MICA
Minerals Intelligence Capacity Analysis

FACTSHEET

Reconciliation in Material Flow Analysis

The problem of balancing flows and stocks in material flow analysis when there is conflicting input information.

Scope and description

1. Introduction

Material flow analysis (MFA; see Factsheet "Material and Substance Flow Analysis") 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 (Dubois et al., 2014). This method was developed in the sixties to study the metabolism of urban systems (e.g., 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 material and energy balances. 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, m flows out and a stock level s of the form:

$$\sum_{i=1}^{n} IN_i = \sum_{j=1}^{m} OUT_j + \Delta S \qquad (1)$$

where : IN_i is inflow i (mass per unit time); OUT_j is outlow j (mass per unit time); ΔS is variation of stock (mass).

In general, by convention, if the sum of outflows exceeds the sum of inflows, the variation of stock is negative (the system has released mass). In the opposite case the variation of stock is positive (the system has stocked mass).

The most important (and time-consuming) step in MFA is the collection of data regarding the various flows in the system. In most situations, estimated inflows, outflows and stocks do not initially satisfy Eq. (1) (conservation of mass), in which case the analysist uses some form of data reconciliation.

2. Reconciliation in MFA

Data reconciliation has been defined as a "technique to optimally adjust measured process data so that they are consistent with known contraints" (Kelly, 2004). The traditional approach to data reconciliation (Narasimhan and Jordache, 2000) assumes that data come from measurements and that measurement errors follow a Gaussian distribution with zero average and a diagonal covariance matrix. The precision of each measurement *F* (understood as a mean value) is characterized by its standard deviation (σ_i). Data reconciliation is then performed by minimizing an objective function. Considering the simple case of a process with *m* entering flows and *n* exiting flows (with respective averages F_i and standard deviations σ_i; i=1..m+n) and designating initial estimators for the reconciled flows as F_i*, reconciliation is obtained by minimizing the following objective function:

$$\chi^2 = \sum_{i=1}^{m+n} \left[\frac{1}{\sigma_i} \left(F_i - F_i^* \right) \right]^2 \tag{2}$$

under the constraint that the sum of flows entering the process equals the sum of flows exiting the process (mass conservation):

$$\sum_{j=1}^{m} F_j^* = \sum_{k=m+1}^{m+n} F_k^* \tag{3}$$

Several MFA tools (e.g. BILCO; Durance et al., 2004; *STAN*; Brunner and Rechberger, 2004) perform this type of calculation. A drawback of the methodology in the context of MFA is related to the fact that in practical situations of MFA projects, available information often does not justify a representation of flows using single Gaussian distributions. Information is typically incomplete and/or imprecise and therefore other tools for representing uncertainty may be preferred to single probability distributions. The factsheet "*Parameter uncertainty in mineral intelligence analysis*" presents such alternative tools.

A practical tool for representing incomplete/imprecise information, especially coming from experts, is the well-known min-max interval. But as shown in the uncertainty factsheet, an expert may have information that allows him/her to express preferences within the interval. This yields the so-called possibility distributions (or fuzzy numbers) that are illustrated in Figure 5 of the uncertainty factsheet. Assuming information on flows and stocks in MFA are represented by possibility distributions (see Factsheet "Parameter uncertainty in mineral intelligence analysis"), reconciliation under fuzzy constraints can be performed using the method of Dubois et al. (2014).

To illustrate this method, Figure 1 shows the simple case of a single process with one inflow, one outflow and no stock. As seen in Figure 1, the inflow and outflow are affected by uncertainty. There are two ways of viewing this uncertainty:

- (i) the indicated values are preferred values within the intervals, resp., [45; 55] and [50;70]
- (ii) the indicated values are the mean values of Gaussian distributions with standard deviation, resp., 5 and 10.

