projet ASTER (ANR-11-ECOT-002) a pour objectif de réaliser une analyse des flux de matière (MFA) de certaines terres rares à l'échelle de l'Union Européenne des 28. S'intéressant aux potentialités de recyclage des terres rares, l'approche considère l'association entre terres rares et applications. Les trois couples considérés sont :

- Terbium, Europium, Yttrium dans les luminophores des lampes basse-consommation ;

- Néodyme, Dysprosium dans les aimants permanents NdFeB ;

- Néodyme dans les batteries NiMH.

Tandis que les Tâches 4 et 5 sont dédiées à la collecte des informations relatives à ces flux et stocks et la Tâche 3 au développement d'une méthodologie de réconciliation des données du MFA, la Tâche 6 objet du présent rapport, met en œuvre cette méthodologie sur les données collectées.

Les résultats sont présentés sous la forme de diagrammes de Sankey, qui fournissent une vision globale des transferts de ces terres rares le long de la chaîne de valeur. Ces diagrammes permettent notamment :

- d'identifier les flux potentiellement recyclables ;

- de mettre en évidence la dépendance de l'UE-28 par rapport aux importations à différents stades de la chaîne de valeur ;

- d'identifier la complémentarité entre ressources primaires (extraites ou potentiellement extractibles) et secondaires (issues de déchets).

Ces informations sont utilisées dans le cadre de la Tâche 7 relative à l'analyse de marché.

Sommaire

Liste des illustrations

Liste des annexes

1. Introduction

Il est rappelé en introduction que le projet ASTER (« Analyse Systémique des TErres Rares – flux et stocks ») vise à établir, à l'échelle de l'Union Européenne des 28, une analyse des flux et stocks de certaines terres rares jugées critiques (NRC, 2008; EC, 2010; Graedel et al., 2012; DOE, 2011). Compte tenu de l'objectif du projet d'identifier des potentialités de recyclage des terres rares, l'approche choisie dans ASTER a consisté à combiner des éléments de terres rares et des applications industrielles qui utilisent ces éléments. Les trois combinaisons terres rares – applications considérées dans le projet ASTER sont :

- Terbium (Tb), Europium (Eu), Yttrium (Y) dans les luminophores des lampes fluorescentes ;

- Néodyme (Nd), Dysprosium (Dy) dans les aimants permanents NdFeB ;

- Néodyme (Nd) dans les batteries NiMH.

L'outil utilisé pour cette analyse ; le MFA (analyse des flux de matière ; Brunner et Rechberger, 2004) consiste notamment à :

- définir les limites du système étudié (dans notre cas les frontières de l'UE-28) ;

- identifier les différents processus impliqués dans le métabolisme de ce système ;

- caractériser les flux entrant et sortant de ces processus, les différences entre les flux entrant et sortant constituant des ajouts ou des retraits aux stocks au sein des processus.

Les processus identifiés et les résultats de la collecte de données pour les flux et stocks dans la technosphère font l'objet du rapport de Tâche 5 (Planchon et al., 2014a, 2014b, 2014c), tandis que la Tâche 4 s'intéresse aux potentialités géologiques (stocks lithosphériques ; Tuduri et al., 2015). A noter que les aimants permanents NdFeB étant utilisés dans un nombre important de produits, l'acquisition d'informations dans ASTER prend en compte cette diversité (lecteurs de disques durs, véhicules électriques ou hybrides, turbines d'éoliennes, ordinateurs fixes ou portables, appareils audio, etc.).

Compte tenu de la diversité des produits contenant des terres rares et aussi de la dispersion des terres rares dans ces produits, les informations issues de la phase de collecte de données sont entachées d'incertitude. Or ces incertitudes sont de nature « épistémique » : elles sont liées au caractère incomplet / imprécis de l'information disponible. Ce type d'incertitude se distingue de l'incertitude dite « stochastique », qui est liée à la variabilité aléatoire des grandeurs qu'on cherche à caractériser (voir en Annexe 1).

Le principe fondamental du MFA est la loi de conservation de la masse. Dans le cas de « n » flux entrant « *EN* » dans un processus et de « m » flux sortants « *SO* », on a :

$$\sum_{i=1}^{n} EN_i = \sum_{j=1}^{m} SO_j + \Delta ST$$

où ΔST est la quantité de variation de stock (positive si les sorties < entrées et négative dans le cas contraire).

Mais en raison, notamment, des incertitudes et erreurs sur les estimations de flux et de stocks, les MFA peuvent ne pas « boucler » et dans ce cas on procède à une « réconciliation » des données. Tandis que des méthodologies de réconciliation du MFA en présence d'incertitudes stochastiques (liées à de la variabilité aléatoire) existent de longue date (minimisation de fonctions objectifs par la méthode des moindres carrés ; Narasimhan et Jordache, 2000), l'objectif de la Tâche 3 a consisté à développer une méthode de réconciliation adaptée à de l'incertitude de type épistémique (liée à des défauts de connaissance). Le rapport de Tâche 3 (Dubois et al., 2013) présente en détail l'utilisation de distributions de possibilité (ensembles flous) pour représenter de l'information incomplète / imprécise et le développement d'une méthode de réconciliation sous contraintes floues. Une synthèse, qui représente la forme la plus aboutie à ce jour de ce travail de développement, est présentée en Annexe 1 du présent rapport, sous la forme d'un article publié dans un journal scientifique à comité de lecture.

L'objectif de ce rapport de Tâche 6 est l'application de la méthodologie, développée dans la Tâche 3, aux données collectées dans le cadre de la Tâche 5. Ces données proviennent d'une multitude de sources d'informations (bases de données statistiques ; EUROSTAT, World Trade Atlas ; rapports d'entreprises qui commercialisent des produits contenant des terres rares ; rapports d'entreprises minières qui explorent des gisements à terres rares ; jugements d'expert, etc.).

Cependant, il faut souligner qu'en raison du stade de développement (durant le projet ASTER) de la méthodologie de réconciliation proposée, une réconciliation « à la main » a été réalisée durant le processus de collecte des informations, afin de « guider » ce processus en vue d'obtenir des MFA globalement cohérents. En effet, un bilan incohérent sur tel ou tel processus constitue une indication qu'un flux manque ou au contraire est trop élevé. Une tâche qui reste à réaliser à l'issue de ce projet a trait à la manière de mieux intégrer la nouvelle méthodologie proposée au sein du processus de collecte des informations.

La méthodologie est appliquée ici aux MFA des couples « substance-application » : Tb-luminophores, Nd-aimants et Nd-batteries. En effet, les MFA des couples Eu- et Y-luminophores ainsi que Dy-aimants, se déduisent des précédents. L'application est réalisée à l'aide d'un code de calcul développé dans un environnement MATLAB et qui est présenté en Annexe 2. Comme il est expliqué dans le rapport de Tâche 3 et en Annexe 1, le processus de réconciliation consiste à trouver les valeurs, ou gammes de valeurs, pour les flux et stocks, qui permettent de respecter les contraintes spécifiées sur ces grandeurs et sur les bilans de masse, avec un niveau de plausibilité globale maximale. Un processus itératif (appelé « Leximin ») est ensuite appliqué dans un deuxième temps afin d'ajuster les valeurs de manière optimale.

2. Applications

2.1. INTRODUCTION

Le point de départ de l'application de la méthode de réconciliation est l'ensemble des données de flux et stocks issus de la phase de data mining (voir rapports de Tâche 5) et exprimées sous la forme d'intervalles flous (Figure 1a) par :

- une borne inférieure du support. Le « support » est l'intervalle de valeurs dont on estime qu'il englobe nécessairement la valeur recherchée. C'est l'intervalle [30000-60000] dans la Figure 1a ;

- une borne supérieure du support ;

- une valeur préférée au sein de l'intervalle support (il peut s'agir également d'un intervalle). Cette valeur est appelée le « noyau » (45000 dans la Figure 1a).

Par défaut et sauf information particulière, la valeur préférée est prise au centre de l'intervalle définissant le support. Comme il est montré dans le rapport de Tâche 3, ce type de représentation est jugé plus cohérent par rapport aux informations dont on dispose dans la pratique, qu'une représentation par des distributions de probabilité uniques (par exemple des distributions normales représentées par une valeur moyenne et un écart type).

En termes de distributions de probabilité, à un intervalle flou correspond une famille de distributions, définie par une limite de probabilité haute et une limite de probabilité basse (Figure 1b). Ce type de représentation est cohérent avec le fait que les informations disponibles ne permettent pas de sélectionner une distribution de probabilité unique, mais néanmoins de fournir des bornes.

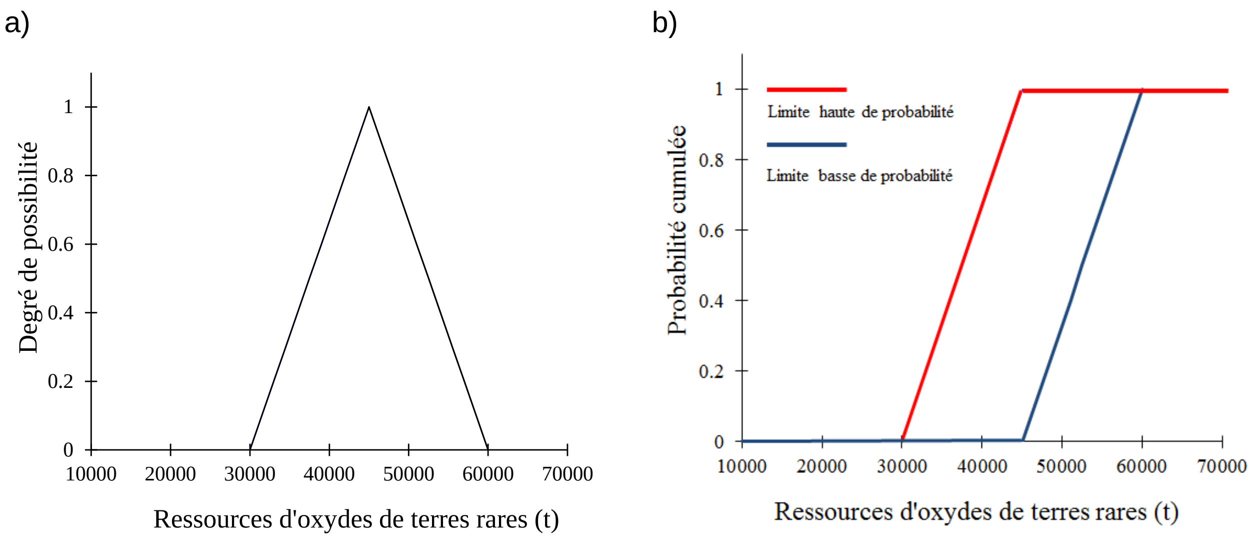


Figure 1 – Exemple d'intervalle flou (a) et limites haute et basse de la famille de distributions de probabilité correspondante

La méthode de réconciliation se base sur les étapes suivantes :

- Analyse de la cohérence entre les équations d'équilibre décrivant les processus du réseau étudié et les données imprécises sur les flux et stocks de matière, interprétées comme des contraintes flexibles à respecter. On calcule le degré de cohérence entre données et réseau ;

- Mise à jour en conséquence des supports des informations sur les flux et stocks ;

- Détermination des valeurs de flux et stocks plausibles, en minimisant les écarts maximaux avec les valeurs plausibles fournies dans les données. C'est un processus itératif (dit « leximin ») qui fixe peu à peu les valeurs de flux et stocks ; les plus contraintes d'abord, puis on recalcule le degré de cohérence sur les valeurs non-encore fixées, on en déduit d'autres valeurs plausibles de flux et stocks, etc., jusqu'à fixer toutes ces valeurs. Cette méthode calcule en fait des valeurs optimales au sens d'une « norme de Tchebychev ».

2.2. LUMINOPHORES

2.2.1. Le système terbium-luminophores

Le système Tb-luminophores a été décrit dans Planchon et al. (2014a) et est rappelé en Figure 2. A noter que dans cette figure, le processus « Lithosphère » est à cheval sur les limites de l'UE-28 (trait tiré), car sont inclus dans « Lithosphère » le bouclier scandinave (avec la Péninsule de Kola) ainsi que le Groenland. Cependant, aucune flèche ne relie la case « Lithosphère » et les autres cases du schéma. En effet, aucune mine en activité dans le processus « Lithosphère » ainsi défini ne contribue de minerai servant à extraire du Tb pour des applications luminophores. Comme il est indiqué dans la section 2.3, une flèche existe pour le cas du néodyme.

Pour les applications « luminophores », l'Europe importe des oxydes de terres rares en mélange et la société SOLVAY (qui a racheté RHODIA en 2011) effectue une séparation en oxydes de Tb, Eu et Y.

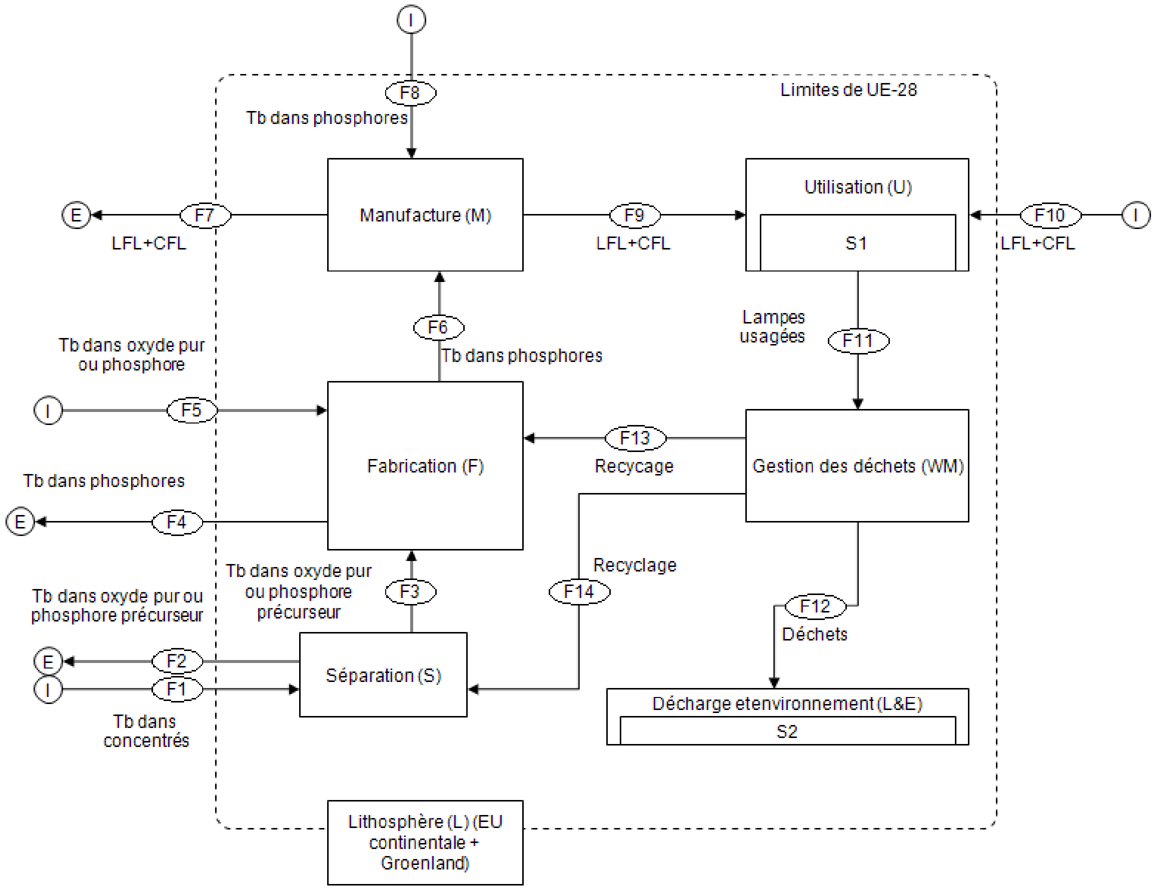


Figure 2 - Schéma du système Tb-luminophores

Notes : LFL = linear fluorescent lamp ; CFM = compact fluorescent lamp : F1 = flux 1 ; S1 = stock 1

2.2.2. Valeurs issues de la Tâche 5

Les données initiales issues de la Tâche 5 et servant de point de départ à la réconciliation sont présentées dans le Tableau 1.

Tableau 1 - Valeurs issues de la Tâche 5 pour Tb-luminophores

Flux/Stock	Description	Borne inf	Borne sup	Valeur préférée
F1	Import vers Séparation	11.0	15.0	13.0
F2	Séparation vers Export	6.0	9.0	7.5
F3	Séparation vers Fabrication	2.0	7.0	4.5
F4	Fabrication vers Export	2.0	5.0	3.5
F5	Import vers Fabrication	4.0	8.0	6.0
F6	Fabrication vers Manufacture	5.0	11.0	8.0
F7	Manufacture vers Export	3.0	12.0	7.5
F8	Import vers Manufacture	10.0	15.0	12.5
F9	Manufacture vers Utilisation	8.0	15.0	11.5
F10	Import vers Utilisation	10.0	30.0	20.0
F11	Utilisation vers Gestion déchets	9.0	13.0	11.0
F12	Gestion déchets vers décharge	9.0	13.0	11.0
S1	∆stock Utilisation	15.0	30.0	22.5
S2	∆stock Décharge et Environnement	5.0	15.0	10.0

Notes : F = Flux ; S = stock ; valeurs en tonnes Tb métal

Pour la réconciliation, les informations de la Figure 2 et du Tableau 1 sont codées dans un fichier d’entrée au format Excel (Annexe 3).

2.2.3. MFA terbium-luminophores réconcilié

Les résultats de la réconciliation sont issus de l’application de la méthode présentée dans Dubois et al. (2014) et en Annexe 1. Une mesure du niveau de cohérence globale entre les données d’entrée est obtenue de la valeur de possibilité (notée α*) calculée lors de la première itération de la réconciliation. A noter que si on obtient $\alpha^* = 0$, cela signifie que les données d’entrée sont incompatibles : pour les gammes indiquées on ne peut trouver aucune combinaison de valeurs permettant de respecter les bilans de masse.

Pour les valeurs du Tableau 1 on obtient $\alpha^* = 0.83$; valeur élevée qui s’explique par la réconciliation réalisée manuellement durant le processus de collecte de données. Durant le processus itératif (Leximin), le niveau α* augmente jusqu’à atteindre un palier (Figure 3).

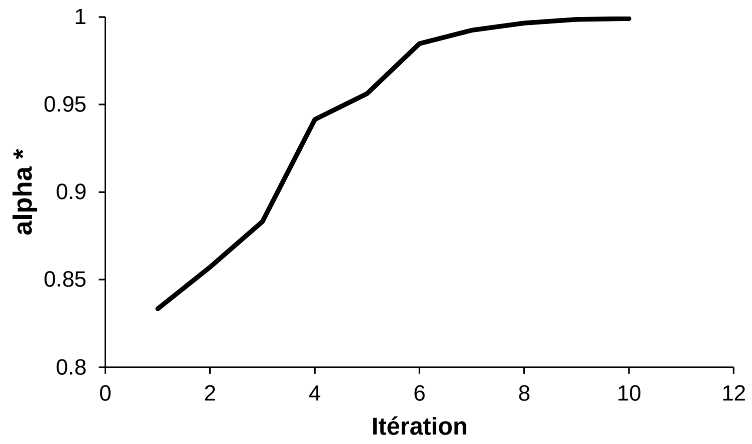
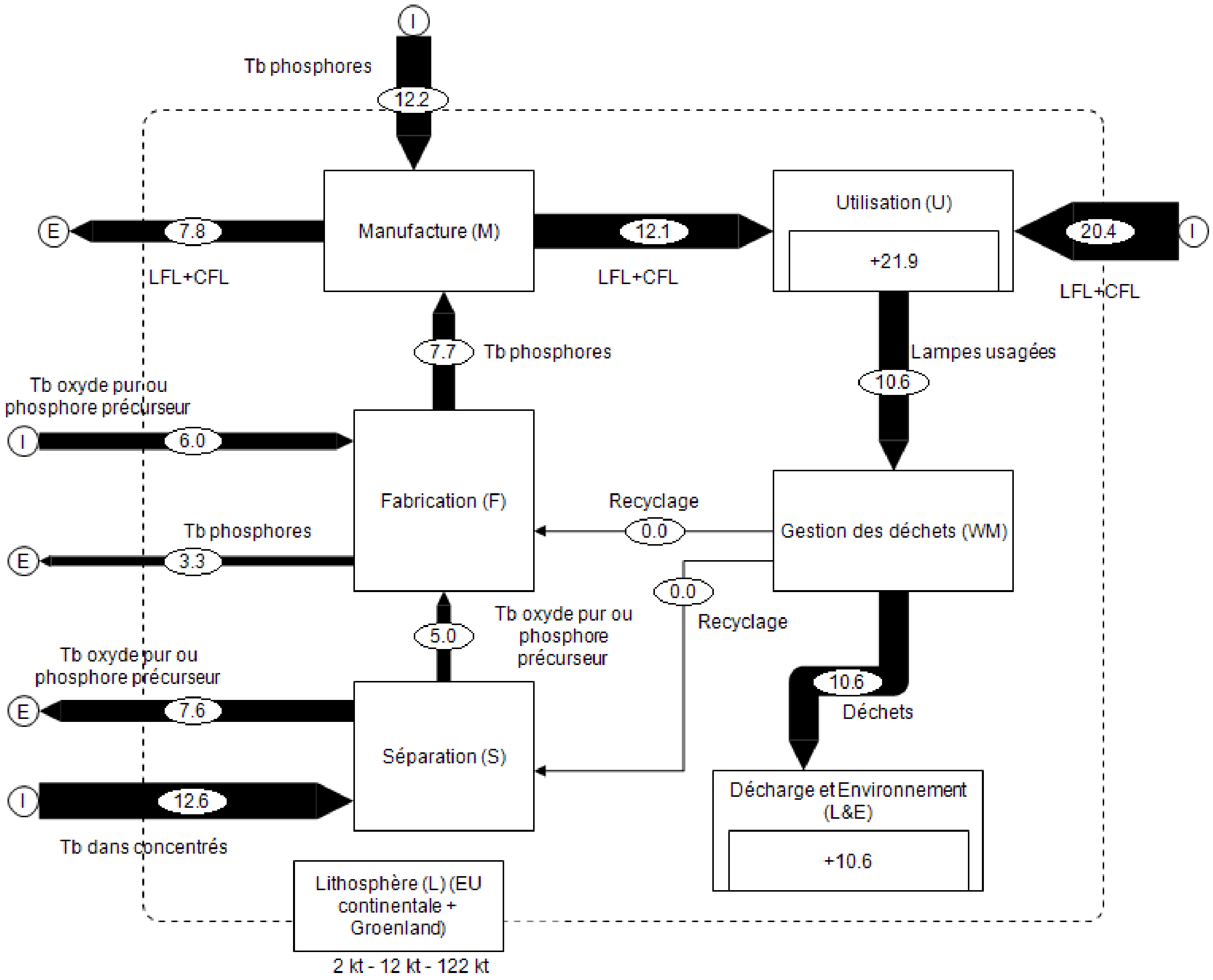


*Figure 3 – Evolution de l'indicateur de cohérence α**

Tableau 2 - Valeurs issues de la réconciliation optimale pour Tb-luminophores

Flux/Stock	Description	Support min	Support max	Noyau
F1	Import vers Séparation	11.0	15.0	12.6
F2	Séparation vers Export	6.0	9.0	7.6
F3	Séparation vers Fabrication	2.0	7.0	5.0
F4	Fabrication vers Export	2.0	5.0	3.3
F5	Import vers Fabrication	4.0	8.0	6.0
F6	Fabrication vers Manufacture	5.0	11.0	7.7
F7	Manufacture vers Export	3.0	12.0	7.8
F8	Import vers Manufacture	10.0	15.0	12.2
F9	Manufacture vers Utilisation	8.0	15.0	12.0
F10	Import vers Utilisation	10.0	30.0	20.4
F11	Utilisation vers Gestion déchets	9.0	13.0	10.6
F12	Gestion déchets vers décharge	9.0	13.0	10.6
S1	∆stock Utilisation	15.0	30.0	21.9
S2	∆stock Décharge et Environnement	9.0	13.0	10.6

Dans le Tableau 2 par rapport au Tableau 1, on constate un ajustement du support de la variation de stock pour « Décharge et environnement » et de légères variations des valeurs de flux. Le diagramme de Sankey correspondant aux valeurs réconciliées est présenté en Figure 4. A noter que les chiffres qui apparaissent sous le processus « Lithosphère » correspondent à des « stocks géologiques » liés à trois projets miniers : Norra Kärr en suède, Kvanefjeld et Kringlerne au Groenland (voir Tuduri et al., 2015). Mais en l'absence de production minière de terres rares en Europe et Groenland, ces chiffres ne participent pas au processus de réconciliation (le processus « Lithosphère » est déconnecté du reste de la chaîne de valeur).



Notes : LFL = linear fluorescent lamp ; CFL = compact fluorescent lamp ; valeurs en tonnes Tb métal

Figure 4 - MFA Tb-luminophores après réconciliation

2.2.4. Analyse de la sensibilité aux valeurs aberrantes

Une analyse de l'effet de valeurs aberrantes a été réalisée en introduisant des erreurs volontaires sur trois des flux du Tableau 1. Ces valeurs sont indiquées en gras dans le Tableau 3. Pour les trois flux F2, F5 et F10, la limite haute du support du flux est fixée à une valeur très élevée (100) tandis qu'on a des différences entre les « valeurs préférées » par rapport à celles du Tableau 1, situées entre 60% et 88%.

Naturellement, l'indicateur du niveau de cohérence entre les informations (α^*) donne une valeur initiale beaucoup plus faible que précédemment : $\alpha^* = 0.14$. L'évolution de la valeur de α^* durant le processus Leximin est présentée dans la Figure 5, tandis que les valeurs issues de la réconciliation sont présentées dans le Tableau 4.

On note une certaine « robustesse » de la méthode de réconciliation, dans la mesure où les valeurs aberrantes ont été « contraintes » par les équilibres générés par les autres valeurs de flux. Les supports sont ramenés à des limites hautes beaucoup plus « raisonnables », tandis que les centres des noyaux réconciliés sont beaucoup plus proches des valeurs du Tableau 2.

Tableau 3 - Valeurs utilisées pour l'analyse de sensisbilité aux valeurs aberrantes

Flux/Stock	Description	Borne inf	Borne sup	Valeur préférée
F1	Import vers Séparation	11.0	15.0	13.0
F2	**Séparation vers Export**	**6.0**	**100.0**	**53.0**
F3	Séparation vers Fabrication	2.0	7.0	4.5
F4	Fabrication vers Export	2.0	5.0	3.5
F5	**Import vers Fabrication**	**4.0**	**100.0**	**52.0**
F6	Fabrication vers Manufacture	5.0	11.0	8.0
F7	Manufacture vers Export	3.0	12.0	7.5
F8	Import vers Manufacture	10.0	15.0	12.5
F9	Manufacture vers Utilisation	8.0	15.0	11.5
F10	**Import vers Utilisation**	**10.0**	**100.0**	**55.0**
F11	Utilisation vers Gestion déchets	9.0	13.0	11.0
F12	Gestion déchets vers décharge	9.0	13.0	11.0
S1	∆stock Utilisation	15.0	30.0	22.5
S2	∆stock Décharge et Environnement	5.0	15.0	10.0

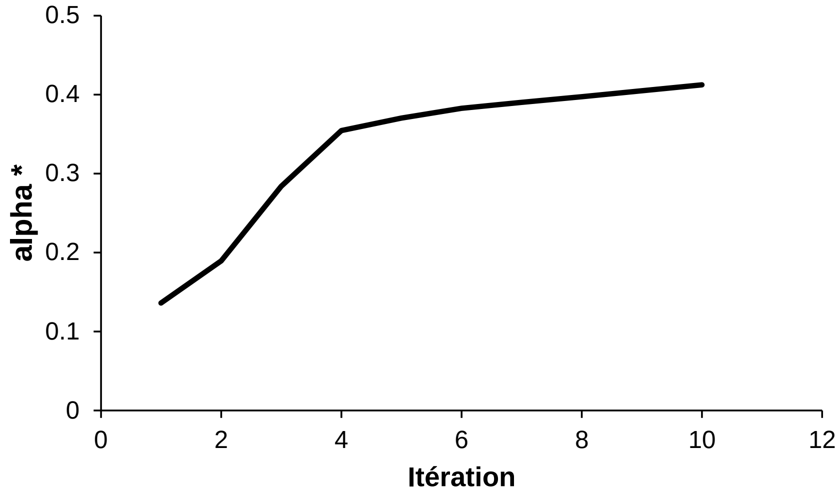


*Figure 5 – Evolution de l'indicateur de cohérence α**

Tableau 4 – Réconciliation avec valeurs aberrantes

Flux/Stock	Description	Support min	Support max	Noyau min	Noyau max
F1	Import vers Séparation	11.0	15.0	12.5	13.9
F2	**Séparation vers Export**	**6.0**	**13.0**	**8.6**	**11.3**
F3	Séparation vers Fabrication	2.0	7.0	2.0	5.0
F4	Fabrication vers Export	2.0	5.0	3.1	4.9
F5	**Import vers Fabrication**	**4.0**	**14.0**	**7.7**	**9.1**
F6	Fabrication vers Manufacture	5.0	11.0	7.2	9.7
F7	Manufacture vers Export	3.0	12.0	7.7	10.1
F8	Import vers Manufacture	10.0	15.0	11.0	12.4
F9	Manufacture vers Utilisation	8.0	15.0	9.4	10.5
F10	**Import vers Utilisation**	**10.0**	**35.0**	**28.5**	**29.6**
F11	Utilisation vers Gestion déchets	9.0	13.0	11.1	12.2
F12	Gestion déchets vers décharge	9.0	13.0	11.1	12.2
S1	Δstock Utilisation	15.0	30.0	25.8	26.9
S2	Δstock Décharge et Environnement	9.0	13.0	11.1	12.2

2.3. NEODYME DANS LES AIMANTS PERMANENTS

2.3.1. Le système Néodyme-aimants

Le système Nd-aimants est décrit dans la Figure 6. Comme indiqué précédemment, les limites du processus « Lithosphère » sont à cheval sur les limites de l'UE-28 (trait tiré) car y sont inclus le bouclier scandinave (et la Péninsule de Kola) ainsi que le Groenland. Or du minerai contenant du Nd extrait à Lovozero dans la Péninsule de Kola est envoyé en Estonie où la société SILMET fait de la séparation en oxydes de Nd. Il y a donc une flèche reliant les processus « Lithosphère » et « Séparation » dans la Figure 5.

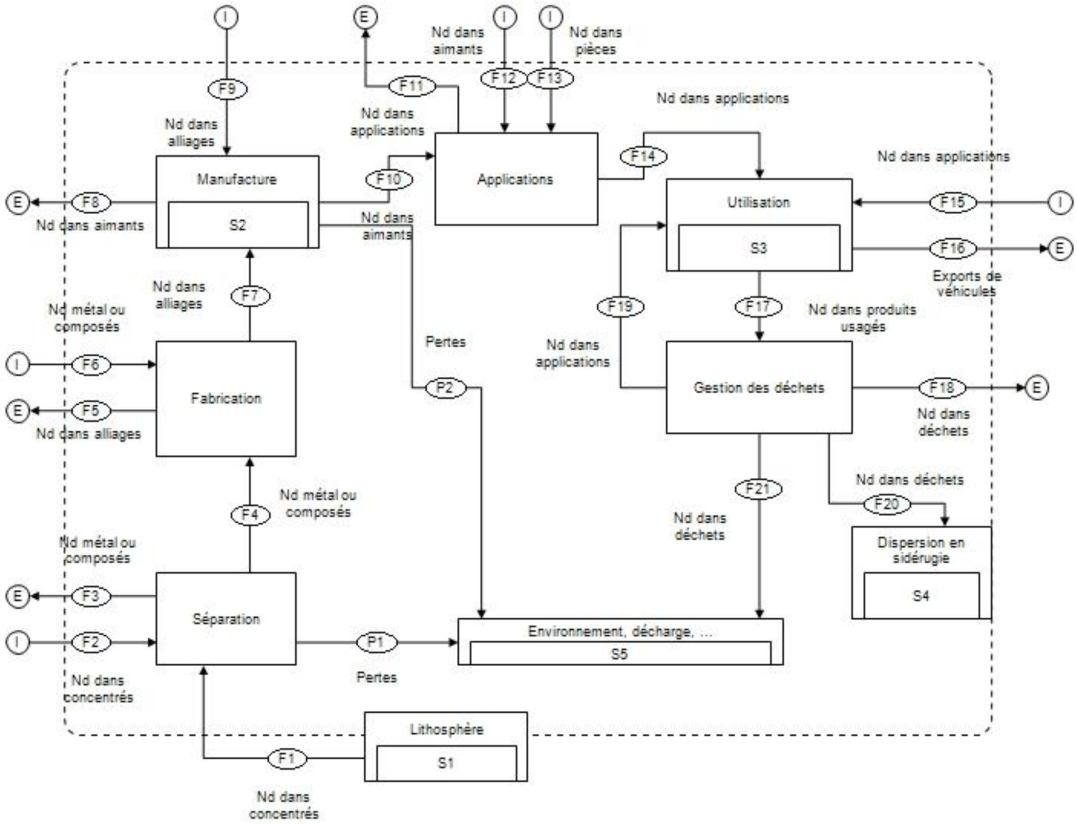


Figure 6 – Schéma du système Nd-aimants

2.3.2. Valeurs issues de la Tâche 5

Les données initiales issues de la Tâche 5 et servant de point de départ à la réconciliation sont présentées dans le Tableau 5.

Tableau 5 - Valeurs issues de la Tâche 5 pour Nd-aimants

Flux/Stock	Valeur min	Valeur max	Valeur préférée	Flux/Stock	Valeur min	Valeur max	Valeur préférée
F1	180.0	250.0	215.0	F15	320.0	390.0	355.0
F2	2.0	5.0	3.5	F16	290.0	440.0	365.0
F3	160.0	230.0	195.0	F17	520.0	630.0	575.0
F4	9.0	25.0	17.0	F18	14.0	18.0	16.0
F5	45.0	65.0	55.0	F19	3.0	5.0	4.0
F6	190.0	240.0	215.0	F20	350.0	430.0	390.0
F7	150.0	200.0	175.0	F21	150.0	180.0	165.0
F8	3.0	7.0	5.0	S1	180.0	250.0	215.0
F9	160.0	180.0	170.0	S2	80.0	95.0	87.0
F10	180.0	230.0	205.0	S3	200.0	350.0	275.0
F11	250.0	310.0	280.0	S4	350.0	430.0	390.0
F12	610.0	630.0	620.0	S5	180.0	230.0	205.0
F13	240.0	400.0	320.0	P1	5.0	8.0	6.5
F14	780.0	960.0	870.0	P2	32.0	38.0	35.0

Notes : F = Flux ; S = stock ; P = pertes

2.3.3. MFA Nd-aimants réconcilié

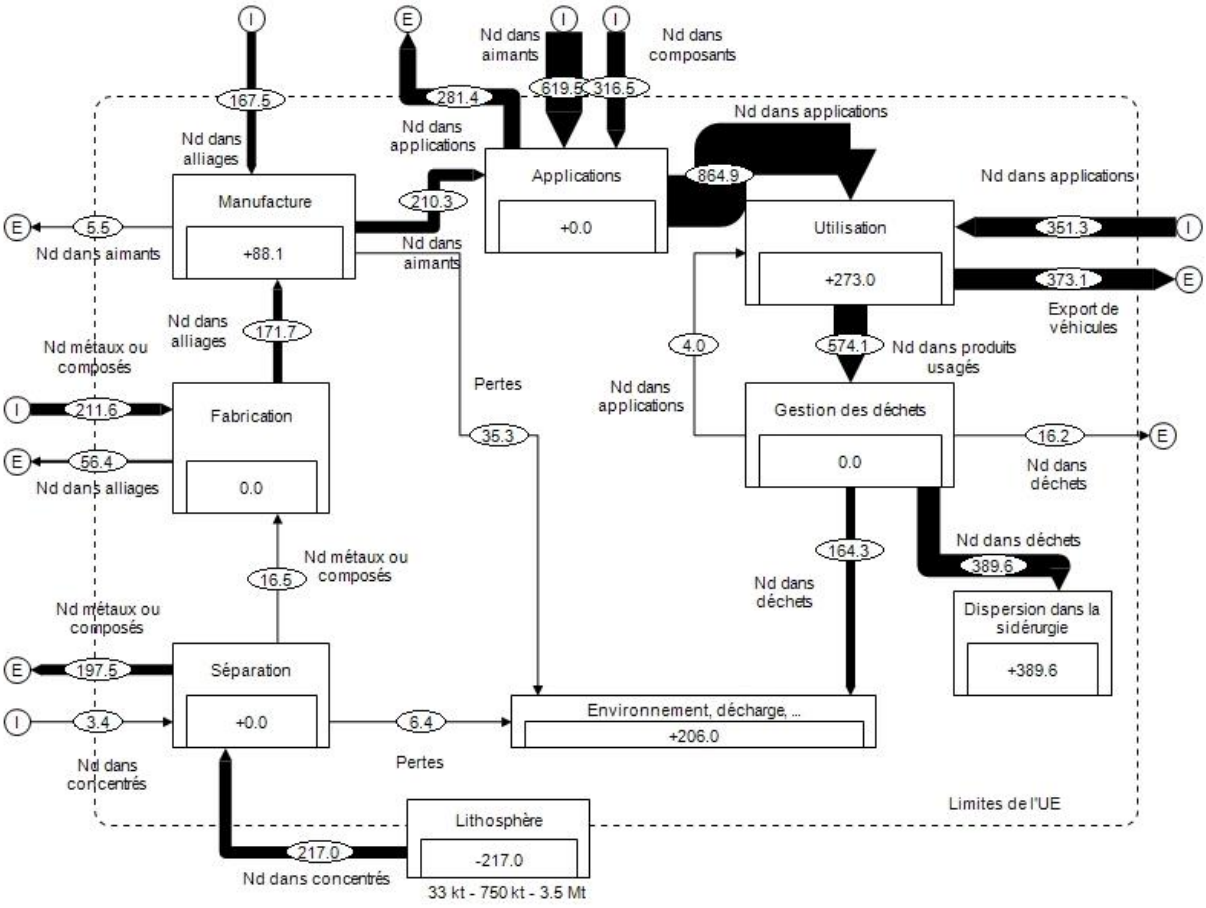
Les valeurs réconciliées sont présentées dans le Tableau 6. Le nombre d'itérations nécessaire pour obtenir les noyaux optimisés est dans ce cas de 31.

Tableau 6 - Valeurs issues de la réconciliation optimale pour Nd-aimants

Flux/Stock	Support		Noyau	Flux/Stock	Support		Noyau
	min	max			min	max	
F1	180.0	250.0	217.0	F15	320.0	390.0	351.3
F2	2.0	5.0	3.4	F16	290.0	440.0	373.1
F3	160.0	230.0	197.5	F17	520.0	630.0	574.1
F4	9.0	25.0	16.5	F18	14.0	18.0	16.2
F5	45.0	65.0	56.4	F19	3.0	5.0	4.0
F6	190.0	240.0	211.6	F20	350.0	430.0	389.6
F7	150.0	200.0	171.8	F21	150.0	180.0	164.3
F8	3.0	7.0	5.5	S1	180.0	250.0	217.0
F9	160.0	180.0	167.5	S2	80.0	95.0	88.1
F10	180.0	230.0	210.3	S3	200.0	350.0	273.0
F11	250.0	310.0	281.4	S4	350.0	430.0	389.6
F12	610.0	630.0	619.5	S5	187.0	226.0	205.9
F13	240.0	400.0	316.4	P1	5.0	8.0	6.4
F14	780.0	960.0	864.9	P2	32.0	38.0	35.3

Notes : F = Flux ; S = stock ; P = pertes

La valeur de α^* correspondant à la première itération est $\alpha^* = 0.82$. Les centres des noyaux du Tableau 6 sont représentés dans le diagramme de Sankey de la Figure 7.



Note : valeurs en tonnes Nd métal

Figure 7 – MFA Nd-aimants après réconciliation

2.4. NEODYME DANS LES BATTERIES NIMH

2.4.1. Le système néodyme-batteries

Le système Nd-batteries est décrit dans la Figure 8.

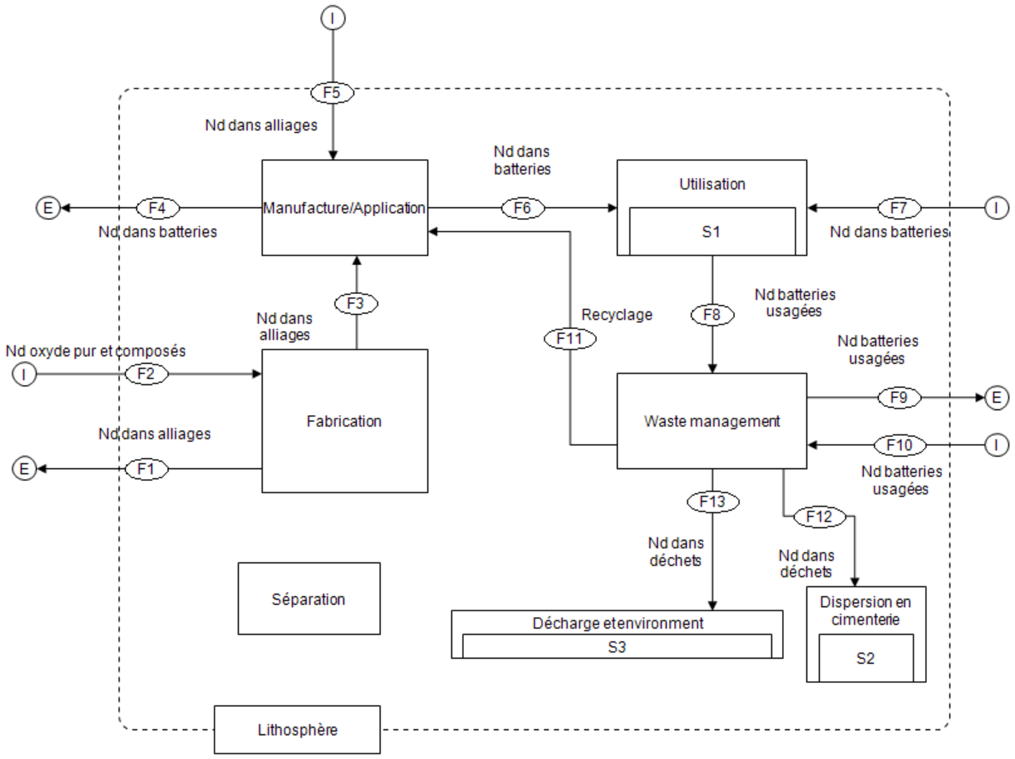


Figure 8 – Schéma du système Nd-batteries

2.4.2. Valeurs issues de la Tâche 5

Les données initiales issues de la Tâche 5 et servant de point de départ à la réconciliation sont présentées dans le Tableau 7.