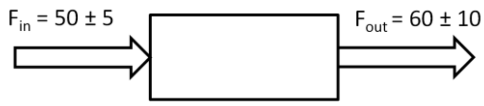

Figure 1 – Single process with one inflow, one outflow and no stock

Assuming the first interpretation, reconciliation is obtained by identifying the values that satisfy mass conservation and flow membership information (see Dubois et al., 2014). As illustrated in Figure 2, these conditions are satisfied for all values located within the intersection between the two possibility distributions: i.e. the interval [50; 55] with a "preferred" (most possible) value of 53.3. A value of e.g. 48 is not possible, as it does not lie within this intersection.

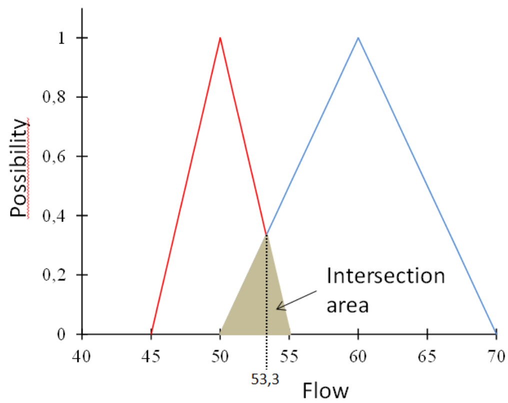

Figure 2 – Schematic illustration of reconciliation under fuzzy constraints

Considering now the second interpretation and applying equations (2) and (3) yields a reconciled flow of mean 52 and standard deviation 4.5. In this simple example, the two interpretations of uncertainty and their ensuing treatment yield very similar results, but this is not always the case. In particular if there were large discrepancies between estimates of inflow and outflow, the first method might indicate that it is not possible to find an intersection: either the model or the flow estimates are erroneous. On the other hand the second method will always yield a result because Gaussian probability distributions are defined over the interval [-∞; +∞]: a solution will be found albeit in areas of very low probability. This may be a problem in the case of outliers (erratic values) and therefore tools such as *STAN* incorporate checks to verify that the reconciled "solution" is not too remote from initial estimates.

3. Case study: application to rare earths in the EU-28

From 2011 through 2015, the ASTER project on rare earth flows and stocks in the EU-28 (see Guyonnet et al., 2015) was led by BRGM in partnership with Solvay, BIO by Deloitte and the University of Toulouse, with the suppor of the French Research Agency (ANR). Consistent with the standard MFA procedure, the system under inverstigation was first defined and then information regarding individual flows was collected. For the case of Neodymium in magnet applications, the defined system is depicted in Figure 3.

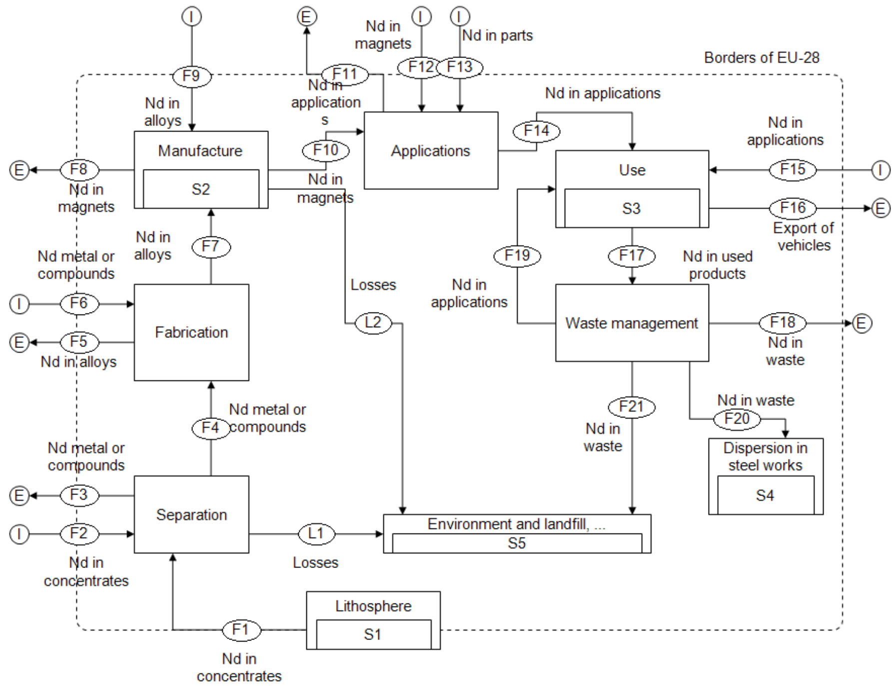

Figure 3 – System investigated for the case of Neodymium in magnet applications. Notes: F = flow; L = loss; I = import; E = export; S = stock variation

The following table lists the applications considered in the analysis.