Tableau 7 - Valeurs issues de la réconciliation optimale pour Nd-batteries

Flux/Stock	Valeur min	Valeur max	Valeur préférée	Flux/Stock	Valeur min	Valeur max	Valeur préférée
F1	4.0	8.0	6.0	F9	0.0	0.2	0.1
F2	20.0	45.0	32.5	F10	5.0	6.0	5.5
F3	18.0	35.0	26.5	F11	14.0	25.0	19.5
F4	20.0	30.0	25.0	F12	52.0	57.0	54.5
F5	20.0	40.0	30.0	S1	30.0	70.0	50.0
F6	28.0	46.0	37.0	S2	15.0	25.0	20.0
F7	75.0	95.0	85.0	S3	35.0	75.0	55.0
F8	64.0	71.0	67.5				

Notes : F = Flux ; S = stock ; P = pertes

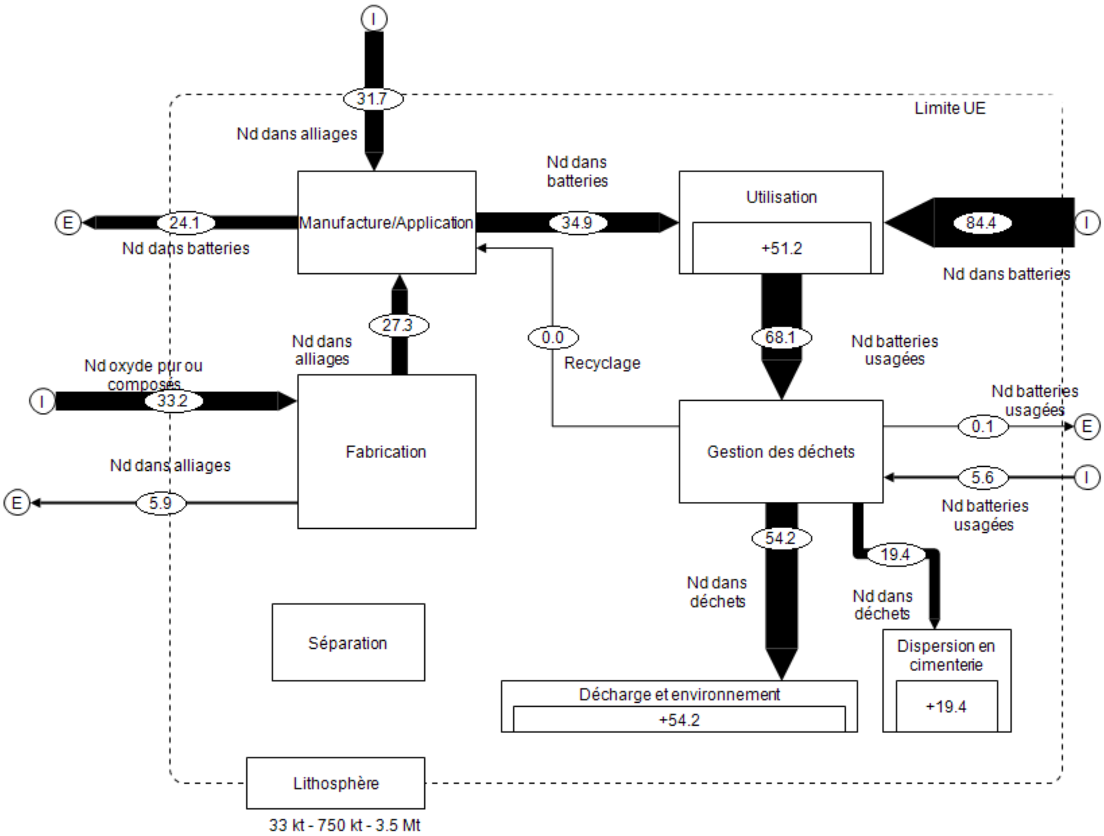
2.4.3. MFA Nd-batteries réconcilié

Les valeurs réconciliées sont présentées dans le Tableau 8. De nouveau, les supports ne changent pas car ils sont compatibles ; par contre on a quelques modifications pour les valeurs les plus « vraisemblables » (noyau) par rapport aux valeurs préférées du Tableau 7. Les valeurs des noyaux sont représentées dans le diagramme de Sankey de la Figure 9

Tableau 8 - Valeurs issues de la réconciliation optimale pour Nd-batteries

Flux/Stock	Support min	Support max	Noyau	Flux/Stock	Support min	Support max	Noyau
F1	4.0	8.0	5.9	F9	0.0	0.2	0.1
F2	22.0	43.0	33.2	F10	5.0	6.0	5.6
F3	18.0	35.0	27.4	F11	15.0	25.0	19.4
F4	20.0	30.0	24.1	F12	52.0	57.0	54.2
F5	20.0	40.0	31.7	S1	32.0	70.0	51.3
F6	28.0	46.0	35.0	S2	15.0	25.0	19.4
F7	75.0	95.0	84.4	S3	52.0	57.0	54.2
F8	64.0	71.0	68.1				

Notes : F = Flux ; S = stock ; P = pertes



Note : valeurs en tonnes Nd métal

Figure 9 - MFA Nd-batteries après réconciliation

3. Conclusions

Un code de réconciliation des données du MFA sous contraintes floues a été développé et appliqué aux valeurs collectées dans le cadre de la Tâche 5. S'agissant d'une méthodologie en cours de développement, cette première application constitue un test préliminaire et d'autres cas seront nécessaires pour améliorer, d'une part, la robustesse de la méthode et, d'autre part, sa bonne intégration dans le processus de collecte des informations.

A ce stade on peut dire qu'un avantage significatif de la méthodologie proposée tient à la bonne adéquation entre le type d'information requise pour son application et la nature des informations généralement disponibles dans la pratique des études de MFA. En effet, demander à un expert de traduire directement sa connaissance par un intervalle contenant la grandeur recherchée (par exemple un flux) ainsi que par une ou des préférence(s) au sein de cet intervalle, est souvent plus naturel que de lui demander une valeur moyenne et un écart-type relatifs à un processus aléatoire hypothétique générant cette connaissance incomplète.

Dans l'avenir il est projeté d'intégrer cette méthodologie au sein d'un panel de méthodes plus vaste, faisant intervenir des outils à la fois probabilistes et possibilistes, de manière à offrir un choix au modélisateur du MFA souhaitant aborder la question des incertitudes.

4. Bibliographie

Brunner, P.H., Rechberger, H. (2004). Practical Handbook of Material Flow Analysis. Lewis Publishers.

DOE (2011). Critical materials strategy. U.S. Department of Energy Report, December 2011.

Dubois, D., Fargier, H., Guyonnet, D., Ababou, M. (2013). Analyse des flux de matière (MFA) : incertitudes épistémiques et réconciliation. Projet ASTER « Analyse Systémique des Terres Rares – flux et stocks ». Rapport de la Tâche 3. Rapport de recherche IRIT 2013 RR-2013-36-FR.

EC (2010). Critical raw materials for the EU. Report of the ad-hoc Working Group on defining critical raw materials. European Commission, June 2010.

Graedel, T.E., Barr, R., Chandler, C., Chase, T., Choi, J., Christoffersen, L., Friedlander, E., Henly, C., Jun, C., Nassar, N.T., Schechner, D., Warren, S., Yang, M.Y., Zhu, C. (2012). Methodology of metal criticality determination. Environ. Sci. Technol., 46, 1063-1070.

Narasimhan, S., Jordache, C. (2000). Data reconciliation and gross error detection; an intelligent use of process data. Gulf Publishing Company, Houston.

NRC (2008). Minerals, Critical Minerals, and the U.S. Economy. U.S. National Research Council. Washington, D.C., National Academies Press, 245pp.

Planchon, M., Guyonnet, D., Rollat, A., Tuduri, J. (2014a). Projet ANR ASTER. Rapport de Tâche 5 (flux et stocks de terres rares dans l'anthroposphère). Résultats pour Tb dans les poudres luminophores. Rapport Final BIO IS, Paris (document à accès réservé).

Planchon, M., Guyonnet, D., Rollat, A., Tuduri, J. (2014b). Projet ANR ASTER. Rapport de Tâche 5 (flux et stocks de terres rares dans l'anthroposphère). Résultats pour Nd dans les aimants permanents. Rapport Final BIO IS, Paris (document à accès réservé).

Planchon, M., Guyonnet, D., Rollat, A., Tuduri, J. (2014c). Projet ANR ASTER. Rapport de Tâche 5 (flux et stocks de terres rares dans l'anthroposphère). Résultats pour Nd dans les batteries NiMH. Rapport Final BIO IS, Paris (document à accès réservé).

Tuduri, J., Charles, N., Pourret. O. (2014). Projet ANR ASTER. Rapport de Tâche 4. Potentialités géologiques de terres rares dans l'UE-28 en incluant le bouclier scandinave et le Groenland. Rapport BRGM/RP (à paraître).

Annexe 1

Article Scientifique : méthode de réconciliation sous contraintes floues

International Journal of General Systems, 2014
http://dx.doi.org/10.1080/03081079.2014.920840

A fuzzy constraint-based approach to data reconciliation in material flow analysis

Didier Dubois[a*], Hélène Fargier[a], Meïssa Ababou[b] and Dominique Guyonnet[b]

[a] *IRIT, CNRS & Université de Toulouse, Toulouse, France;* [b] *BRGM-ENAG, Orléans, France*

(Received 3 December 2013; accepted 30 April 2014)

Data reconciliation consists in modifying noisy or unreliable data in order to make them consistent with a mathematical model (herein a material flow network). The conventional approach relies on least-squares minimization. Here, we use a fuzzy set-based approach, replacing Gaussian likelihood functions by fuzzy intervals, and a leximin criterion. We show that the setting of fuzzy sets provides a generalized approach to the choice of estimated values, that is more flexible and less dependent on oftentimes debatable probabilistic justifications. It potentially encompasses interval-based formulations and the least squares method, by choosing appropriate membership functions and aggregation operations. This paper also lays bare the fact that data reconciliation under the fuzzy set approach is viewed as an information fusion problem, as opposed to the statistical tradition which solves an estimation problem.

Keywords: material flow analysis; data reconciliation; least squares; fuzzy constraints

1. Introduction

Material flow analysis (MFA) consists in calculating the quantities of a certain product transiting within a defined system made up of a network of local entities referred to as processes, considering input and output flows and including the presence of material stocks. This method was developed in the sixties to study the metabolism of urban systems, like (Wolman 1965) for water networks. A material flow system is defined by a number of related processes. Material conservation is the basis of MFA: constraints are typically related to conservation laws such as steady-state material, energy and component balance. In MFA, the unknowns to be determined are the values of the flows and stocks at each process. These flows and stocks must be balanced, through a set of linear equations. The basic principle that provides constraints on the flows is that what goes into a process must come out, up to the variations of stock. This is translated into mass-balance equations relative to a process with n flows in, k flows out and a stock level s of the form:

$$\sum_{i=1}^{n} IN_i = \sum_{j=1}^{k} OUT_j + \Delta s \tag{1}$$

where Δs is the amount of stock variation (positive if $\sum_{j=1}^{k} OUT_j < \sum_{i=1}^{n} IN_i$ and negative otherwise).

Such flow balancing equations in a process network define a linear system of the form $Ay^t = B$, y being the vector of N flows and stock variations. In order to evaluate balanced flows and stocks,

*Corresponding author. Email: dubois@irit.fr

data are collected regarding the material flow transiting the network and missing flow or stock variation values are calculated. But this task may face two opposite kinds of difficulties:

- There may not be sufficient information to determine all the missing flows or stock variations.
- There may be on the contrary too much information available and the system of balance equations is incompatible with the available data. This is because the available information is often not sufficiently reliable.

In this paper, we address the second case. If the data are in conflict with the mass-balance equations, it may be because they are erroneous and should be corrected: this is the problem of *data reconciliation*, a well-known problem in science as early as the end of the eighteenth century when this question was addressed using the method of least squares. The idea was to find solutions to a system of linear equations as close as possible to measured values according to the Euclidean distance. The same method is still used today, but the justification is statistical and usually based on the Central Limit Theorem, the use of Gaussian functions and the principle of maximum likelihood.

Data reconciliation has been defined as "a technique to optimally adjust measured process data so that they are consistent with known constraints" by Kelly (2004). According to Crowe (1996), such adjustment consists in a "constrained minimization problem that is usually one of constrained least squares"; see also (Narasimhan and Jordache 2000). Ayres and Kneese (1969) extended the application of MFA to national economies while Baccini and Brunner (1991) used it to study the metabolism of the anthroposphere, i.e. that portion of the environment that is made or modified by humans for use in human activities and human habitats. MFA has become an important tool in the field of industrial ecology, e.g. (Frosch and Gallopoulos 1989). More recently, researchers have applied MFA to study the global flows and stocks of metals (e.g. Bonnin et al. 2013; Graedel et al. 2004). Some have generalized MFA to several periods of time via a dynamic approach (Bai, Thibault, and McLean 2006). Data reconciliation software is now available such as STAN (Brunner and Rechberger 2004) or BILCO (Durance, Brochot, and Mugabi 2004).

In this paper, we examine the limitations of the classical approach and propose an alternative one that takes into account data uncertainty more explicitly, using intervals or fuzzy intervals[1]. Under this view, imprecise data are considered as (flexible) constraints, to the same extent as balance equations, contrary to the statistical methodology that considers the former as random variables. Under this approach, the problem can then be solved using crisp or fuzzy linear programming. The idea of using fuzzy intervals to address data reconciliation problems can be traced back to a paper by Kikuchi (2000) in connection with traffic modelling, and has not been much used in MFA since then, one exception being life cycle inventory analysis (Tan, Briones, and Culaba 2007). The aim of this paper is

- To better position the least squares approach in a historical perspective, questioning its usual justifications in the MFA context.
- To motivate the need for reconciliation methods different from the least squares approach in this context.
- To present and improve the fuzzy interval approach to data reconciliation.
- To outline a general framework based on fuzzy intervals for the derivation of precise estimates that encompasses the least-squares estimate as a special case.
- To show that the problems addressed by the statistical and fuzzy approaches are fundamentally different despite the fact that the computed estimates in each approach are special cases of a more general setting.

The paper is organized as follows: in Section 2, we recall the formulation of the material flow problem and its least squares solution. We question the appropriateness of the statistical justification of this technique in the data reconciliation problem. Section 3 assumes imprecise data are modelled by means of intervals, and presents the interval reconciliation problem where model constraints and imprecise data are handled as crisp constraints. Section 4 extends the interval-based reconciliation method to fuzzy intervals. Section 5 provides simple examples where the specificity of the fuzzy data reconciliation problem can be highlighted. Section 6 describes the application of the proposed methodology to an example inspired by a real case of copper flow analysis. Finally, in Section 7, we compare the statistical and the fuzzy approaches both on the issue of estimating plausible flow values and on that of computing the resulting uncertainty on the flow values. The appendix recalls definitions pertaining to the modelling of uncertain data by fuzzy sets.

2. Data reconciliation via least squares: a discussion

As recalled above, data reconciliation in the MFA context consists in modifying measured or estimated quantities in order to balance the mass flows in a given network. We denote by y the vector of flows and stocks and subdivide it into two sub-vectors x and u, i.e. k informed quantities x_i and $N - k$ totally unknown quantities u_j, to be determined. We denote by $\hat{x}$ the vector of available measurements $\hat{x}_i$. In this paper, we focus on the case when the system $A(xu)^t = B$ has no solution such that $x = \hat{x}$. This absence of solution is assumed to be due to measurement errors or information defects. The problem to be solved is to modify x, while remaining as close as possible to $\hat{x}$, so that the mass balance equations $A(xu)^t = B$ are satisfied.

2.1. The least squares approach

The traditional approach to data reconciliation (Narasimhan and Jordache 2000) considers that data come from measurements, and measurement errors follow a Gaussian distribution with zero average and a diagonal covariance matrix. The precision of each measurement $\hat{x}_i$, understood as a mean value, is characterized by its standard deviation σ_i. Data reconciliation is then formulated as a problem of quadratic optimization under linear constraints. In the simplest case, assuming no variables u (that is, some piece of information is available for all flows and stocks):

$$\text{Find } x \text{ minimizing} \sum_{i=1}^{k} w_i (x_i - \hat{x}_i)^2$$
$$\text{such that } Ax^t = b$$

The solution is known to be of the form (Narasimhan and Jordache 2000):

$$x^* = \hat{x} - W^{-1}A^t \left(AW^{-1}A^t\right)^{-1} A(\hat{x} - b),$$

where W is a diagonal matrix containing terms $1/w_i$. Weights are often of the form $w_i = (\sigma_i)^{-2}$. It is the method of weighted least-squares used to reconcile data in several MFA tools such as STAN (Brunner and Rechberger 2004) or BILCO (Durance, Brochot, and Mugabi 2004).

Such packages sometimes also reconcile variances as explained in (Narasimhan and Jordache 2000). It assumes that the vector of estimated values $\hat{x}$ has a multivariate normal distribution characterized by a covariance matrix C generalizing W whose diagonal contains the variances σ_i^2. The balance flows being linear, the reconciled values x^* depend on the estimated values via a linear transformation, say $x^* = B\hat{x}$; hence, the x^* also have a normal distribution and the covariance matrix of x^* is of the form $C^* = BCB^t$.

2.2. Limitations of the approach

The method of least squares is often justified based on the principle of maximum likelihood, applied to normal distributions. The shape of the latter is in turn justified by the Central Limit Theorem (CLT). If p_i is the probability density function associated with error $\epsilon_i = x_i - \hat{x}_i$, the maximum likelihood is calculated on the function $L(x) = \prod_{i=1}^{k} p_i(x_i - \hat{x}_i)$. If the p_i's are normal with mean 0 and standard deviation σ_i, then $p_i(x_i - \hat{x}_i)$ is proportional to $e^{-\frac{(x_i-\hat{x}_i)^2}{\sigma_i^2}}$. As a consequence, the maximum of $L(x)$ coincides with the solution to the least squares method. The Gaussian assumption seems to be made because of the popularity of Gauss' law. The universal character of this approach, albeit reasonable in certain situations, is nevertheless dubious:

- It is not consistent with the history of statistics (Stigler 1990). The least squares method, developed by Legendre (1805) and Gauss (end of eighteenth century), was discovered prior to the CLT, and the normal law was found independently. Invented precisely to solve a problem of data reconciliation in astronomy, the least squares method sounded natural since it was in accordance with the Euclidean distance. Moreover, it led to solutions that could be calculated analytically and it could justify the use of the intuitively appealing average in the estimation of quantities based on several independent measures. The normal law was discovered by Gauss as the only error function compatible with the average estimator. However, the CLT is a mathematical result obtained independently by Laplace, who later on made the connection between his mathematical result and the least squares method, based on Gauss finding.
- The CLT presupposes a statistical process with a finite mean value E and standard deviation σ. In this case, the average of n random variables v_i has standard deviation $\sigma/\sqrt{n}$ and the distribution of the variable $\frac{\sum_{i=1}^{n} v_i - nE}{\sqrt{n}}$ is asymptotically Gaussian as n increases. The fundamental hypothesis behind the normal distribution is the existence of a finite σ. In practice, this implies that for N observations a_i of v, the empirical variance $msd = \frac{2\sum_{i<j}(a_i - a_j)^2}{N(N-1)}$ remains bounded as N increases. This assumption is neither always true nor easily verifiable; but it is obviously true if the measured quantity is bounded from below and from above due to physical constraints (but then its distribution is not Gaussian, strictly speaking – even if it is often approximated by a Gaussian function).
- The Gaussian hypothesis is only valid in the case of an unbounded random variable. If v_i is positive or bounded, assuming that the quantity $E_n = \frac{\sum_{i=1}^{n} v_i}{n}$ asymptotically follows a normal distribution with standard deviation $\sigma/\sqrt{n}$ is an approximation that may be useful in practice but does not constitute a general principle.

Based on the remarks above, it is natural to look for alternative methods for reconciling data that do not come from the repetitive use of a single measurement process. Indeed, in the reconciliation problem, we rather face the case of having single assessments of many quantities, rather than many measurements of a single quantity. Given the fact that these assessments are not assumed to come from physical sensors (they may come from documents or experts), and that actual values are not independent since related via balance equations, applying some kind of ergodicity that would justify the classical estimation method makes little sense here. In fact, other probabilistic techniques can be envisaged when the Gaussian assumption does not apply. For instance, Gottschalk, Scholz, and Nowack (2010) use a Bayesian approach to represent uncertain data, in the form of various distributions (e.g. the uniform one in case of total uncertainty within limits); they apply Monte-Carlo methods to solve the reconciliation problem. Alhaj-Dibo, Maquin, and Ragot (2008) propose a mixture of two Gaussians to account for noise and gross errors in a separate way, which enable the latter to be coped with in the reconciliation process.

An alternative, more straightforward approach consists in representing error-tainted data by means of intervals and checking the compatibility between these intervals and the material flow model. This is quite different from the standard statistical approach, where the least-squares solution is taken for granted and variance reconciliation is the result of a kind of probabilistic sensitivity analysis around it.

3. Interval reconciliation

In practice, information on mass flows is seldom precise: the data-gathering process often relies on subjective expert knowledge or on scarce measurements published in various documents that moreover might be obsolete. Or the flow information deals with various products grouped together. Each flow value provided by a source can thus be more safely represented, as a gross approximation, by an interval $\hat{X}_i$ that can be considered as encompassing the actual flow value: of course, the less precise the available information, the wider the interval. Missing values u_i can also be taken into account: we then select as its attached interval the domain of possible values of the corresponding parameter (for example, the unknown grade of an ore extracted from a mine and sent to the treatment plant can, by default, be represented by the interval [0,100]%). In the weighted least squares approach to data reconciliation, weights reflect the assumed variance of a Gaussian phenomenon; if such information on variances σ_i^2 is available, we can set $\hat{X}_i = [\hat{x}_i - 3\sigma_i, \hat{x}_i + 3\sigma_i]$ as a realistic interval containing x_i. This choice captures 99.74% of the normal distribution. Actually, the distribution of x_i is often assumed to be Gaussian for practical reasons, even when the actual parameter is known to be positive or bounded due to physical constraints. Thus, knowledge about each of the N variables y_i of the vector $y = xu$ can be approximately modelled by an interval $\hat{Y}_i$. In this setting, there is clearly a uniform treatment of balance equations and data pertaining to measured or non measured flow values.

The representation of flow data by intervals leads us to consider the reconciliation as a problem of constraint satisfaction; the mass balance equations must be satisfied for flux and stock values that lie within the specified intervals – or, to be more precise, we can restrain these intervals to the sole values that are compatible with the balancing model, given the feasibility ranges of other variables in the form of intervals. Formally, the reconciliation problem can be expressed as follows:

For each $i = 1, \ldots N$, find the smallest and largest values for y_i, such that:

$$\begin{aligned} Ay^t &= B \\ y_i &\in \hat{Y}_i, \quad i = 1, \ldots, N \end{aligned}$$

The calculation of consistent minimum and maximum values of y_i is sufficient: since all the equations are linear, we can show that if there exist two flow vectors y and y', each being a solution to the above system of equations, then any vector v lying between y and y' componentwise is a solution of the system of equations $Ay^t = B$.

The problem can of course be solved using linear programming. Due to the linearity of the constraints, it may also be solved by methods based on interval propagation (Benhamou, Granvilliers, and Goualard 2000). For each variable y_i, equation j of the system $Ay^t = B$ can be expressed as

$$y_i = \frac{\sum_{k \neq i} b_j - a_{jk} y_k}{a_{ji}}, \quad i = 1, \ldots, N.$$

We can then project this constraint on y_i and find the possible values of y_i consistent with it. Due to the m linear constraints, the values of y_i can be restricted to lie in the interval:

$$Y_i = \hat{Y}_i \cap \left(\cap_{j=1\ldots,m} \frac{\sum_{k \neq i} b_j - a_{jk} \hat{Y}_k}{a_{ji}} \right),$$

where $\frac{\sum_{k \neq i} b_j - a_{jk} \hat{Y}_k}{a_{ji}}$ is calculated according to the laws of interval arithmetic (Jaulin et al. 2001); if the new interval of possible values of y_i has become more precise ($Y_i \subset \hat{Y}_i$), it is in turn propagated to the other variables. This procedure, known as "arc consistency", is iterated until intervals are stabilized; when there are no disjunctive constraints, it converges within a finite number of steps to a unique set of intervals (Lhomme 1993) (for additional details, see (Benhamou, Granvilliers, and Goualard 2000; Granvilliers and Benhamou 2006)). This approach has actually been applied to reconciliation problems in the area of measurement in the early 2000s (Ragot and Maquin 2004; Ragot, Maquin, and Alhaj-Dibo 2005). Again, contrary to the statistical approach, the model constraints and the imprecise data are handled on a par, the latter being viewed as unary constraints.

4. Fuzzy interval reconciliation

The interval approach of Section 3 does not yield the same type of answer as the least-squares method because it provides intervals rather than precise values. Such intervals may look similar to reconciled variances provided by current software for MFA and data reconciliation like STAN (Brunner and Rechberger 2004) (but as we shall see later, this comparison is misleading).

A natural way to obtain both reconciled values and intervals is to enrich the representation of the information pertaining to flow estimates using the notion of fuzzy interval: the more-or-less possible values of each flow or stock y_i will be limited by a fuzzy interval $\tilde{Y}_i$. For some of these quantities, these constraints will be satisfied to a certain degree, rather than simply either satisfied or violated. In practice, it means that for each informed quantity, not only an interval should be provided, but also a plausible value (or a shorter interval thereof). Such information can be modelled by means of a triangular or trapezoidal fuzzy interval. See the appendix, for an introductory discussion of fuzzy intervals as representing incomplete information, and the basic definitions used in this section.

The problem of searching for a possible solution then becomes an optimization problem – we seek an optimal value within all the (fuzzy) intervals of possible ones. If no solution provides entire satisfaction for all intervals, some of them will be relaxed if necessary (Dubois, Fargier, and Prade 1996).

4.1. The max-min setting

In this approach, the linear equations describing the material flow for each process are considered as integrity constraints that must necessarily be satisfied, but the information relative to possible values of each flow or stock quantity y_i is now represented in the form of a fuzzy interval $\tilde{Y}_i$, understood as a possibility distribution π_i that expresses a flexible unary constraint. This fuzzy interval may coincide with the domain of the quantity, in the case of total ignorance.

An assignment y for all y_i is feasible, provided it satisfies all the constraints. In other words, the degree of plausibility of an assignment $y = xu$ can be obtained by a conjunctive aggregation of the local satisfaction degrees. The simplest approach is to use the minimum as the standard fuzzy conjunction, following the pioneering paper of Bellman and Zadeh (1970). It has the merit of being insensitive to possible dependencies between the involved fuzzily assessed quantities.

The corresponding optimization problem for determining a most plausible estimate was already formulated some years ago by Kikuchi (2000): Find y^* that maximizes

$$\pi_{\min}(y) = \min_{i=1}^{N} \pi_i(y_i) \quad \text{where } Ay^t = B. \tag{2}$$

Let $\alpha^* = \min_{i=1}^{N} \pi_i(y_i^*)$ be the maximal plausibility value and y^* an optimal solution. The value α^* can be interpreted as the degree of consistency of the flexible constraint problem. This implies that we cannot hope for a plausibility value $\alpha > \alpha^*$, since there will be no simultaneous choice of the y_i in the α-cuts of $\tilde{Y}_i$ that will form a consistent vector in the sense of the network defined by $Ay^t = B$, whereas there exists at least one consistent assignment of flows y^* at level α^*. This approach was in fact already used in an algorithm for tuning the cutting parameters of machine tools, based on expert preference, under crisp constraints pertaining to the production rate (Dubois 1987).

Note that, contrary to what examples in the paper by (Kikuchi 2000) may suggest, there may exist several solutions y^* that achieve a global level of satisfaction α^*. Once α^* is known, we can indeed assign to each flow an interval of optimal values $(\tilde{Y}_i)_{\alpha^*} = \{y_i : \pi_i(y_i) \geq \alpha^*\}$ by solving for each y_i the following interval reconciliation problem: Find the minimum (resp. maximum) values of y_i such that $Ay^t = B$ and:

$$\pi_j(y_j) \geq \alpha^*, \quad j = 1, \ldots, N.$$

Rather than providing the user with one amongst several optimal solutions, it is often more informative to have reconciled flows in the form of fuzzy intervals $\tilde{Y}_i^*$ obtained by projection of $\pi_{\min}$ on the domain of each y_i:

$$\forall v \in S(\tilde{Y}_i), \quad \mu_{\tilde{Y}_i^*}(v) = \max_{y \text{ s.t. } y_i = v \text{ and } Ay^t = B} \min_{j=1}^{N} \pi_j(y_j)$$

The supports of the fuzzy intervals $\tilde{Y}_i^*$ containing the y_i's can be obtained if we use the supports of the $\tilde{Y}_i$ in the procedure of the previous section. This mathematical program contains on the one hand the mass flow model $Ay^t = B$ which, as seen previously, is linear; moreover, we force the y_i's to belong to the supports $[\underline{s}_i, \overline{s}_i]$, $i = 1, \ldots N$ of the fuzzy intervals $\tilde{Y}_i$'s. Finding the fuzzy reconciled flows $\tilde{Y}_i^*$ provides both plausible ranges and uncertainty around them. These fuzzy domains are subnormalized if $\alpha^* < 1$: they all have heights $h_i = \sup_{y_i} \pi_i^*(y_i) = \alpha^*$ and at least one of them contains a single value y_i^*, while the α^*-cuts of others are intervals of optimal values.

The fuzzy reconciliation method fails to deliver a solution if $\alpha^* = 0$. In that case, we may consider that the data are inconsistent with the material flow model, and we must either re-assess the validity of data items (deleting some of them considered unreliable, or relaxing the tolerance distributions) or revise the material flow model. The possibility of this occurrence contrasts with the least squares method which always provides a solution. But this solution may yield values far away from the original estimates in case the data are strongly conflicting with the balance equations. It will correspond to a very low likelihood value in the Gaussian interpretation. While the possible failure of the fuzzy constraint-based reconciliation method may be regarded as a drawback, one may on the contrary see it as a virtue as it is capable of warning the user when an inconsistency occurs without having to prescribe an arbitrary likelihood global threshold on the [0, 1] scale.

4.2. Resolution methods

From a technical standpoint, the fuzzy interval reconciliation problem can be solved using three alternative approaches:

4.2.1. Using a fuzzy interval propagation algorithm

As in the crisp case, fuzzy intervals of possible values $\tilde{Y}_i$ can be improved by projecting the fuzzy domains of other variables over the domain of y_i via the balancing equations:

$$\tilde{Y}_i' = \tilde{Y}_i \cap \left(\cap_{j=1\dots,m} \frac{\sum_{k \neq i} b_j - a_{jk} \tilde{Y}_k}{a_{ji}} \right),$$

where $\frac{\sum_{k \neq i} b_j - a_{jk} \tilde{Y}_k}{a_{ji}}$ is a fuzzy interval $\tilde{A}_j$ that can be easily obtained by means of fuzzy interval arithmetics (Dubois et al. 2000) since equations are linear. Expressions such as $\tilde{Y}_i \cap (\cap_{j=1\dots,m} \tilde{A}_j)$ have possibility distribution $\pi_i' = \min \left(\pi_i, \min_{j=1\dots,m} \pi_{\tilde{A}_j}\right)$.

The propagation algorithm iterates these updates by propagating the new fuzzy intervals on all the neighbouring y_i's, until their domains no longer evolve. This procedure presupposes efficient fuzzy interval representation schemes must be used. Typically we should use piecewise linear fuzzy intervals (Steyaert et al. 1995) including subnormalized ones. Eventually, the optimal (maximally precise) fuzzy intervals $\tilde{Y}_i^*$ (fuzzy domains of the reconciled flows defined in the previous subsection) are obtained with heights not greater than α^*. Indeed, fuzzy arithmetic methods applied to fuzzy intervals of various heights only preserve the least height (Dubois et al. 2000).

4.2.2. Using α-cuts

In order to take advantage of the calculation power of modern linear programming packages, a simple solution is to proceed by dichotomy on the α-cuts of the fuzzy intervals: once each $\tilde{Y}_i$ is cut at a given level α, we obtain a system of equations as in Section 3, replacing $\hat{Y}_i$ by the interval $(\tilde{Y}_i)_\alpha$; this system can therefore be solved by calling an efficient linear programming solver. If the solver finds a solution, the level α is increased; if not, i.e. if it detects an inconsistency in the system of equations, the value α is decreased, etc. until the maximum value α^* is obtained with sufficient precision, along with the corresponding intervals $(\tilde{Y}_i^*)_{\alpha^*}$.

4.2.3. Using fuzzy linear programming

When the fuzzy intervals are triangular or trapezoidal (or even homothetic, as in the case of L-R fuzzy numbers), it is possible to write a (classical) linear program in order to obtain the value of α^*, then obtain the optimal ranges $(\tilde{Y}_i)_{\alpha^*}$'s for the reconciled flows. It is necessary to model the fact that the global degree of plausibility of the optimal reconciled values is the least among the local degrees of possibility, i.e. we should maximize a value less than all the $\pi_i(y_i)$, hence we should write N constraints $\alpha \leq \pi_i(y_i), i = 1, \dots, N$ (a trick as old as (Zimmermann 1978)). When the original fuzzy intervals are triangular with core $\hat{y}_i$ and support $[\underline{s}_i, \overline{s}_i]$, each constraint is written in the form of two linear inequalities, one for each side of the fuzzy intervals, as already proposed in (Kikuchi 2000; Tan, Briones, and Culaba 2007). All these equations being linear, we can then use a linear solver to maximize the value α such that:

$$
\begin{aligned}
Ay^t &= B \\
\underline{s}_i \le y_i \le \overline{s}_i, &\quad i = 1, \ldots, N \\
\alpha(\hat{y}_i - \underline{s}_i) \le y_i - \underline{s}_i, &\quad i = 1, \ldots, N \\
\alpha(\overline{s}_i - \hat{y}_i) \le \overline{s}_i - y_i, &\quad i = 1, \ldots, N
\end{aligned}
$$

The same type of modelling yields the inf and sup limits of the α^*-cuts for the reconciled intervals $\tilde{Y}_i^*$ (maximizing and minimizing y_i, letting $\alpha = \alpha^*$ in the constraints above). By virtue of the linearity of the system of equations and of the membership functions, we can reconstruct the reconciled $\tilde{Y}_i^*$ up to possibility level α^* by linear interpolation between the cores and the optimal supports obtained by deleting the third and fourth constraints in the above program (although the reconciled fuzzy intervals might only be piecewise linear).

Among the three approaches, the latter based on fuzzy linear programming looks like the most convenient one.

4.3. Iterating the optimization process

It is possible (and recommended) to iterate the method and update again some of the fuzzy ranges $\tilde{Y}_i{}^*$. Namely, one may refine the optimal intervals $(\tilde{Y}_i{}^*)_{\alpha^*}$ not reduced to a single value yet, and obtain more precise plausible estimates. The idea, described in (Dubois and Fortemps 1999), is that, while some intervals $(\tilde{Y}_i)_{\alpha^*}$ reduce to singletons $\{y_i^*\}$ that can be considered as fully determined flows, other intervals $(\tilde{Y}_i)_{\alpha^*}$ obtained after the previous optimization step can be further reduced to precise values as well.

Namely, let $V_1 = \{i : (\tilde{Y}_i{}^*)_{\alpha^*} = y_i^*\}$ be the indices of parameters whose values are fixed by considering α^*-cuts. This set is not empty for otherwise, since the fuzzy sets $\tilde{Y}_i$ are of triangular shape, one could still raise the level α^* without creating an inconsistency, which by assumption is not the case as α^* is maximal. So, we define a second optimization problem, where we assign their optimal values y_i^* to flows $y_i \in V_1$, and leave other values free in their original fuzzy ranges. We thus solve the following partially instantiated program: maximize the value β such that

$$
\begin{aligned}
Ay^t &= B \\
\underline{s}_i \le y_i \le \overline{s}_i, &\quad i \notin V_1 \\
y_i = y_i^*, &\quad i \in V_1 \\
\beta(\hat{y}_i - \underline{s}_i) \le y_i - \underline{s}_i, &\quad i \notin V_1 \\
\beta(\overline{s}_i - \hat{y}_i) \le \overline{s}_i - y_i, &\quad i \notin V_1 \\
\beta &\ge \alpha^*
\end{aligned}
$$

Then we get a new optimal value $\beta^* > \alpha^*$ that pertains to flows not in V_1. Indeed, there are several possible values $y_i \in (\tilde{Y}_i{}^*)_{\alpha^*}$, when $i \notin V_1$, and the new optimization problem tends to select the ones that have higher membership grades inside $(\tilde{Y}_i{}^*)_{\alpha^*} \cap \tilde{Y}_i$. We thus get narrower optimal ranges $Y_i^2 \subseteq (\tilde{Y}_i{}^*)_{\alpha^*} \cap (\tilde{Y}_i)_{\beta^*}$, $i \notin V_1$ some of which (forming a subset V_2 of flows) again reduce to singletons. So, at this second step, we have instantiated a set $V_2 \cup V_1$ of variables. We can iterate this procedure until all variables y_i are instantiated, at various levels of optimal possibility α_i^*, $i = 1, \ldots k$, with $\alpha_k^* > \alpha_{k-1}^* > \cdots > \alpha_2^* = \beta^* > \alpha_1^* = \alpha^*$. Eventually, it delivers for each variable y_i a 4-tuple $(\underline{s}_i, \overline{s}_i, y_i^*, \alpha_j^*)$, assuming a precise value y_i^* was found at step j. It can be approximated by a triangular fuzzy interval $\tilde{Y}_i^{**}$ such that $[\underline{s}_i, \overline{s}_i]$ is its support as well as the support of $\tilde{Y}_i{}^*$ (found in the first pass), y_i^* is its core and α_i^* its height, that is, precise reconciled values along with their maximal range of possible values around them[2].

The plausible estimates obtained at the end of this recursive procedure are Pareto-optimal in the sense of the vector-maximization of the vectors $(\pi_1(y_1), \dots \pi_N(y_N))$, and leximin-optimal for the maximization. Namely, there does not exist another tuple of values y such that $\pi_i(y_i) \geq \pi_i(y_i^*)$, $\forall i = 1, \dots, N$, and $(\pi_1(y_1), \dots \pi_N(y_N)) \neq (\pi_1(y_1^*), \dots \pi_N(y_N^*))$, on the one hand, and moreover, $(\pi_1(y_1^*), \dots \pi_N(y_N^*))$ is maximal for the leximin order defined by $(a_1, \dots, a_N) \geq_{lmin} (b_1, \dots, b_N)$ if and only if $(a_{\sigma(1)}, \dots, a_{\sigma(N)})$ is lexicographically greater than $(b_{\tau(1)}, \dots, b_{\tau(N)})$[3] where $(a_{\sigma(1)}, \dots, a_{\sigma(N)})$, and $(b_{\tau(1)}, \dots, b_{\tau(N)})$ are the two vectors reshuffled in the increasing order: $a_{\sigma(1)} \leq \dots \leq a_{\sigma(N)}$, and $b_{\tau(1)} \leq \dots \leq b_{\tau(N)}$ (see Dubois and Fortemps (1999) for details on the leximin order in the setting of max-min optimization).

5. Some examples

We present simple examples in order to compare the statistical and fuzzy approaches.

5.1. One-process case

We consider the example illustrated in Figure 1, which is composed of four flows (y_1, y_2, y_3 and y_4) and one process (P1). Flows y_1 and y_2 enter the process, while y_3 and y_4 exit the process. There are no stocks. In this example, we have symmetric triangular fuzzy intervals $\tilde{Y}_1 = 24 \pm 2$, $\tilde{Y}_2 = 16 \pm 3$, $\tilde{Y}_3 = 15 \pm 4$, $\tilde{Y}_4 = 22 \pm 5$.

With the fuzzy interval approach, the calculation of α^* using linear programming is obtained by solving the following linear problem: Maximize α such that:

$$\begin{aligned}
y_1 + y_2 &= y_3 + y_4 \\
22 &\leq y_1 \leq 26 \\
\alpha \cdot (26 - 24) &\leq 26 - y_1 \\
\alpha \cdot (24 - 22) &\leq y_1 - 22 \\
13 &\leq y_2 \leq 19 \\
\alpha \cdot (19 - 16) &\leq 19 - y_2 \\
\alpha \cdot (16 - 13) &\leq y_2 - 13 \\
11 &\leq y_3 \leq 19 \\
\alpha \cdot (19 - 15) &\leq 19 - y_3 \\
\alpha \cdot (15 - 11) &\leq y_3 - 11 \\
17 &\leq y_4 \leq 27 \\
\alpha \cdot (27 - 22) &\leq 27 - y_4 \\
\alpha \cdot (22 - 17) &\leq y_4 - 17
\end{aligned}$$

The results obtained using the two methods (least-squares and fuzzy interval reconciliation) are provided in Table 1. We note that the alpha-cuts of the fuzzy intervals at level α^* after propagation are singletons (no need for a second pass) and that the maximum distance between the initial and reconciled values is smaller in the case of the fuzzy method than with the least-squares method, which is expected since the aim of the max-min approach is precisely to minimize the largest deviation from initial values. Moreover, in this example, the original supports are left unchanged by the reconciliation procedure.