Table 1. Neodymium-containing applications considered in this study

Applications using NdFeB permanent magnets	Applications using NiMH batteries
Electric and Non electric vehicles	Portable batteries (rechargeable batteries):
Hard drives	- Cameras
Cell phones	- Electric shavers
Laptops and desktops	- Cell phones and cordless phones
Wind turbine generators (REE magnet-based)	- Laptops
MRI machines	- Handheld tools
Refrigerators	- Remote-controlled toys
Washing machines	- Emergency lighting equipment
Air conditioners	Industrial batteries:
Cameras	- Hybrid vehicles (HEV)
Headphones and earphones	- Electrical aircraft systems
CD player	- Satellite pinpointing systems
Fax, printers, scanner	
Shavers and electric epilators	

Data sources included statistical databases (e.g. import, export and production data from EUROSTAT, World Trade Atlas, USGS, BGS, …), specialized reports (e.g., ROSKILL, company reports, …), data published in the literature regarding (i) quantities of REEs in components used in applications, (ii) weights of these components in applications and (iii) quantities of applications sold or used per year as reported by manufacturers, expert information, etc. An invaluable source of information in this study was Solvay's knowledge of the REE markets. The experts participating in the project were asked to provide estimates for flows, not as single values, but instead to:

- provide an interval which, based on their analysis, must include the actual flow value;
- express a preference within this interval.

While the experts provided estimates for flow intervals, they expressed the preferred values at the centre of the intervals. The resulting data are presented in Table 2. The year investigated is 2010.

Table 2. Values from the data mining (tons Nd metal, year 2010).

Flow/Stock	Min value	Max value	Preferred value	Flow/Stock	Min value	Max value	Preferred value
F1	100	300	200	F15	300	400	350
F2	2	10	6	F16	250	450	350
F3	150	250	200	F17	500	650	575
F4	5	25	15	F18	12	20	16
F5	40	70	55	F19	3	5	4
F6	150	250	200	F20	350	450	400
F7	120	220	170	F21	150	190	170
F8	2	10	6	S1	-100	-300	-200
F9	150	200	175	S2	70	120	87
F10	150	250	200	S3	180	400	290
F11	220	350	285	S4	300	500	400
F12	550	650	600	S5	150	260	205
F13	200	450	325	L1	5	15	10
F14	750	1000	875	L2	25	40	32.5

Notes : F = Flow ; S = stock ; L = losses; data derived from Guyonnet et al. (2015) but not identical.

The reconciliation of this data is presented below using the two methods illustrated above. It is reminded that each method corresponds to a distinct interpretation of the indicated uncertainty : imprecision in the first method, and random variability in the second method. For the second method, values were defined by considering that:

- the preferred value in Table 2 is a "mean" value;
- the ranges between the mean and the min values (or max values) represent 3 standard deviations (i.e., 99,7% of the data in a normal distribution).

Results of the reconciliation under fuzzy constraints (Dubois et al., 2014) yields values in Table 3.