Figure 1. Example 1: one process with 4 flows.

Table 1. Reconciliated flows for Example 1.

	y_1	y_2	y_3	y_4
Least squares method				
Original data	24 ± 0.67	16 ± 1	15 ± 1.33	22 ± 1.67
Reconciliated values and S.D's	23.8 ± 0.54	15.5 ± 0.91	$15, 9 \pm 1.12$	23.4 ± 1.22
Fuzzy set method				
Original fuzzy intervals	(22, 24, 26)	(13, 16, 19)	(11, 15, 19)	(17, 22, 27)
$\alpha^* : \frac{11}{14}$				
Reconciliated values	$23 + 4/7$	$15 + 5/14$	$15 + 6/7$	$23 + 1/14$
Reconciliated supports	[22, 26]	[13, 19]	[11, 19]	[17, 27]

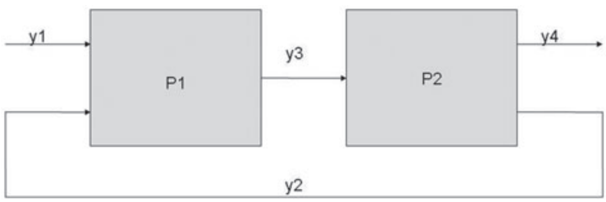


Figure 2. Example 2: two processes with 4 flows.

5.2. Two-process example

We consider the example in Figure 2, composed of four flows (y_1, y_2, y_3 and y_4) and two processes (P1 and P2). Flows y_1 and y_2 both enter process P1; y_3 exits P1 to enter P2, two flows exit P2: y_4 and y_2, while the latter is recycled into P1. In this example, $\tilde{Y}_1 = 20 \pm 3$, $\tilde{Y}_2 = 10 \pm 2$, $\tilde{Y}_3 \in 28 \pm 4$, $\tilde{Y}_4 = 16 \pm 3$.

For the approach using fuzzy intervals, the calculation of α^* by linear programming is obtained by solving a system of equations similar to that of the previous case. We obtain $\alpha^* = 1/3$. We can also obtain this value by calculating the height of $\tilde{Y}_1 \cap \tilde{Y}_4$. Indicated in Table 2 are the cuts at level $1/3$ and the supports of the reconciled fuzzy intervals. We note that reconciled values obtained by least-squares are at the centre of the supports of the reconciled intervals obtained using the fuzzy interval method.

However, it is possible to refine the remaining intervals. If we retain the information $y_1^* = y_4^* = 18$ and run the fuzzy interval propagation procedure again, we verify that the intersection $\tilde{Y}_2 \cap (\tilde{Y}_3 - 18)$ has a height of unity, obtained for $y_2 = 10$. We can also fix $y_3 = 28$ considering $\tilde{Y}_3 \cap (\tilde{Y}_2 + 18)$. We can therefore verify that $\pi_1(18) = \pi_4(18) = 1/3, \pi_2(10) = \pi_3(28) = 1$ and therefore that the least-squares solution coincides in this particular example with the

Table 2. Reconciliated flows For Example 2.

	y_1	y_2	y_3	y_4
Least squares method				
Original data	20 ± 1	10 ± 0.67	28 ± 1.33	16 ± 1.67
Reconciliated values	18 ± 0.64	10 ± 0.61	28 ± 0.79	18 ± 0.64
Fuzzy set method				
Original fuzzy intervals	(17, 20, 23)	(8, 10, 12)	(24, 28, 32)	(13, 16, 19)
$\alpha * : \frac{1}{3}$				
Reconciliated supports	[17, 19]	[8, 12]	[25, 31]	[17, 19]
Reconciliated cores (1st round)	[18, 18]	[8.66, 11.44]	[26.66, 29.33]	[18, 18]
Reconciliated cores: 2d round	[18, 18]	[10, 10]	[28, 28]	[18, 18]

Pareto-optimal solution of the fuzzy data reconciliation problem, due to the symmetry of the network and of the fuzzy intervals. The next example shows that this is rather seldom the case.

5.3. Comparing reconciled values: a simple generic example

Consider a single process with n inputs x_i and a single output $x_0 = \sum_{i=1}^{n} x_i$. Suppose all measured inputs are $\hat{x}_i = a > 0$ while $\hat{x}_0 = ka > 0$. One may argue that, assuming the x_i's have the same variance, x_0 has a variance n times larger. This is what is assumed in the following.

It is easy to obtain least-squares estimates, minimizing $\sum_{i=1}^{n}(x_i - a)^2 + \frac{(x_0 - ka)^2}{n}$ under the balancing constraint. It is easy to find that

$$x_0^{LS} = \frac{a(k+n)}{2} \text{ and } x_i^{LS} = \frac{a}{2} + \frac{ak}{2n}.$$

Note that $\lim_{n\to\infty} x_i^{LS} = a/2$ and in fact $\frac{a}{2} < x_i^{LS} \leq \frac{a(k+1)}{2}$. All reconciled flows linearly increase to infinity if k increases.

In the fuzzy interval approach, we can assume general triangular membership functions: $\tilde{X}_i$ has mode a and support $[a - \alpha, a + \beta]$, where the magnitudes of α, β depend on the available knowledge. Suppose that the relative error of the data is everywhere the same so that $\hat{X}_0$ has mode ka and support $[k(a - \alpha), k(a + \beta)]$. The reconciled value for x_0 is obtained as the value for which the intersection $\hat{X}_0 \cap n\tilde{X}_i$ has maximal positive possibility degree. There are two cases:

$$x_0^* = \begin{cases} \dfrac{nka(\alpha+\beta)}{n\alpha + k\beta} & \text{if } k \leq n \text{ and } k(a+\beta) > n(a-\alpha) \\ \dfrac{nka(\alpha+\beta)}{k\alpha + n\beta} & \text{if } k \geq n \text{ and } k(a-\alpha) < n(a+\beta). \end{cases}$$

It can be checked that the least-squares solution is encompassed by the fuzzy interval approach:

- If $k \leq n$, $x_0^* = x_0^{LS}$ if and only if α, β are chosen such that $n\alpha = k\beta > \frac{a(n-k)}{2}$ (the latter inequality makes the fuzzy reconciliation problem feasible).
- Likewise, if $n \geq k$, the condition is $k\alpha = n\beta > \frac{a(n-k)}{2}$.

These findings are at odds with the least-squares method. Indeed note that in order to get the same estimates in the two approaches, we cannot assume the triangular fuzzy intervals are symmetric, while the translation of normal laws into symmetric fuzzy intervals ($\alpha = \beta$ with spreads $\alpha = 3\sigma$) would enforce symmetry.

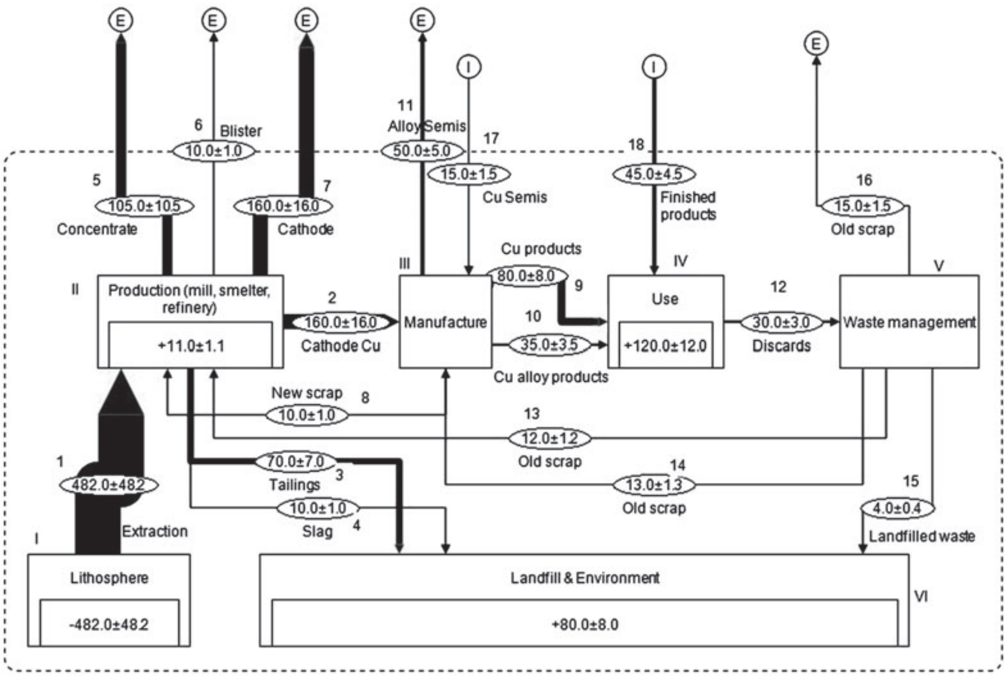


Figure 3. Initial flows and stocks of copper in the Australian economy (adapted from van Beers, van Berkel, and Graedel 2005). Numbers in thousand metric tons.

Finally we can check when it is the case that x_0^* is closer to the estimated value ka than x_0^{LS}. For instance, if $k \leq n$ then $x_0^* > ka$ and $x_0^{LS} > ka$; then it holds that $x_0^{LS} > x_0^* > ka$ provided that the condition $k\beta < n\alpha$ holds.

6. A case study

The least squares and fuzzy reconciliation approaches were compared using data adapted from van Beers, van Berkel, and Graedel (2005) relative to flows and stocks of copper in Australia in the mid-90s (see Figure 3 adapted from Figure 5A in that paper). The processes considered by these authors in their analysis of the major flows of copper over the entire copper life-cycle in Australia are: extraction from the lithosphere (the subsurface), treatment (production) of the copper ore, manufacturing of semi- and finished products (e.g. copper wire, tubing, etc.), use of these products in the economy, waste management and finally landfill and the environment, overall 24 quantities to be reconciliated.

For the purpose of the application, flows and stocks in Figure 5A of van Beers, van Berkel, and Graedel (2005) were shifted arbitrarily from their initial values so that the MFA is no longer balanced. This initial MFA, which served for the reconciliation, is depicted in the Sankey diagram of Figure 3, which was constructed using the STAN software (Brunner and Rechberger 2004). Sankey diagrams (Baccini and Brunner 1991) are a specific type of flow diagram, in which the widths of the arrows are proportional to the flow quantities. Such diagrams are typically used to visualize energy or material transfers between processes. They are particularly useful to help identify potentials for recycling.

For example, Figure 3 suggests that the yearly change in stock in the “Landfill and Environment” process is significant when compared to the copper extracted from the subsurface. Such data and diagrams can help motivate efforts with respect to so-called “landfill mining” operations, e.g. (Jain, Townsend, and Johnson 2013). Figure 3 suggests that in the mid-90s, approximately 500 thousand tons of copper were extracted each year as copper ore from Australian mines. This ore was processed in mills, smelters and refineries to produce intermediate copper products (concentrate, blister and copper cathode). Such processing generated discards in the form of tailings and slags that ended up in the “Landfill & Environment” process. The intermediate copper products were sent to manufacturing processes, located within Australia, to generate finished products that entered the economy to be used. Some finished products were exported outside Australia, the limits of which are symbolized by the dashed line in Figure 3. End-of-life products (discards) entered the waste management system, which generated old scrap that was either exported or else recycled within the domestic production and manufacturing processes. Waste containing copper also ended up in landfills.

Figure 3 also provides the means and standard deviations of the flows and stocks used for the least-squares reconciliation. Since no information pertaining to standard deviations is provided by van Beers, van Berkel, and Graedel (2005), standard deviations (before reconciliation) were assumed to be equal to 10% of the mean. For the possibilistic case, the original possibility distributions were assumed to be triangular. The means provided by Figure 3 were assumed to represent the central preferred values of the distribution; the supports were taken as plus or minus three times the standard deviations (see Section 3).

Comparative results of the least-squares and possibilistic reconciliation methods are presented on Table 3. The results of the possibilistic reconciliation are those of the max-min method with leximin iteration. The first pass delivers the supports of the fuzzy intervals, and the global consistency level (α^*) for the fuzzy constraint reconciliation, which is close to 0.4, as shown on Figure 4. It reflects a moderate conflict between original items of information. On this figure, we can see that the flows 12–16 are set to this consistency level and have precise reconciliated values that can be seen on the 2d column from the right on Table 3. These flows are inputs and outputs of the waste management process on Figure 3, which indicates the location of the most conflicting information. The other variables are still assigned intervals corresponding to the optimal consistency value. The three runs needed to reach precise estimates are patent from Figure 4, each run corresponding to a higher possibility value. The right-hand column of Table 3 provides precise estimates resulting from several leximin iterations. Table 3 also illustrates the point that with the possibilistic method, ranges around the preferred values are not necessarily symmetrical, unlike the least-squares method. Work is currently under way to identify the most appropriate graphical representation of such results in a Sankey diagram.

While reconciled flows are generally close to initial flows, there are some significant differences; as for example in the case of “Discards from Use to Waste management”, which vary by nearly 30% compared to the initial value. The values for this flow before and following reconciliation are depicted in Figure 5. For the purpose of the comparison, the probability density functions were normalized to unity. As can be seen in this figure, in the case of least-squares reconciliation, the distribution is shifted laterally and moves outside the support of the initial interval, whereas in the case of the fuzzy reconciliation method, the reconciled flow and its possibility distribution (always) remain within the boundaries of the initial distribution.

This suggests the fact that the distribution of least squares estimated values may overlap a domain of values considered impossible or very unlikely, in any case inconsistent with the original imprecise data.

Besides, Figure 6 pictures relative differences between reconciliated values and original values. It lays bare the fact that while many least-square estimates remain close to original values,

Table 3. Results of reconciliation using the least-squares and the possibilistic methods. Flow and process numbers refer to Figure 3.

Flow	Rec. Mean	Rec. σ	Rec. Support	Optimal Cut First Pass	Final Leximin
1. Extraction	491.1	18.1	[337.4, 626.6]	[390.8, 573.2]	478.6
2. Cu cath to Mnf.	150.6	7.7	[112.0, 207.9]	[129.7, 185.4]	154.1
3. Tailings	68.1	5.2	[49.0, 91.0]	[56.8, 83.2]	68.8
4. Slag	10.0	1.0	[7.0, 13.0]	[8.1, 11.9]	9.8
5. Export Cu con.	104.1	10.1	[73.5, 136.5]	[85.1, 124.9]	105.7
6. Export of blister	10.0	1.0	[7.0, 13.0]	[8.1, 11.9]	10.1
7. Exp.Cu cath.	158.0	14.7	[112.0, 208.0]	[129.7, 190.3]	161.1
8. New scrap to Prod.	10.0	1.0	[7.0, 13.0]	[8.1, 11.9]	10.4
9. Cu products	80.9	6.4	[56.0, 104.0]	[64.9, 95.1]	80.8
10. Cu alloy products	35.2	3.4	[24.5, 45.5]	[28.4, 41.6]	35.4
11. Export alloy	50.7	4.8	[35.0, 65.0]	[40.5, 59.5]	52.2
12. Discards	38.7	1.8	[30.8, 39]	35.7	35.6
13. Old scrap to Prod.	10.6	1.1	[8.4, 15.6]	9.7	9.7
14. Old scrap to Mnf.	11.3	1.2	[9.1, 16.9]	10.5	10.4
15. Landfilled waste	3.8	0.4	[2.8, 5.2]	3.2	3.2
16. Export of old scrap	12.9	1.4	[10.5, 18.7]	12.2	12.3
17. Import Cu semis	14.9	1.5	[10.5, 19.5]	[12.2, 17.8]	14.3
18. Import finished prod.	44.7	4.3	[31.5, 58.5]	[36.5, 53.5]	43.5
I. Ch. In Litho. Stock	491.1	18.1	[337.4, 626.6]	[390.8, 573.2]	478.6
II. Ch. In Prod. Stock	11.0	1.1	[7.7, 14.3]	[8.9, 13.1]	10.9
IV. Change in Use stock	122.1	7.2	[84.0, 156.0]	[97.3, 142.7]	124.1
VI. Change in Landfill Stock	81.9	5.3	[58.8, 104.0]	[68.1, 95.1]	81.9
Total Imports (I)	59.6	4.5	[42, 78]	[48.6, 71.4]	57.8
Total Exports (E)	335.7	17.3	[238, 441.2]	[275.7, 398.6]	341.4

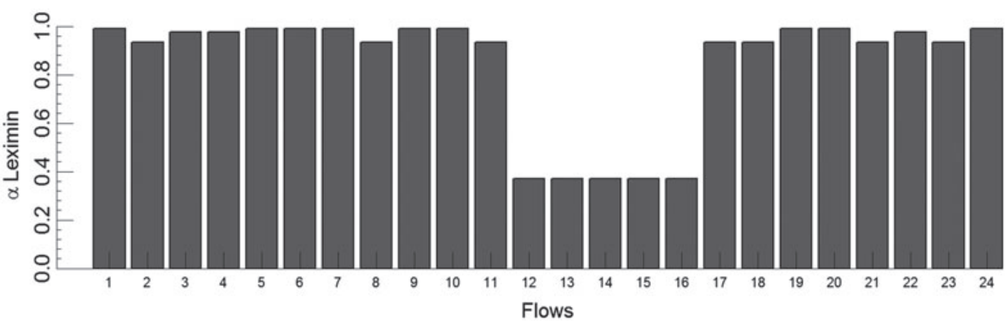


Figure 4. Convergence of the leximin iterative method.

some of them are far away (especially the “Discards from Use to Waste management”, which is the worst result in relative value, while the fuzzy set approach does better). This is because the least square method clearly suggests some initial values of parameter are outliers, while the fuzzy approach tries to build a trade-off between all initial estimates, considered as valuable so long as $\alpha^* > 0$. This view may be considered more natural if initial estimates come from experts and not from measurement devices subject to gross errors.

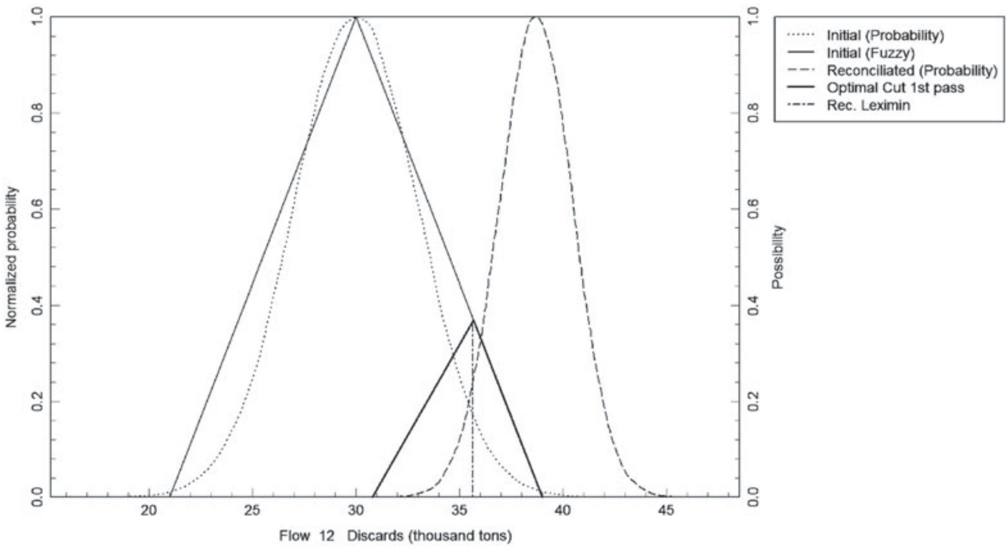


Figure 5. Comparison of discards flow from Use, obtained using the least-squares and possibilistic methods.

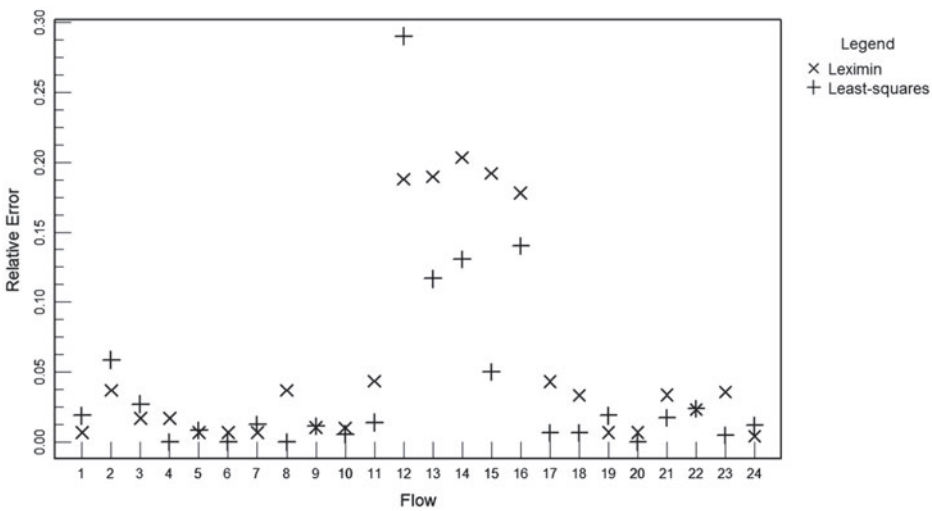


Figure 6. Relative distances of the LS and Leximin solutions to original values.