Table 3. Flows and stocks of Nd in the EU-28: results of reconciliation under fuzzy constraints (tons Nd metal, year 2010)

Flow/Stock	Support		Optimal core	Flow /Stock	Support		Optimal core
	min	max			min	max	
F1	150	288	215.8	F15	300	400	351.3
F2	2	10	6.4	F16	250	450	358.5
F3	150	250	198.7	F17	515	650	579.8
F4	5	25	14.5	F18	12	20	16.1
F5	40	70	55.7	F19	3	5	3.9
F6	150	250	202.8	F20	350	450	393.2
F7	120	220	161.8	F21	150	190	166.6
F8	2	10	6.6	S1	-150	-288	-215.8
F9	150	200	173.3	S2	70	120	88.3
F10	150	250	208.0	S3	180	400	282.3
F11	220	350	285.0	S4	350	450	393.2
F12	550	650	605.0	S5	180	245	207.6
F13	200	450	337.6	L1	5	15	9.0
F14	750	1000	865.6	L2	25	40	32.0

Notes : F = Flow ; S = stock ; L = losses

The optimal core values are represented in the Sankey diagram below.

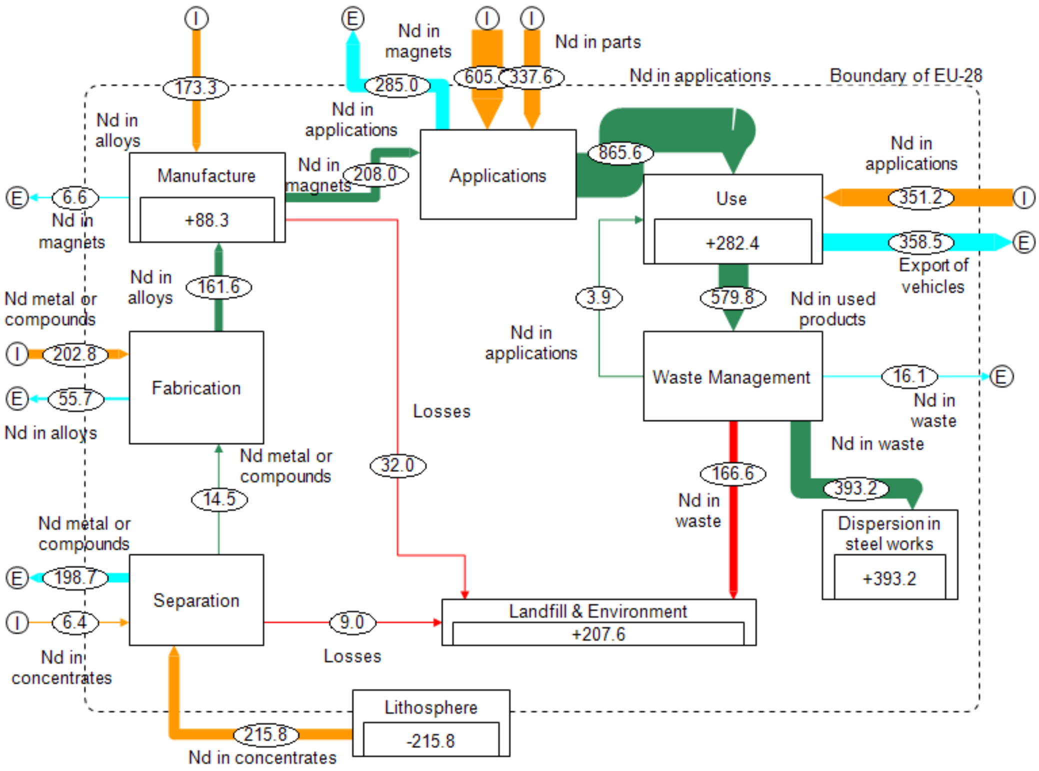

Figure 4 – Sankey diagram with values from reconciliation under fuzzy constraints (tons Nd metal, year 2010, diagram built with *STAN*)

The input data for reconciliation using the second method is shown in Table 4. It is reminded that “mean” values are the same as the “preferred” values of Table 2.