7. Discussion: least squares or fuzzy set approach

Beyond comparing the results obtained by the two data reconciliation methods on practical examples, it is interesting to see to what extent these methods differ in their principles and the problems they address. The least squares estimation has two possible readings: a distance-based one and a statistical one. The distance-based one turns out to be close to the fuzzy set-based approach in the derivation of the most plausible reconciled values, and as shown below both can be put in the same general formal setting. However, the variance-reconciliation step requires a statistical understanding of the least-squares procedure, and it does not consider measurement data as (unary) constraints in the same sense as balance equations. As explained in this section, beyond the possibility of obtaining different results, the conceptual frameworks underlying the statistical approach to the least-squares method and the fuzzy constraint approach are radically different.

7.1. A unified framework for reconciliation

The max-min formulation (2) of Section 4.1 of the fuzzy constraint approach can be extended, replacing the minimum by a more general fuzzy conjunction. Namely, we may instead consider a likelihood function of the form $L(y) = \star_{i=1}^{N} \pi(y_i)$, where the operation $\star$ is associative, commutative and increasing on [0, 1] – a t-norm (Klement, Mesiar, and Pap 2000). In fact, it is well known that a likelihood function is a special case of a possibility distribution (Dubois, Moral, and Prade 1997). We may then calculate the most plausible reconciled vectors and the associated degree of possibility by solving the following problem: Find the values $y = xu$ that maximize:

$$\pi_\star(y) = \star_{i=1}^{N} \pi_i(y_i) \quad \text{such that } Ay^t = B$$

Restricting to continuous Archimedean t-norms, maximising $\pi_\star(y)$ comes down to minimizing a sum of the form $\sum_{i=1}^{N} g(\pi_i(y_i))$ where g is a t-norm generator (a continuous decreasing mapping from $[0, 1]$ to $[0, +\infty)$ with $g(1) = 0$ (Klement, Mesiar, and Pap 2000)). It comes close to generalized forms of least squares discussed, for instance, in (Alhaj-Dibo, Maquin, and Ragot 2008). Under suitable choice of functions π_i and g, the composition $g(\pi(y))$ is of the form $(\frac{y_i - \hat{y}_i}{\sigma_i})^p$ for some value $p > 1$: the problem comes down to minimizing an l_p norm.

If we select the product for operation $\star$ in the general formulation, the reconciliation problem boils down to maximizing the expression $\pi_\odot(y) = \prod_{i=1}^{N} \pi_i(y_i)$ under constraints $Ay^t = B$. If, in addition, we choose to use Gaussian shapes $\pi_i(y) = e^{-\frac{(y_i - \hat{y}_i)^2}{\sigma_i^2}}$ for the fuzzy intervals, it becomes clear that this formulation brings us precisely back to the standard maximum likelihood expression of the least squares method. Therefore the general fuzzy interval framework captures the least-squares estimation method as a special case, minimizing the Euclidean distance to estimated values.

With $\star = \min$ and triangular fuzzy intervals $\tilde{Y}_i$ centred around measured values $\hat{y}_i$, solving the max-min fuzzy constraint problem, reduces to minimizing the maximal weighted absolute deviation:

$$e_\infty(y) = \max_{i=1,\dots,N} \frac{|y_i - \hat{y}_i|}{\sigma_i},$$

using a Chebyshev l_∞ norm (i.e. $\lim_{p\to\infty} e_p(y) = (\sum_{i=1,\dots,N} \frac{|y_i - \hat{y}_i|^p}{\sigma_i^p})^{1/p}$) instead of the Euclidean l_2 norm. Here σ_i is interpreted as half the support length of the fuzzy interval $\tilde{Y}_i$. The precise estimates obtained by repeating the max-min optimization step (Section 4.3) yields the strict Chebyshev norm solution already known in numerical analysis (Descloux 1963; Rice 1962). In his paper, Rice (1962) even describes a recursive procedure similar to the one we outline.

Similarly, choosing $a \star b = \max(0, a + b - 1)$ under the same hypotheses comes down to minimizing a weighted sum of absolute errors, i.e. use the l_1 norm:

$$e_1(y) = \sum_{i=1,\dots,N} \frac{|y_i - \hat{y}_i|}{\sigma_i}.$$

More generally, recent works on penalty-based aggregation (Calvo and Beliakov 2010) may help us find an even more general setting for devising reconciliation methods in terms of general penalty schemes when deviating from the measured data flows. It is well known that applied to the estimation of a single quantity for which several measurements are available, the three methods, respectively, yield the average (l_2 norm) the median (l_1 norm) and the mid-point between extreme measurement values (l_∞ norm). So the approach using l_1 norm is insensitive to outliers while the one based on the l_∞ norm is very sensitive to extreme values. This state of facts could be viewed as a major drawback for the fuzzy maxmin approach in some context as repeated measurements with gross errors. However, our context is not the one of repeated measurements of a single quantity, but the case of single imprecise expert information items about several quantities related by linear constraints. In this context, it is perfectly reasonable to provide estimates of each quantity that respect as much as possible the opinion of each expert (which is what the fuzzy approach does). However, the fuzzy approach does detect the presence of outlier experts, when the collected information is inconsistent with the model equations. But then, outlier elimination is the result of a reasoned process, not of an automatic data processing method.

7.2. The reconciliation problem : estimation vs. information fusion

Despite the above formal unification of methods computing reconciled values, the statistical approach seems to solve a problem that is radically different from the problem solved by the fuzzy approach, when it comes to modelling the resulting uncertainty on estimated values:

- The statistical approach envisages data reconciliation as an estimation problem, i.e. that of finding the ideal unbiased estimate (namely, the least-squares solution) and computing its distribution as induced by the distributions of the data.
- In the fuzzy approach, the aim is to find the widest fuzzy ranges for reconciled values by projecting the result of merging model constraints and data constraints. The reconciled values are then maximal likelihood estimates *resulting* from these reconciled local fuzzy ranges.

So the statistical approach first computes estimates, while the fuzzy approach primarily computes fuzzy ranges resulting from an information fusion process. The choice of a paradigm (estimation or fusion) actually does not depend on the (probabilistic or not) formal setting. Indeed, one could use a fusion approach in the probabilistic setting and an uncertainty propagation approach in the possibilistic setting.

The probabilistic counterpart of the fuzzy approach, that is a probabilistic information fusion method, may run as follows. Let $\mathcal{D}$ denote the domain encompassed by the flow and stock balance equations, and $P(x)$ be the joint distribution of the measured data:

(1) Condition the joint distribution $P(x)$ of data on the balanced flow domain $\mathcal{D}$;
(2) Compute the projections of $P(x|\mathcal{D})$ on the parameter ranges;
(3) Extract means for all parameters, and the covariance matrix.

Clearly, if $P(x)$ is a multidimensional Gaussian function, it may fail to be the case for the resulting distributions (e.g. if the domain is $\mathcal{D}$ bounded, or for instance the symmetry of distributions may be lost). On the contrary, the usual statistical approach to the reconciliation process preserves the Gaussian nature of the inputs when the model equations are linear (Narasimhan and Jordache 2000). In any case, computing the distribution of the maximum likelihood estimate is different from projecting the conditional probability over the domain of reconciled values on each parameter space. We have seen in the case study that the distribution of best least-squares estimates may fail to fit inside the ranges of the input data, when they are approximated by Gaussian functions.

Conversely, one can envisage possibilistic reconciliation in the spirit of an estimation procedure followed by a sensitivity analysis step:

- Choose a preferred norm (via a t-norm and a shape of π_i) and form the corresponding error criterion.
- Compute a vector of optimal reconciled values y^* as a function, say f, of measured values $\hat{x}$.
- Compute the possibilistic uncertainty on reconciled values by sensitivity analysis using the imprecision of measured values: $\tilde{Y}^* = f(\tilde{X})$ where $\tilde{X}$ is the fuzzy set vector around the initial estimate $\hat{x}$.

Just as in the statistical method, the resulting fuzzy intervals are not necessarily upper bounded by the fuzzy sets originally assigned to input values. The study and implementation of these alternative approaches to reconciliation is left for further research.

8. Conclusion

In the context of the material flow reconciliation problem, we often deal with scarce data of various origins, pertaining to different quantities that we can hardly assume to be generated by a standard random process. It seems more natural to treat the problem as one of information fusion than as a pure statistical estimation based on random measurements. As a consequence, it sounds more reasonable to practically justify the choice of a distance ($l_1, l_2, l_\infty, \ldots$) for minimizing the error rather than to invoke the Central Limit Theorem to justify the least-squares method. A fuzzy-set approach to data reconciliation has been proposed. Its advantages are:

- Its flexible setting for representing various kinds of imprecise information items.
- Its clear conceptual framework as an information fusion problem. The reconciled ranges around the reconciled values are also more easy to interpret than the reconciled variances, as they result from the conjunctive merging of all available information items.
- Its general framework: in a formal sense, it recovers the least-squares method by a proper choice of a shape for membership functions and of a conjunction operation, without betraying the principle of maximum likelihood.
- The possibility of solving the problem in the max-min case using standard linear programming methods and software.

However, this fuzzy constraint-based data reconciliation framework is conceptually at odds with the usual probabilistic reconciliation methods where the flow measurements are viewed as random variables affecting the optimal estimates, and not as additional constraints to be merged with the flow model. Further developments are needed in order to

- Study more examples where the max-min and the least-squares approaches provide disagreeing results, so as to refine the comparison outlined here.
- Compare the information fusion and the estimation approaches inside the probabilistic and possibilistic paradigms, respectively, so as to better understand when one approach is more cogent than the other one.

A software for fuzzy constraint-based approach has been built with a view to apply it to the analysis of material flow of rare earth elements in the anthroposphere of the EU-27.

Acknowledgements

This work is supported by the French National Research Agency (ANR), as part of Project ANR-11-ECOT-002 ASTER “Systemic Analysis of Rare Earths – flows and stocks”. Thanks to Romain Leroux for his help on the copper example.

Notes

1. A first draft of this paper (Dubois, Fargier, and Guyonnet 2013) was presented at the EUSFLAT conference in Milano, September 2013.
2. In fact we also possess all cuts $(\tilde{Y}_i^{\,*})_{\alpha_\ell}, \ell < j$.
3. That is, there exists an index k, such that $a_{\sigma(i)} = b_{\tau(i)}, \forall i < k$, and $a_{\sigma(k)} > b_{\tau(k)}$.

Notes on contributors

Didier Dubois is a CNRS senior research scientist at IRIT, the computer science laboratory of Paul Sabatier University in Toulouse, France. He is the co-author, with Henri Prade, of two books on *Fuzzy Sets and Possibility Theory*, and more than 15 edited volumes on uncertain reasoning and fuzzy sets. Also with Henri Prade, he coordinated the *Handbook of Fuzzy Sets Series* published by Kluwer (7 volumes, 1998-2000). He has contributed more than 200 technical journal papers on uncertainty theories and applications. Since 1999, he has been co-editor-in-chief of the journal *Fuzzy Sets and Systems*. He is a former president of the International Fuzzy Systems Association (1995-1997). In 2002, he received the Pioneer Award of the IEEE Neural Network Society. His topics of interest range from artificial intelligence to operations research and decision sciences, with emphasis on the modelling, representation and processing of imprecise, and uncertain information in reasoning, problem-solving and risk analysis tasks.

Hélène Fargier received her engineering school degree from Compiegne University of Technology in 1988 and a PhD in Computer Science (Artificial Intelligence) in 1994. She worked as a research engineer at Alcatel Alsthom during four years and then moved to the University of Toulouse, where she has now a permanent position as a CNRS senior researcher. Her scientific activity is shared between uncertainty management in Artificial Intelligence on one side and algorithmics for knowledge compilation on the other side.

Dominique Guyonnet (PhD in Hydrogeology) is currently the director of the French National School for Applied Geosciences (ENAG/BRGM-School) located in Orléans, France. His main research interests relate to the representation and propagation of uncertainties in environmental risk assessments, to the performance of mineral barriers used for waste containment and to material flows of resources in the anthroposphere.

Meïssa Ababou holds an engineering degree from the National Polytechnic School of Algiers, Algeria, and more recently, she specialized in the area of sustainable management of mineral resources at ENAG (the French National School for Applied Geosciences), the School of BRGM in Orléans, France. She is currently working with METSO CISA, a company specilized in the development and control of industrial processes in the mining industry.

References

Alhaj-Dibo, M., D. Maquin, and J. Ragot. 2008. “Data Reconciliation: A Robust Approach Using a Contaminated Distribution.” *Control Engineering Practice* 16 (2): 159–170.

Ayres, R. U., and A. V. Kneese. 1969. “Production, Consumption and Externalities.” *American Economic Review* 59 (3): 282–297.

Baccini, P., and H. P. Brunner. 1991. *Metabolism of the Anthroposphere*. New York: Springer.

Bellman, R. E., and L. A. Zadeh. 1970. “Decision Making in a Fuzzy Environment.” *Management Science* 17: B141–B164.

Benhamou, F., L. Granvilliers, and F. Goualard. 2000. “Interval Constraints: Results and Perspectives.” In *New Trends in Constraints, LNAI*, edited by K. R. Apt, A. C. Kakas, E. Monfroy, and F. Rossi, vol. 1865, 1–16. Berlin: Springer-Verlag.

Bai, S., J. Thibault, and D. D. McLean. 2006. “Dynamic Data Reconciliation: Alternative to Kalman Filter.” *Journal of Process Control* 16: 485–498.

Bonnin, M., C. Azzaro-Pantel, L. Pibouleau, S. Domenech, and J. Villeneuve. 2013. "Development of a Dynamic Material Flow Analysis Model for French Copper Cycle." *Chemical Engineering Research and Design* 91 (8): 1390–1402.

Brunner, P. H., and H. Rechberger. 2004. *Practical Handbook of Material Flow Analysis*. Boca Raton, FL: Lewis Publishers.

Calvo, T., and G. Beliakov. 2010. "Aggregation Functions Based on Penalties." *Fuzzy Sets and Systems* 161 (10): 1420–1436.

Crowe, C. 1996. "Data Reconciliation – Progress and Challenges." *Journal of Process Control* 6: 89–98.

Durance, M. V., S. Brochot, and M. Mugabi. 2004. Material Balance Approach for Parameter Determination in Bioleaching Process. In *Computer Applications in Biotechnology*. Proceedings of 9th IFAC International Symposium, edited by M. N. Pons and J. F. Van Impe. Oxford: Pergamon.

Descloux, J. 1963. "Approximations in l_p and Chebyshev Approximations." *Journal of the Society for Industrial and Applied Mathematics* 11: 1017–1026.

Dubois, D. 1987. "An Application of Fuzzy Arithmetics to the Optimization of Industrial Machining Processes." *Mathematical Modelling* 9: 461–475.

Dubois, D. 2006. "Possibility Theory and Statistical Reasoning." *Computational Statistics & Data Analysis* 51: 47–69.

Dubois, D., H. Fargier, and D. Guyonnet. 2013. Data Reconciliation under Fuzzy Constraints in Material Flow Analysis. In *Proceedings of European Society For Fuzzy Logic And Technology Conference (EUSFLAT 2013)*, edited by J. Montero, 25–32. Milan: Atlantis Press.

Dubois, D., H. Fargier, and H. Prade. 1996. "Possibility Theory in Constraint Satisfaction Problems: Handling Priority, Preference and Uncertainty." *Applied Intelligence* 6: 287–309.

Dubois, D., and P. Fortemps. 1999. "Computing Improved Optimal Solutions to Max-min Flexible Constraint Satisfaction Problems." *European Journal of Operation Research* 118: 95–126.

Dubois, D., L. Foulloy, G. Mauris, and H. Prade. 2004. "Probability-possibility Transformations, Triangular Fuzzy Sets, and Probabilistic Inequalities." *Reliable Computing* 10: 273–297.

Dubois, D., E. Kerre, R. Mesiar, and H. Prade. 2000. "Fuzzy Interval Analysis." In *Fundamentals of Fuzzy Sets. The Handbooks of Fuzzy Sets Series*, edited by D. Dubois and H. Prade, 483–581. Boston, MA: Kluwer.

Dubois, D., S. Moral, and H. Prade. 1997. "A Semantics for Possibility Theory Based on Likelihoods." *Journal of Mathematical Analysis Applications* 205: 359–380.

Dubois, D., H. Prade, and P. Smets. 2008. "A Definition of Subjective Possibility." *International Journal of Approximate Reasoning* 48 (2): 352–364.

Frosch, R. A., and N. E. Gallopoulos. 1989. "Strategies for Manufacturing." *Scientific American* 261 (3): 144–152.

Gottschalk, F., R. W. Scholz, and B. Nowack. 2010. "Probabilistic Material Flow Modeling for Assessing the Environmental Exposure to Compounds: Methodology and an Application to Engineered Nano-TiO_2 Particles." *Environmental Modelling & Software* 25 (3): 320–332.

Graedel, T. E., D. van Beers, M. Bertram, K. Fuse, R. B. Gordon, A. Gritsinin, A. Kapur et al. 2004. "The Multilevel Cycle of Anthropogenic Copper." *Environmental Science & Technology* 38: 1253–1261.

Granvilliers, L., and F. Benhamou. 2006. "Algorithm 852: RealPaver: An Interval Solver Using Constraint Satisfaction Techniques." *ACM Trans. on Mathematical Software* 32 (1): 138–156.

Jain, P., T. G. Townsend, and P. Johnson. 2013. "Case Study of Landfill Reclamation at a Florida Landfill Site." *Waste Management* 33 (1): 109–116.

Jaulin, L., M. Kieffer, O. Didrit, and E. Walter. 2001. *Applied Interval Analysis*. London: Springer.

Kelly, J. D. 2004. "Techniques for Solving Industrial Nonlinear Data Reconciliation Problems." *Computers and Chemical Engineering* 28: 2837–2843.

Kikuchi, S. 2000. "A Method to Defuzzify the Fuzzy Number: Transportation Problem Application." *Fuzzy Sets and Systems* 116 (1): 3–9.

Klement, E. P., R. Mesiar, and E. Pap. 2000. *Triangular Norms*. Dordrecht: Kluwer.

Lhomme, O. 1993. "Consistency Techniques for Numeric CSPs." *Proceedings of International Joint Conference on Artificial Intelligence(IJCAI 1993)*, 232–238, Chambéry.

Mauris, G. 2007. "Expression of Measurement Uncertainty in a Very Limited Knowledge Context: A Possibility Theory-Based Approach." *IEEE Transactions on Instrumentation and Measurement* 56 (3): 731–735.

Mauris, G. 2011. "Possibility Distributions: A Unified Representation of Usual Direct-probability-based Parameter Estimation Methods." *International Journal of Approximate Reasoning* 52 (9): 1232–1242.

Narasimhan, S., and C. Jordache. 2000. *Data Reconciliation and Gross Error Detection: An Intelligent Use of Process Data*. Houston: Gulf Publishing Company.

Ragot, J., and D. Maquin. 2004. "Reformulation of Data Reconciliation Problem with Unknown-but-bounded Errors." *Industrial and Engineering Chemistry Research* 43 (6): 1530–1536.

Ragot, J., D. Maquin, and M. Alhaj-Dibo. 2005. "Linear Mass Balance Equilibration: A New Approach for an Old Problem." *ISA Transactions* 44 (1): 23–35.

Rice, J. 1962. "Tschebyschev Approximation in a Compact Metric Space." *Bulletin of the American Mathematical Society* 68: 405–410.

Steyaert, H., F. Van Parys, R. Baekeland, and E. Kerre. 1995. "Implementation of Piecewise Linear Fuzzy Quantities." *International Journal of Intelligent Systems* 10: 1049–1059.

Stigler, S. M. 1990. *The History of Statistics: The Measurement of Uncertainty before 1900*. Cambridge, MA: Belknap Press.

Tan, R. R., L. M. A. Briones, and A. B. Culaba. 2007. "Fuzzy Data Reconciliation in Reacting and Non-reacting Process Data for Life Cycle Inventory Analysis." *Journal of Cleaner Production* 15 (10): 944–949.

van Beers, D., R. van Berkel, and T. E. Graedel. 2005. "The Application of Material Flow Analysis for the Evaluation of the Recovery Potential of Secondary Metals in Australia." In *4th Australian LCA Conference*. Sydney, Australia.

Walley, P. 1991. *Statistical Reasoning with Imprecise Probabilities*. London: Chapman and Hall.

Wolman, A. 1965. "The Metabolism of Cities." *Scientific American* 213 (3): 179–190.

Yager, R. R. 1982. "Level Sets for Membership Evaluation of Fuzzy Subsets." In *Fuzzy Sets and Possibility Theory : Recent Developments*, edited by R. R. Yager, 90–97. Oxford: Pergamon Press.