Table 4. Input for least-squares reconcilition

Flow /Stock	Mean	Sigma	Flow /Stock	Mean	Sigma
F1	200	33.3	F15	350	16.7
F2	6	1.3	F16	350	33.3
F3	200	16.7	F17	575	25.0
F4	15	3.3	F18	16	1.3
F5	55	5.0	F19	4	0.3
F6	200	16.7	F20	400	16.7
F7	170	16.7	F21	170	6.7
F8	6	1.3	S1	-200	33.3
F9	175	8.3	S2	87	5.7
F10	200	16.7	S3	290	36.7
F11	285	21.7	S4	400	33.3
F12	600	16.7	S5	205	18.3
F13	325	41.7	L1	10	1.7
F14	875	41.7	L2	32.5	2.5

Notes : F = Flow ; S = stock ; L = losses

Results of the reconciliation are presented in Table 5 and graphically in the Sankey diagram of Figure 5.

Table 5. Results of least-squares reconciliation

Flow /Stock	Mean	Sigma	Flow /Stock	Mean	Sigma
F1	215.2	15.2	F15	350.0	16.7
F2	6.0	1.3	F16	350.0	33.3
F3	196.2	14.9	F17	584.9	14.6
F4	15.0	3.3	F18	16.0	1.3
F5	54.6	4.9	F19	4.0	0.3
F6	204.8	12.2	F20	395.6	14.0
F7	165.2	12.2	F21	169.3	6.5
F8	6.0	1.3	S1	-215.2	15.2
F9	175.0	8.3	S2	99.6	22.1
F10	202.2	16.2	S3	280.6	51.7
F11	281.3	20.5	S4	395.6	14.0
F12	602.2	16.2	S5	211.8	7.2
F13	338.5	32.7	L1	10.0	1.7
F14	861.5	32.7	L2	32.5	2.5

Notes : F = Flow ; S = stock ; L = losses

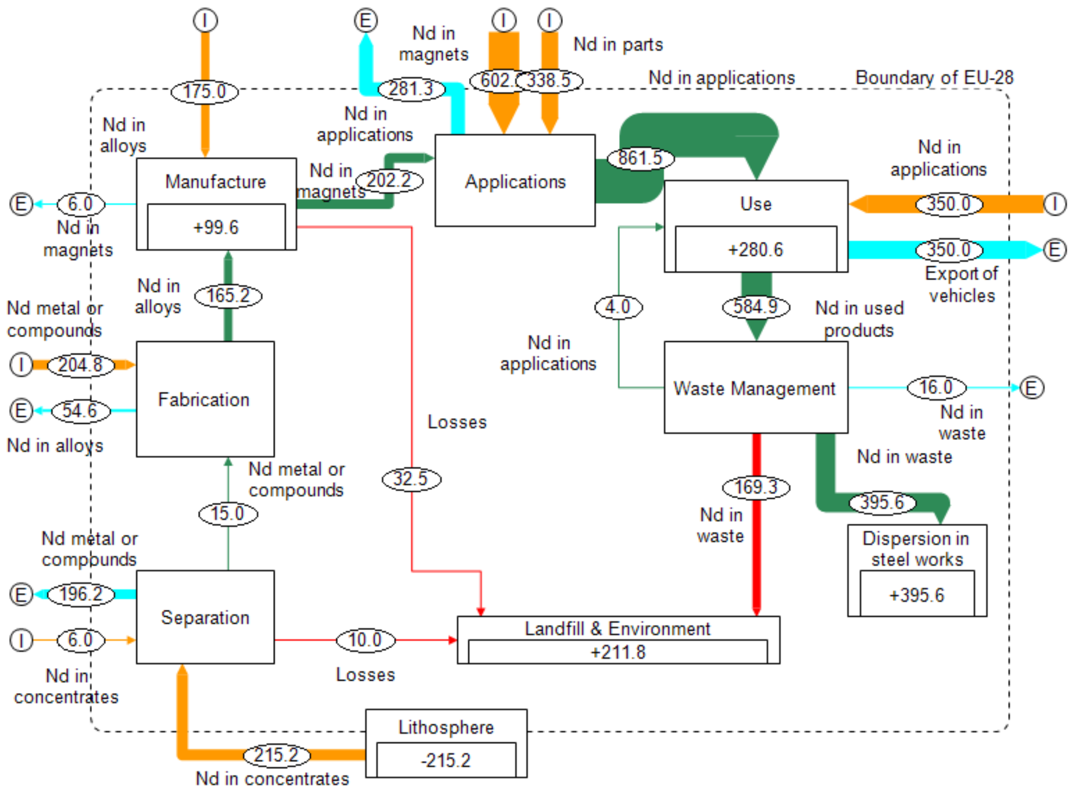

Figure 5 – Sankey diagram for values from least-squares reconciliation (tons Nd metal, year 2010, diagram built with *STAN*)