Zadeh, L. A. 1978. "Fuzzy Sets as a Basis for a Theory of Possibility." *Fuzzy Sets and Systems* 1: 1–28.

Zimmermann, H.-J. 1978. "Fuzzy Programming and Linear Programming with Several Objective Functions." *Fuzzy Sets and Systems* 1: 45–55.

Appendix 1. Modelling data using fuzzy intervals

Intervals have a limited expressive power. One is led to a dilemma between safeness and precision. Namely, short intervals are unreliable, and large intervals are uninformative. However, a very simple and convenient, yet much more expressive, generalization of intervals consists of fuzzy intervals (Dubois 2006) representing possibility distributions on the real line. A possibility distribution is a mapping $\pi : \mathbb{R} \to [0, 1]$ such that $\pi(r^*) = 1$ for some $r^* \in \mathbb{R}$: it is a normal fuzzy set (Zadeh 1978). It represents the current information on a quantity x. The idea is that $\pi(r) = 0$ if and only if $x = r$ is impossible, while $\pi(r) = 1$ if $x = r$ is a totally normal, expected, unsurprising value. One rationale for this framework is that the set $I_\alpha = \{r : \pi(r) \geq \alpha\}$ (α-cut) contains x with level of confidence $1 - \alpha$, that can be interpreted as a lower probability bound (Dubois et al. 2004). In particular, it is sure that $x \in \{r, \pi(r) > 0\} = S(\pi)$, the support of the possibility distribution.

A fuzzy interval is a possibility distribution whose α-cuts I_α are closed intervals. They form a nested family of intervals containing the core $C(\pi) = \{r, \pi(r) = 1\}$ and contained in the support. The simplest representation of a fuzzy interval is a trapezoid defined by its core and its support. Note that this format is very convenient to gather information from experts in the form of nested confidence intervals, or more basically in the form of one safe interval and a plausible value.

Given a possibility distribution π, the degree of possibility of an event A is $\Pi(A) = \sup_{r \in A} \pi(r)$. The degree of certainty of event A is $N(A) = 1 - \Pi(A^c)$, where A^c is the complement of A. A possibility distribution can be viewed as encoding a convex probability family $\mathcal{P}(\pi) = \{P : P(A) \geq N(A), \forall A$ measurable$\}$; see (Dubois 2006) for references. Functions Π and N can be shown to compute exact probability bounds in the sense that:

$$\Pi(A) = \sup_{P \in \mathcal{P}(\pi)} P(A) \text{ and } N(A) = \inf_{P \in \mathcal{P}(\pi)} P(A).$$

In fact, it can be shown (Dubois et al. 2004) that $\mathcal{P}(\pi)$ is characterized by the α-cuts of π:

$$\mathcal{P}(\pi) = \{P : P(\{r : \pi(r) \geq \alpha\}) \geq 1 - \alpha, \forall \alpha > 0\},$$

thus suggesting that a possibility distribution is a kind of two-sided cumulative probability distribution. Probabilistic inequalities yield examples of such possibility distributions. For instance, knowing the mean value and the standard deviation of a random quantity, Chebyshev inequality gives a possibility distribution that encompasses all probability distributions having such characteristics (Dubois et al. 2004). Gauss inequality also provides such possibility distributions encompassing probability distributions with fixed mode and standard deviation as pointed out in (Mauris 2011). It yields a triangular (bounded) fuzzy interval if probability distributions have bounded support. Hence, a possibility distribution may account for incomplete statistical data (Dubois et al. 2004).

In the framework of measurement problems, Mauris (2007) has suggested that in the case of competing error functions (empirical probability distributions $p_i, i = 1, \ldots, k$, such as Gaussian, uniform, double exponential, etc.), one may refrain from choosing one of them and consider a family of probabilities $\mathcal{P}$ instead, to represent our knowledge about x, where $p_i \in \mathcal{P}, \forall i$. In general, such a representation can be extremely complex. For instance, in the setting of imprecise probability theory, $\mathcal{P}$ should be convex, typically the convex hull of $\{p_i, i = 1, \ldots, k\}$ (Walley 1991). Alternatively, when several error functions are possible, one may choose to represent them by a possibility distribution that encompasses them. This is the idea developed by Mauris (2007). This representation is much simpler, even if more imprecise. This remark gives some foundation to the idea of using fuzzy intervals for representing measurement-based imprecise statistical information.

Conversely, if an expert provides a probability distribution that represents subjective belief, it is possible to reconstruct a possibility distribution by reversing the Laplace principle of indifference (Dubois, Prade, and Smets 2008). When the available knowledge is an interval $[a, b]$, and the expert is forced to propose a probability distribution, the most likely proposal is a uniform distribution over $[a, b]$ due to symmetry. If the available knowledge is a possibility distribution π, this symmetry argument leads to replace π by a probability distribution constructed by (i) picking at random a threshold $\alpha \in [0, 1]$ and (ii) a number at random in the α-cut I_α of π (Yager 1982). One may argue that we should bet on the basis of this probability function in the absence of any other information. Conversely, a subjective probability provided by an expert can be represented by the (unique) possibility distribution that would yield this probability distribution using this two-stepped random Monte-Carlo process (Dubois, Prade, and Smets 2008). Note that the symmetric triangular possibility distribution over a bounded interval encompasses the uniform distribution on this interval (it is the most precise choice that retains symmetry) (Mauris 2007).

In summary, fuzzy intervals, and specifically triangular or trapezoidal possibility distributions, may account for uncertain information coming from various origins.

Annexe 2

Code MATLAB de réconciliation sous contraintes floues

```
%****************************************************************************************
% Code
%****************************************************************************************
% Fuzzy constraint-based approach to data reconciliation in material flow analysis
%****************************************************************************************

%****************************************************************************************
% References
%****************************************************************************************
% Dubois, Didier and Fargier, Hélène and Ababou, Meïssa and Guyonnet, Dominique
% A fuzzy constraint-based approach to data reconciliation in material flow analysis
% International Journal of General Systems
%****************************************************************************************
 % Programmed by Romain Leroux, 2014
%****************************************************************************************
% Input and output
%****************************************************************************************
% Input
%   Matrix M : data sheet .xls
% Output
%   alpha_iter : optimal alpha cut optimal
%   reconcilated_supports : reconciled supports
%   reconcilated_cores : reconciled cores at last round
%   alpha_cut_flow : alpha leximin for the flows
%****************************************************************************************

clear all
clc

%************************************************************************

% Input Data and Read Data

%************************************************************************

%********************************************************************************
% Folder for input and output data
%********************************************************************************

repertoire.data.MFA =  input('Name of the folder for MFA ','s');
nom.fichier.MFA = input('Name of the file MFA   ','s');

filename.data.MFA = fullfile(repertoire.data.MFA, nom.fichier.MFA);

nom_fichier_leximin = input('Name of the output file for alpha leximin  ','s');
nom_fichier_reconcilated_cores = input('Name of the output file for the reconciled cores ','s');
nom_fichier_reconcilated_supports = input('Name of the output file for the reconciled supports  ','s');
nom_fichier_optimal_cut_first_pass = input('Name of the output file for the optimal cut first pass  ','s');
nom_fichier_liste_alpha_cut = input('Name of the output file for the list of the alpha cut ','s');

f1 = char(nom_fichier_leximin);
f2 = char(nom_fichier_reconcilated_cores);
f3 = char(nom_fichier_reconcilated_supports);
f4 = char(nom_fichier_optimal_cut_first_pass);
f5 = char(nom_fichier_liste_alpha_cut);

[M Nom_des_flux_stocks] = xlsread(filename.data.MFA,'Parametres du systeme','','basic');

%************************************************************************

% Read Data

% M : data for MFA

% Nom_des_flux_stocks : matrix of the names for flows and stocks
```

```
%**************************************************************************

% MFA parameters

%**************************************************************************

%**************************************************************************
% Number of block
%**************************************************************************

nb_bloc = M(1,1);

%**************************************************************************
% Number of flows
%**************************************************************************

nb_flux = M(2,1);

%**************************************************************************
% Number of stocks
%**************************************************************************

nb_stock = M(3,1);

%**************************************************************************
% Matrix of data
%**************************************************************************

nb_total_ligne_M = size(M,1);
nb_total_colonne_M = size(M,2);

M=M(1:nb_total_ligne_M ,2:nb_total_colonne_M);

%**************************************************************************
% Number of input and output flows and stocks
%**************************************************************************

nb_flux_entrants = M(1:(nb_bloc),3);
nb_flux_sortants = M(1:(nb_bloc),4);

%**************************************************************************
% List of input and output flows and stocks
%**************************************************************************

max_nb_flux_entrants = max(nb_flux_entrants);
max_nb_flux_sortants = max(nb_flux_sortants);

flux_entrants=M(1:nb_bloc,6:(5+max_nb_flux_entrants));
flux_sortants=M(1:nb_bloc,(7+max_nb_flux_entrants):(6+max_nb_flux_entrants+max_nb_flux_sortants));

stocks =
M(1:nb_bloc,(10+max_nb_flux_entrants+max_nb_flux_sortants):(9+max_nb_flux_entrants+max_nb_flux_sortants+nb_bloc));

%**************************************************************************
% Lower and upper bounds for the supports of the stocks
%**************************************************************************

bornes_inf_support_stock = M(1:nb_bloc,9+max_nb_flux_entrants+max_nb_flux_sortants+nb_bloc+1);
bornes_sup_support_stock = M(1:nb_bloc,9+max_nb_flux_entrants+max_nb_flux_sortants+nb_bloc+2);

ind_bornes_inf_support_stock = find(bornes_inf_support_stock ~= 0);
ind_bornes_sup_support_stock = find(bornes_sup_support_stock ~= 0);

bornes_inf_support_stock = bornes_inf_support_stock(ind_bornes_inf_support_stock,:);
bornes_sup_support_stock = bornes_sup_support_stock(ind_bornes_sup_support_stock,:);

%**************************************************************************
% Lower and upper bounds for the cores of the stocks
%**************************************************************************

bornes_inf_noyaux_stocks = M(1:nb_bloc,9+max_nb_flux_entrants+max_nb_flux_sortants+nb_bloc+3);
bornes_sup_noyaux_stocks = M(1:nb_bloc,9+max_nb_flux_entrants+max_nb_flux_sortants+nb_bloc+4);
```

```
ind_bornes_inf_noyaux_stocks = find(bornes_inf_noyaux_stocks ~= 0);
ind_bornes_sup_noyaux_stocks = find(bornes_sup_noyaux_stocks ~= 0);

bornes_inf_noyaux_stocks = bornes_inf_noyaux_stocks(ind_bornes_inf_noyaux_stocks,:);
bornes_sup_noyaux_stocks = bornes_sup_noyaux_stocks(ind_bornes_sup_noyaux_stocks,:);

%***************************************************************************
% Lower and upper bounds for the supports of the flows
%***************************************************************************

bornes_inf_support_flux = M(1:nb_flux,10+max_nb_flux_entrants+max_nb_flux_sortants+nb_bloc+6);
bornes_sup_support_flux = M(1:nb_flux,10+max_nb_flux_entrants+max_nb_flux_sortants+nb_bloc+7);

%***************************************************************************
% Lower and upper bounds for the cores of the flows
%***************************************************************************

bornes_inf_noyaux_flux = M(1:nb_flux,10+max_nb_flux_entrants+max_nb_flux_sortants+nb_bloc+8);
bornes_sup_noyaux_flux = M(1:nb_flux,10+max_nb_flux_entrants+max_nb_flux_sortants+nb_bloc+9);

bornes_inf_support_flux = [ bornes_inf_support_flux ; bornes_inf_support_stock ];
bornes_sup_support_flux = [ bornes_sup_support_flux ; bornes_sup_support_stock ];

bornes_inf_noyaux_flux = [bornes_inf_noyaux_flux ;bornes_inf_noyaux_stocks ];
bornes_sup_noyaux_flux = [bornes_sup_noyaux_flux ;bornes_sup_noyaux_stocks ];

%***************************************************************************
% Test for the order of the lower and bounds of the flows
%***************************************************************************

if ( find(bornes_inf_noyaux_flux < bornes_inf_support_flux  )~=0 |...
   find(bornes_inf_noyaux_flux > bornes_sup_support_flux  )~=0 |...
   find(bornes_sup_noyaux_flux < bornes_inf_support_flux )~=0 |...
   find(bornes_sup_noyaux_flux > bornes_sup_support_flux  )~=0)

disp('Error in lower and upper bounds for the support the flows')

pause('on');
pause;

break

end

%***************************************************************************
% Number max of iterations
%***************************************************************************

nb_iter_alpha = M(1,10+max_nb_flux_entrants+max_nb_flux_sortants+nb_bloc+10);

%***************************************************************************
% Epslion for the lower and upper bounds of the supports
%***************************************************************************

epsilon_bound = M(1,10+max_nb_flux_entrants+max_nb_flux_sortants+nb_bloc+11);

%***************************************************************************
% Matlab options for enable/disable warning messages
%***************************************************************************

options = optimset('Display','off');

%***************************************************************************
% List of the flowd
%***************************************************************************

n=17+max_nb_flux_entrants+max_nb_flux_sortants+nb_bloc;

nb_total_flux_stocks = nb_flux + nb_stock;

liste_flux = [Nom_des_flux_stocks(3:(size(Nom_des_flux_stocks(:,n),1)),n) ;
Nom_des_flux_stocks(2+ind_bornes_inf_noyaux_stocks,4)];
```

```
%**************************************************************************

% Parameters for the Linear Programming

%**************************************************************************

%**************************************************************************
% Matrix Ainq
%**************************************************************************

vec.flux = [-1,1]';

Ainq = kron(eye(nb_total_flux_stocks),vec.flux);

%**************************************************************************
% Matrix binq
%**************************************************************************

B = [bornes_inf_support_flux,bornes_sup_support_flux];

n1 =  size(bornes_inf_support_flux,1);
n2 =  2*size(bornes_inf_support_flux,1);

binq = (reshape((reshape(B,n1,2))',n2,1));

T = toeplitz((-1).^(1:n2));

binq = binq .* T(:,1);

%**************************************************************************
% Matrix Aeq
%**************************************************************************

for ( i=1:nb_bloc)

Aeq(i,nonzeros(flux_entrants(i,1:nb_flux_entrants(i))))=1;
Aeq(i,nonzeros(flux_sortants(i,1:nb_flux_sortants(i))))=-1;

end

[row,col] = find(stocks);
stocks= -1 * stocks(:,col);

Aeq = [Aeq,stocks];

%**************************************************************************
% Matrix beq
% Lower Bound lb et Upper bound ub
%**************************************************************************

beq = zeros(nb_bloc, 1);

lower_bound(1,1:n1) = B(1:n1,1) - epsilon_bound;
upper_bound(1,1:n1) = B(1:n1,2) + epsilon_bound;

%**************************************************************************

% Reconciliation of the supports

%**************************************************************************

f_min = zeros(1, nb_total_flux_stocks);
f_max = zeros(1, nb_total_flux_stocks);

for i = 1:nb_total_flux_stocks

   f_min(i) = 1;
   f_max(i) = -1;

   minimum = linprog(f_min,Ainq,binq,Aeq,beq,lower_bound,upper_bound,[],options);
   maximum = linprog(f_max,Ainq,binq,Aeq,beq,lower_bound,upper_bound,[],options);
```

```
    val_min(i) = minimum(i);
    val_max(i) = maximum(i);

    f_min = zeros(1, nb_total_flux_stocks);
    f_max = zeros(1, nb_total_flux_stocks);

end

reconcilated_supports = [val_min;val_max]';

%***************************************************************************

% Leximin :  Optimisation

%***************************************************************************

k = 2*nb_total_flux_stocks;

n1 =  size(bornes_inf_support_flux,1);
n2 =  2*size(bornes_inf_support_flux,1);

%***************************************************************************
% Triangulars intervals
%***************************************************************************

if isequal( bornes_inf_noyaux_flux, bornes_sup_noyaux_flux )==1

temp1 = [bornes_inf_noyaux_flux-bornes_inf_support_flux,bornes_sup_support_flux-bornes_inf_noyaux_flux];

%***************************************************************************
% Trapezoidals intervals
%***************************************************************************

elseif isequal( bornes_inf_noyaux_flux, bornes_sup_noyaux_flux )==0

temp1 = [bornes_inf_noyaux_flux-bornes_inf_support_flux,bornes_sup_support_flux-bornes_sup_noyaux_flux];

end

constraint_alpha = (reshape((reshape(temp1,n1,2))',n2,1));

Ainq_2 = [[Ainq constraint_alpha] ; [Ainq zeros(k,1)]];
binq_2 = [binq;binq];

Aeq_2 = [Aeq  zeros(nb_bloc,1)];

lower_bound = [lower_bound,0]';
upper_bound = [upper_bound,1]';

%***************************************************************************
% Maximisation of the value alpha
% Reconciliationof the cores
%***************************************************************************

f(nb_total_flux_stocks+1) = -1;

disp ('Iteration')
iter=1;
disp(iter)

alpha = linprog(f,Ainq_2,binq_2,Aeq_2,beq,lower_bound,upper_bound,[],options);

alpha_iter = alpha(nb_total_flux_stocks+1);
alpha_leximin(1,1) = alpha_iter;

%***************************************************************************
% Triangulars intervals
%***************************************************************************

x_left = zeros(nb_total_flux_stocks,1);
x_right = zeros(nb_total_flux_stocks,1);

if isequal( bornes_inf_noyaux_flux, bornes_sup_noyaux_flux )==1
```

```
for i = 1:nb_total_flux_stocks

   x_left(i) =  alpha_iter * (bornes_inf_noyaux_flux(i)-bornes_inf_support_flux(i)) + bornes_inf_support_flux(i);

   x_right(i) = -alpha_iter * (bornes_sup_support_flux(i)-bornes_inf_noyaux_flux(i)) + bornes_sup_support_flux(i);

end

%*************************************************************************
% Trapezoidals intervals
%*************************************************************************

elseif isequal( bornes_inf_noyaux_flux, bornes_sup_noyaux_flux )==0

for i = 1:nb_total_flux_stocks

   x_left(i) =  alpha_iter * (bornes_inf_noyaux_flux(i)-bornes_inf_support_flux(i)) + bornes_inf_support_flux(i);

   x_right(i) = -alpha_iter * (bornes_sup_support_flux(i)-bornes_sup_noyaux_flux(i)) + bornes_sup_support_flux(i);

end

end

reconcilated_cores = [x_left ; x_right]';

%*************************************************************************
% Optimal cut first pass
%*************************************************************************

temp2 = (reshape((reshape(reconcilated_cores,n1,2))',n2,1)).* T(:,1);

   for i = 1:nb_total_flux_stocks

   f_min(i) = 1;
   f_max(i) = -1;

   minimum = linprog(f_min,Ainq,temp2,Aeq,beq,x_left,x_right,[],options);
   maximum = linprog(f_max,Ainq,temp2,Aeq,beq,x_left,x_right,[],options);

   val_min(i) = minimum(i);
   val_max(i) = maximum(i);

   f_min = zeros(1, nb_total_flux_stocks);
   f_max = zeros(1, nb_total_flux_stocks);

   end

reconcilated_cores = [val_min;val_max]';

reconciliated_cores_first_pass = reconcilated_cores;

%*************************************************************************
% cores reconciliation : iteration nb_iter_alpha
%*************************************************************************

% Matrix for the variables that are fixes (ie reach the max Leximin) at
% each iteration

indice_variable_changement = zeros( size(reconcilated_cores,1),nb_iter_alpha-1);

ind_value_fixed = find(abs(reconcilated_cores(:,1) - reconcilated_cores(:,2)) < 0.00001 );

indice_variable_changement(ind_value_fixed,iter) = 1;

% Start the loop

for ( iter=2:nb_iter_alpha)

% Test between two consecutives iterations iter-1 and iter for the convergence of the
% Leximin : if the ansolute difference between the results for the reconcilated cores at iter and iter+1 is
% lower than 0.0001 then the loop stop and the results are saved.
```

```
if (max(abs(reconcilated_cores(:,1)-reconcilated_cores(:,2)))>0.00001)

    disp ('Iteration')
    disp(iter)

ind_value_fixed = find(abs(reconcilated_cores(:,1) - reconcilated_cores(:,2)) < 0.00001 );

indice_variable_changement(ind_value_fixed,iter) = 1;

bornes_inf_noyaux_flux(ind_value_fixed) = reconcilated_cores(ind_value_fixed);
bornes_sup_noyaux_flux(ind_value_fixed) = reconcilated_cores(ind_value_fixed);

temp1 = [bornes_inf_noyaux_flux - bornes_inf_support_flux , bornes_sup_support_flux - bornes_inf_noyaux_flux];

n1 =  size(bornes_inf_support_flux,1);
n2 =  2*size(bornes_inf_support_flux,1);

constraint_alpha = (reshape((reshape(temp1,n1,2))',n2,1));

Ainq_2 = [[Ainq constraint_alpha] ; [Ainq zeros(k,1)]];
binq_2 = [binq;binq];

Aeq_2 = [Aeq  zeros(nb_bloc,1)];

lower_bound(ind_value_fixed) = reconcilated_cores(ind_value_fixed);
upper_bound(ind_value_fixed) = reconcilated_cores(ind_value_fixed);

%**************************************************************************
% +/- epsilon sur les bornes inf et sup to ensure the convergence
%**************************************************************************

lower_bound = lower_bound - epsilon_bound;
upper_bound = upper_bound + epsilon_bound;

%**************************************************************************
% Maximisation of the value alpha
%**************************************************************************

f(nb_total_flux_stocks+1) = -1;

alpha = linprog(f,Ainq_2,binq_2,Aeq_2,beq,lower_bound,upper_bound,[],options);

alpha_iter = alpha(nb_total_flux_stocks+1);

alpha_reconciliation(iter,1)=alpha_iter;

%**************************************************************************
% Triangulars intervals
%**************************************************************************

if isequal( bornes_inf_noyaux_flux, bornes_sup_noyaux_flux )==1

for i = 1:nb_total_flux_stocks

   x_left(i) =  alpha_iter * (bornes_inf_noyaux_flux(i)-bornes_inf_support_flux(i)) + bornes_inf_support_flux(i);

   x_right(i) = -alpha_iter * (bornes_sup_support_flux(i)-bornes_inf_noyaux_flux(i)) + bornes_sup_support_flux(i);

end

%**************************************************************************
% Trapezoidals intervals
%**************************************************************************

elseif isequal( bornes_inf_noyaux_flux, bornes_sup_noyaux_flux )==0

for i = 1:nb_total_flux_stocks

    x_left(i) =  alpha_iter * (bornes_inf_noyaux_flux(i)-bornes_inf_support_flux(i)) + bornes_inf_support_flux(i);

    x_right(i) = -alpha_iter * (bornes_sup_support_flux(i)-bornes_sup_noyaux_flux(i)) + bornes_sup_support_flux(i);
```

```
end

end

%************************************************************************
% reconciliated cores
%************************************************************************

reconcilated_cores = [x_left ; x_right]';

n1 =  size(bornes_inf_support_flux,1);
n2 =  2*size(bornes_inf_support_flux,1);
T = toeplitz((-1).^(1:n2));

temp = (reshape((reshape(reconcilated_cores,n1,2))',n2,1)).* T(:,1);

   for i = 1:nb_total_flux_stocks

   f_min(i) = 1;
   f_max(i) = -1;

   minimum = linprog(f_min,Ainq,temp,Aeq,beq,x_left,x_right,[],options);
   maximum = linprog(f_max,Ainq,temp,Aeq,beq,x_left,x_right,[],options);

   val_min(i) = minimum(i);
   val_max(i) = maximum(i);

   f_min = zeros(1, nb_total_flux_stocks);
   f_max = zeros(1, nb_total_flux_stocks);

   end

reconcilated_cores = [val_min;val_max]';

alpha_leximin(:,iter) = alpha_iter;
end
end

disp(' Data Reconciliation ok ')

%************************************************************************
% end loop iter
%************************************************************************

%************************************************************************
% Liste of the alpha cuts
%************************************************************************

[a b]= max( indice_variable_changement==1, [], 2 );

for i=1:size(b,1)

alpha_cut_flow(i,1) = alpha_leximin(b(i));

end

%************************************************************************

% Save the results

%************************************************************************

%************************************************************************
% Output file names .csv
%************************************************************************

filename1 = fullfile(repertoire.data.MFA, f1);
filename2 = fullfile(repertoire.data.MFA, f2);
filename3 = fullfile(repertoire.data.MFA, f3);
filename4 = fullfile(repertoire.data.MFA, f4);
filename5 = fullfile(repertoire.data.MFA, f5);

%************************************************************************
```

```
% alpha leximin
%**************************************************************************

n = arrayfun(@num2str, (1:size(alpha_leximin,2))', 'UniformOutput', false);
Dat = alpha_leximin';
Data = [n,cellfun(@num2str, num2cell(Dat), 'UniformOutput' , false )];
Final_alpha_leximin = [Data];

delimiter = ',';
cellArray = Final_alpha_leximin;
dat = fopen(filename1,'w');

for z=1:size(cellArray,1)
for s=1:size(cellArray,2)
var = eval(['cellArray{z,s}']);
if size(var,1) == 0
var = '';
end
if isnumeric(var) == 1
var = num2str(var);
end
fprintf(dat,var);
if s ~= size(cellArray,2)
fprintf(dat,[delimiter]);
end
end
fprintf(dat,'\n');
end
fclose(dat);

%**************************************************************************
% cores after n passes
%**************************************************************************

DateTime = liste_flux;
Dat = reconcilated_cores(:,1:2);
Data = [DateTime,cellfun(@num2str, num2cell(Dat), 'UniformOutput' , false )];
Final_reconcilated_cores = [Data];

delimiter = ',';
cellArray = Final_reconcilated_cores;
dat = fopen(filename2,'w');

for z=1:size(cellArray,1)
for s=1:size(cellArray,2)
var = eval(['cellArray{z,s}']);
if size(var,1) == 0
var = '';
end
if isnumeric(var) == 1
var = num2str(var);
end
fprintf(dat,var);
if s ~= size(cellArray,2)
fprintf(dat,[delimiter]);
end
end
fprintf(dat,'\n');
end
fclose(dat);

%**************************************************************************
% ncore first passe
%**************************************************************************

n = liste_flux;
Dat = reconciliated_cores_first_pass(:,1:2);
Data = [n,cellfun(@num2str, num2cell(Dat), 'UniformOutput' , false )];
Final_reconciliated_cores_first_pass = [Data];

delimiter = ',';
cellArray = Final_reconciliated_cores_first_pass;
dat = fopen(filename4,'w');
```

```
for z=1:size(cellArray,1)
for s=1:size(cellArray,2)
var = eval(['cellArray{z,s}']);
if size(var,1) == 0
var = '';
end
if isnumeric(var) == 1
var = num2str(var);
end
fprintf(dat,var);
if s ~= size(cellArray,2)
fprintf(dat,[delimiter]);
end
end
fprintf(dat,'\n');
end
fclose(dat);

%****************************************************************************
% supports
%****************************************************************************

n = liste_flux;
Dat = reconcilated_supports(:,1:2);
Data = [n,cellfun(@num2str, num2cell(Dat), 'UniformOutput' , false )];
Final_reconcilated_supports = [Data];

delimiter = ',';
cellArray = Final_reconcilated_supports;
dat = fopen(filename3,'w');

for z=1:size(cellArray,1)
for s=1:size(cellArray,2)
var = eval(['cellArray{z,s}']);
if size(var,1) == 0
var = '';
end
if isnumeric(var) == 1
var = num2str(var);
end
fprintf(dat,var);
if s ~= size(cellArray,2)
fprintf(dat,[delimiter]);
end
end
fprintf(dat,'\n');
end
fclose(dat);

%****************************************************************************
% alpha cut list
%****************************************************************************

n = arrayfun(@num2str, (1:size(Aeq,2))', 'UniformOutput', false);
Dat = alpha_cut_flow;
Data = [n,cellfun(@num2str, num2cell(Dat), 'UniformOutput' , false )];
Final_alpha_cut_flow = [Data];

delimiter = ',';
cellArray = Final_alpha_cut_flow;
dat = fopen(filename5,'w');

for z=1:size(cellArray,1)
for s=1:size(cellArray,2)
var = eval(['cellArray{z,s}']);
if size(var,1) == 0
var = '';
end
if isnumeric(var) == 1
var = num2str(var);
end
fprintf(dat,var);
```

```
if s ~= size(cellArray,2)
fprintf(dat,[delimiter]);
end
end
fprintf(dat,'\n');
end
fclose(dat);

disp(' Write Results ok ' )

pause('on');
pause;

%****************************************************************************************************
% End Code
%****************************************************************************************************
```

Annexe 3

Fichiers Excel d’entrée utilisés pour le code MATLAB

Fichier pour Tb

	A	B	C	D	E	F	G	H	I	J	K	L	M	N	O	P
1	**Parametres du système**		**Blocs**	**Identité**	**Nombre de**	**Nombre de**	**Flux Entrants Blocs**				**Flux Sortants Blocs**					
2					**Flux Entrants**	**Flux Sortants**										
3	Nombre de blocs	8	**1**	**Séparation**	1	2		1	0	0		2	3	0	0	
4	Nombre de flux	12	**2**	**Fabrication**	2	2		3	5	0		4	6	0	0	
5	Nombre de stock	4	**3**	**Manufacture**	2	2		6	8	0		7	9	0	0	
6			**4**	**Utilisation**	2	1		9	10	0		11	0	0	0	
7			**5**	**Gestion des déchets**	1	1		11	0	0		12	0	0	0	
8			**6**	**Décharge et environnement**	1	0		12	0	0		0	0	0	0	
9			**7**	**Total imports**	0	4		0	0	0		1	5	8	10	
10			**8**	**Total exports**	3	0		2	4	7		0	0	0	0	
11																

	Q	R	S	T	U	V	W	X	Y	Z	AA	AB	AC	AD
1	**Stocks**	**Identité**									**Supports des stocks**		**Noyaux des stocks**	
2														
3	**1**	**Séparation**	0	0	0	0	0	0	0	0	0.00	0.00	0.00	0.00
4	**2**	**Fabrication**	0	0	0	0	0	0	0	0	0.00	0.00	0.00	0.00
5	3	**Manufacture**	0	0	0	0	0	0	0	0	0.00	0.00	0.00	0.00
6	**4**	**Utilisation**	0	0	0	1	0	0	0	0	15.00	30.00	22.50	22.50
7	5	**Gestion des déchets**	0	0	0	0	0	0	0	0	0.00	0.00	0.00	0.00
8	**6**	**Landfill & Env**	0	0	0	0	0	1	0	0	5.00	15.00	10.00	10.00
9	**7**	**Imports**	0	0	0	0	0	0	-1	0	30.00	70.00	50.00	50.00
10	**8**	**Exports**	0	0	0	0	0	0	0	1	11.00	25.00	18.00	18.00
11														

	AE	AF	AG	AH	AI	AJ	AK	AL
1	**Flux**	**Identité**	**Supports des flux**		**Noyaux des flux**		**Nombre**	**Espsilon bornes**
2			**Bornes inf**	**Bornes sup.**	**Bornes inf**	**Bornes sup.**	**d'itérations**	**inf. Et sup. Flux**
3	**1**	**I vers S**	11.00	15.00	13.00	13.00	10	0.50
4	**2**	**S vers E**	6.00	9.00	7.50	7.50		
5	**3**	**S vers F**	2.00	7.00	4.50	4.50		
6	**4**	**F vers E**	2.00	5.00	3.50	3.50		
7	**5**	**I vers F**	4.00	8.00	6.00	6.00		
8	**6**	**F vers M**	5.00	11.00	8.00	8.00		
9	**7**	**M vers E**	3.00	12.00	7.50	7.50		
10	**8**	**I vers M**	10.00	15.00	12.50	12.50		
11	**9**	**M vers U**	8.00	15.00	11.50	11.50		
12	**10**	**I vers U**	10.00	30.00	20.00	20.00		
13	**11**	**U vers Déchets**	9.00	13.00	11.00	11.00		
14	**12**	**Déchets vers Décharge**	9.00	13.00	11.00	11.00		
15								

Fichier pour Nd-aimants

	A	B	C	D	E	F	G	H	I	J	K	L	M	N	O	P	Q	R	S	T	U
1	**Parametres du système**		**Blocs**	**Identité**	**Nombre de**	**Nombre de**	**Flux Entrants Blocs**							**Flux Sortants Blocs**							
2					**Flux Entrants**	**Flux Sortants**															
3	Nombre de blocs	11	**1**	**Lithosphère**	0	1		0	0	0	0	0	0		1	0	0	0	0	0	
4	Nombre de flux	23	**2**	**Séparation**	2	3		1	2	0	0	0	0		3	4	5	0	0	0	
5	Nombre de stock	7	**3**	**Fabrication**	2	2		5	7	0	0	0	0		6	8	0	0	0	0	
6			**4**	**Manufacture**	2	3		8	10	0	0	0	0		9	11	12	0	0	0	
7			**5**	**Applications**	3	2		11	14	15	0	0	0		13	16	0	0	0	0	
8			**6**	**Utilisation**	3	2		16	17	21	0	0	0		18	19	0	0	0	0	
9			**7**	**Gestion des déchets**	1	4		19	0	0	0	0	0		20	21	22	23	0	0	
10			**8**	**Aciéries**	1	0		22	0	0	0	0	0		0	0	0	0	0	0	
11			**9**	**Décharge et environnement**	3	0		4	12	23	0	0	0		0	0	0	0	0	0	
12			**10**	**Total imports**	0	6		0	0	0	0	0	0		2	7	10	14	15	17	
13			**11**	**Total exports**	6	0		3	6	9	13	18	20		0	0	0	0	0	0	

	V	W	X	Y	Z	AA	AB	AC	AD	AE	AF	AG	AH	AI	AJ	AK	AL
1	**Stocks**	**Identité**												**Supports des stocks**		**Noyaux des stocks**	
2																	
3	**1**	**Lithosphère**	-1	0	0	0	0	0	0	0	0	0	0	180.00	250.00	215.00	215.00
4	**2**	**Séparation**	0	0	0	0	0	0	0	0	0	0	0	0.00	0.00	0.00	0.00
5	**3**	**Fabrication**	0	0	0	0	0	0	0	0	0	0	0	0.00	0.00	0.00	0.00
6	**4**	**Manufacture**	0	0	0	1	0	0	0	0	0	0	0	80.00	95.00	87.00	87.00
7	**5**	**Applications**	0	0	0	0	0	0	0	0	0	0	0	0.00	0.00	0.00	0.00
8	**6**	**Utilisation**	0	0	0	0	0	1	0	0	0	0	0	200.00	350.00	275.00	275.00
9	**7**	**Gestion des déchets**	0	0	0	0	0	0	0	0	0	0	0	0.00	0.00	0.00	0.00
10	**8**	**Aciéries**	0	0	0	0	0	0	0	1	0	0	0	350.00	430.00	390.00	390.00
11	**9**	**Décharge et environnem**	0	0	0	0	0	0	0	0	1	0	0	180.00	230.00	205.00	205.00
12	**10**	**Total imports**	0	0	0	0	0	0	0	0	0	-1	0	1500.00	1800.00	1650.00	1650.00
13	**11**	**Total exports**	0	0	0	0	0	0	0	0	0	0	1	800.00	1100.00	950.00	950.00
14																	

	AM	AN	AO	AP	AQ	AR	AS	AT
1	**Flux**	**Identité**	**Supports des flux**		**Noyaux des flux**		**Nombre**	**Espsilon bornes**
2			**Bornes inf**	**Bornes sup.**	**Bornes inf**	**Bornes sup.**	**d'itérations**	**inf. Et sup. Flux**
3	**1**	**L vers S**	180.00	250.00	215.00	215.00	10	0.50
4	**2**	**I vers S**	2.00	5.00	3.50	3.50		
5	**3**	**S vers E**	160.00	230.00	195.00	195.00		
6	**4**	**S vers Décharge**	5.00	8.00	6.50	6.50		
7	**5**	**S vers F**	9.00	25.00	17.00	17.00		
8	**6**	**F vers E**	45.00	65.00	55.00	55.00		
9	**7**	**I vers F**	190.00	240.00	215.00	215.00		
10	**8**	**F vers M**	150.00	200.00	175.00	175.00		
11	**9**	**M vers E**	3.00	7.00	5.00	5.00		
12	**10**	**I vers M**	160.00	180.00	170.00	170.00		
13	**11**	**M vers A**	180.00	230.00	205.00	205.00		
14	**12**	**M vers Décharge**	32.00	38.00	35.00	35.00		
15	**13**	**A vers E**	250.00	310.00	280.00	280.00		
16	**14**	**I vers A**	610.00	630.00	620.00	620.00		
17	**15**	**I vers A**	240.00	400.00	320.00	320.00		
18	**16**	**A verse U**	780.00	960.00	870.00	870.00		
19	**17**	**I vers U**	320.00	390.00	355.00	355.00		
20	**18**	**U vers E**	290.00	440.00	365.00	365.00		
21	**19**	**U vers Déchets**	520.00	630.00	575.00	575.00		
22	**20**	**Déchets vers E**	14.00	18.00	16.00	16.00		
23	**21**	**Déchets vers U**	3.00	5.00	4.00	4.00		
24	**22**	**Déchets vers Aciéries**	350.00	430.00	390.00	390.00		
25	**23**	**Déchets vers Décharge**	150.00	180.00	165.00	165.00		
26								

Fichier pour Nd-batteries

	A	B	C	D	E	F	G	H	I	J	K	L	M	N	O	P
1	**Parametres du système**		**Blocs**	**Identité**	**Nombre de**	**Nombre de**	**Flux Entrants Blocs**				**Flux Sortants Blocs**					
2					**Flux Entrants**	**Flux Sortants**										
3	Nombre de blocs	8	**1**	**Fabrication**	1	2		2	0	0		1	3	0	0	
4	Nombre de flux	12	**2**	**Manufacture**	2	2		3	5	0		4	6	0	0	
5	Nombre de stock	5	**3**	**Utilisation**	2	1		6	7	0		8	0	0	0	
6			**4**	**Gestion des déchets**	2	3		8	10	0		9	11	12	0	
7			**5**	**Cimenteries**	1	0		11	0	0		0	0	0	0	
8			**6**	**Décharge et environnement**	1	0		12	0	0		0	0	0	0	
9			**7**	**Total imports**	0	4		0	0	0		2	5	7	10	
10			**8**	**Total exports**	3	0		1	4	9		0	0	0	0	

	Q	R	S	T	U	V	W	X	Y	Z	AA	AB	AC	AD
1	**Stocks**	**Identité**									**Supports des stocks**		**Noyaux des stocks**	
2														
3	**1**	**Fabrication**	0	0	0	0	0	0	0	0	0.00	0.00	0.00	0.00
4	**2**	**Manufacture**	0	0	0	0	0	0	0	0	0.00	0.00	0.00	0.00
5	**3**	**Utilisation**	0	0	1	0	0	0	0	0	30.00	70.00	50.00	50.00
6	**4**	**Gestion des déchets**	0	0	0	0	0	0	0	0	0.00	0.00	0.00	0.00
7	**5**	**Cimenteries**	0	0	0	0	1	0	0	0	15.00	25.00	20.00	20.00
8	**6**	**Décharge et environnem**	0	0	0	0	0	1	0	0	35.00	75.00	55.00	55.00
9	**7**	**Total imports**	0	0	0	0	0	0	-1	0	130.00	180.00	155.00	155.00
10	**8**	**Total exports**	0	0	0	0	0	0	0	1	20.00	40.00	30.00	30.00

	AE	AF	AG	AH	AI	AJ	AK	AL
1	**Flux**	**Identité**	**Supports des flux**		**Noyaux des flux**		**Nombre**	**Espsilon bornes**
2			**Bornes inf**	**Bornes sup.**	**Bornes inf**	**Bornes sup.**	**d'itérations**	**inf. Et sup. Flux**
3	**1**	**F vers E**	4.00	8.00	6.00	6.00	10	0.50
4	**2**	**I vers F**	20.00	45.00	32.50	32.50		
5	**3**	**F vers M**	18.00	35.00	26.50	26.50		
6	**4**	**M vers E**	20.00	30.00	25.00	25.00		
7	**5**	**I vers M**	20.00	40.00	30.00	30.00		
8	**6**	**M vers U**	28.00	46.00	37.00	37.00		
9	**7**	**I vers U**	75.00	95.00	85.00	85.00		
10	**8**	**U vers Déchets**	64.00	71.00	67.50	67.50		
11	**9**	**Déchets vers E**	0.00	0.20	0.10	0.10		
12	**10**	**I vers Déchets**	5.00	6.00	5.50	5.50		
13	**11**	**Déchets vers Cimenteries**	14.00	25.00	19.50	19.50		
14	**12**	**Déchets vers Décharge**	52.00	57.00	54.50	54.50		
15								

Annexe 4

Notice d'utilisation du code MATLAB

Le code de réconciliation par contraintes floues a été codé dans un environnement MATLAB. Pour être utilisé, une version récente de MATLAB doit être installée sur le poste.

1) Placer l'exécutable : Code_Reconciliation_MFA.m (Annexe 2) dans un dossier, ainsi que le fichier Excel (Nom.xls)

2) Dans une fenêtre Explorateur Windows, double-cliquer sur l'exécutable (MATLAB se lance)

3) a l'invite « >> » taper :

>> Code_Reconciliation_MFA

suivi de return

4) Le programme demande le nom du répertoire où se trouve le fichier Nom.xls. Taper Nom_répertoire

5) Le programme demande le nom du fichier Excel. Taper Nom.xls

6) Le programme demande les noms des fichiers sortie :

- le alpha leximin en fonction du nombre d'itérations

- les noyaux réconciliés ;

- les supports réconciliés ;

- les noyaux correspondant à la première itération ;

- les alphas de chaque grandeur à l'issue de la réconciliation.

7) Après le calcul, consulter les fichiers sortie dans le répertoire Nom_Répertoire

3, avenue Claude-Guillemin
BP 36009 – 45060 Orléans Cedex 2 – France – Tél. : 02 38 64 34 34
www.brgm.fr