4. Conclusion

In this specific example, the two reconciliation methods yield very similar results: as seen in Table 6, values differ by 0.3 to 12.8%. But a basic question that should be addressed by the investigator at the data mining stage is: "*does the uncertainty in this data arise from random variability or from the incomplete/imprecise character of my knowledge regarding these parameters?*". If the answer is the latter, then representing the information using intervals (with or without preferences) may seem more "natural" than using means and standard deviations. Hence the method of reconciliation under fuzzy constraints proposed by Dubois et al. (2014).

As mentioned previously, the least-squares reconciliation method will always yield an answer and in some cases this may be misleading because the reconciled values may have very low levels of probability. The fuzzy-constraint method is less "robust" in the sense that it may fail to provide an answer. But this is an indication of inconsistency in the flow values or in the system structure and therefore that the investigator needs to reexamine the data further.

Table 6. Percent differences between results from the two methods

Flow /Stock	% difference	Flow /Stock	% difference
F1	0,3%	F15	0,4%
F2	6,3%	F16	2,4%
F3	1,3%	F17	0,9%
F4	3,4%	F18	0,6%
F5	2,0%	F19	2,6%
F6	1,0%	F20	0,6%
F7	2,1%	F21	1,6%
F8	9,1%	S1	0,3%
F9	1,0%	S2	12,8%
F10	2,8%	S3	0,6%
F11	1,3%	S4	0,6%
F12	0,5%	S5	2,0%
F13	0,3%	L1	11,1%
F14	0,5%	L2	1,6%

Contexts of use, application fields	-> contexts (e.g., environmental, economic, social assessment) -> which types of stakeholder questions are concerned? -> link to published studies that implement the method

Data reconciliation in MFA is used to:

- Optimize industrial ore treatment processes;
- Estimate urban flows/stocks;
- Identify bottlenecks for supply in a system of varying dimensions (plant, city, region, country, …);
- Identify potentialities for secondary resources;
- Etc. (see Factsheet '**Using dynamic MFA or system dynamic modelling**')

The uncertainty representation tools used above can be applied to:

- Measurement errors;
- Risk analysis;
- Life cycle assessment (LCA);
- Cost-benefit analysis;
- Etc.

The context and applicability is transverse, as variability and imprecision/incompleteness affect nearly all areas of mineral intelligence capacity.

Input parameters	-> which parameters are needed to run the method

For the input parameters to material flow analysis, refer to the Factsheet “Using dynamic MFA or system dynamic modelling”. For the application of reconciliation methodologies, input parameters depend on the selected method. For a purely stochastic approach to uncertainties, moments describing the probability distributiuons are required (e.g., averages and standard deviations in the case of Gaussian probability distributions). In the case of possibility distributions based on expert knowledge or scarce information, the support and core of the distributions are required (see Factsheet ‘**Parameter uncertainty in mineral intelligence analysis**’).

Type(s) of related input data or knowledge needed and their possible source(s)	-> which types of data are needed to run the method, from which sources could they come… -> could be qualitative data or quantitative data, and also tacit knowledge, hybrid, etc.

For sources of data and knowledge in material flow analysis, see Factsheet “Using dynamic MFA or system dynamic modelling”.

Model used (if any, geological mathematical, heuristic…)	-> e.g., geological model for mapping -> e.g., mathematical model such as mass balancing, matrix inversion, can be stepwise such as agent -based models, dynamic including time or quasidynamic specifying time series… -> can also be a scenario

The “model” is the “System” defined during the first steps of the material flow analysis. See Factsheet “Using dynamic MFA or system dynamic modelling”.

System and/or parameters considered	-> **the system can be described by its boundaries**. These can refer to a geographic location, like a country, or a city, the time period involved, products, materials, processes etc. involved, like flows and stocks of copper, or the cradle-to-grave chain of a cell phone, or the car fleet, or the construction sector, or the whole economy… -> **parameters** could possibly refer to geographic co-ordinates, scale, commodities considered, genesis of ore deposits and others…

The approach is generic and therefore suited to a wide variety of systems. Systems can be of various scales: an individual process, a processing plant, a city, a country, the world, …

Time / Space / Resolution /Accuracy / Plausibility…	-> to which spatio-temporal domain it applies, with which resolution and/or accuracy (e.g., near future, EU 28, 1 year, country/regional/local level…) -> for foresight methods can also be plausibility, legitimacy and credibility…

Case by case basis. Material Flow Analalysis applies to historical data but may also be used for prospectivity analysis.

Indicators / Outputs / Units	-> this refers to what the method is actually meant for. Units are an important part but that is most of the time not sufficient to express the meaning. For example, **the indicators used in LCA express the cradle-to-grave environmental impacts of a product or service**. This can be expressed in kg CO_2-equivalent. But also in €. Or in millipoints. Or in m^2year land use. -> for foresight methods the outputs are products or processes

The most common units in Material Flow Analysis are masses per unit time. For example tonnes per year.

Treatment of uncertainty, verification, validation	-> evaluation of the uncertainty related to this method, how it can be calculated/estimated

This is the object of the present Factsheet.

References cited	

Brunner, P.H., Rechberger, H. 2004. Practical Handbook of Material Flow Analysis. Lewis Publishers.
Dubois, D., Fargier, H., Guyonnet, D., Ababou, M. 2014. A fuzzy constraint-based approach to data reconciliation in material flow analysis. International Journal of General Systems, 43(8): 787-809.
Durance, M. V., Brochot, S., Mugabi, M. 2004. Material balance approach in paramter determination in bioleaching process. In: Computer Applications in Biotechnology. Proceedings of the 9th IFAC International Symposium. Edited by M. N. Pons and J. F. Van Impte. Oxford, Pergamon Press.
Guyonnet, D., Planchon, M., Rollat, A., Escalon, V., Tuduri, J., Charles, N., Vaxelaire, S., Dubois, D., Fargier, H. 2015. Material flow analysis applied to rare earth elements in the EU-28. Journal of Cleaner Production, 107: 215-228.
Kelly, J.D. 2004. Techniques for solving industrial nonlinear data reconciliation problems. Computers and Chemical Engineering, vol. 28, pp. 2837-2843, 2004.
Narasimhan, S., and C. Jordache. 2000. Data Reconciliation and Gross Error Detection: An Intelligent Use of Process Data. Houston: Gulf Publishing Company.
Wolman, A. 1965. "The Metabolism of Cities." Scientific American 213(3): 179–190.

Related methods	-> List of comparable methods, their particularities… -> link to one or several other existing fact sheet(s)

Related methods are:

- Sensitivity analysis
- Scenario analysis
- Geostatistics
- Probability boxes
- 2D Monte Carlo
- Bayesian methods
- …

Related factSheets:

- FactSheet: '**Parameter uncertainty in mineral intelligence analysis**'
- Factsheet: '**Using dynamic MFA or system dynamic modelling**'

Some examples of operational tools (CAUTION, this list is not exhaustive)	-> e.g., software… Only give a listing and a reference (publication, website/page…) -> **should be provided only if ALL main actors are properly cited**

Examples of MFA tools available either free or commercially:

- STAN (Brunner & Rechberger, 2004)
- BILCO (Durance et al., 2004)

Key relevant contacts	-> list of relevant **types** of organisations that could provide further expertise and help with the methods described above.

Technical University Vienna (STAN): http://www.stan2web.net/
Caspeo (BILCO): http://www.caspeo.net/fr

Glossary of acronyms /abbreviations used	-> Definition