

# Exploring Generalisations for Sustainability Assessment in Medicine Manufacturing Networks

*Ettore Settanni (E.Settanni@eng.cam.ac.uk)*

*Institute for Manufacturing, University of Cambridge, Cambridge, UK*

*Jagjit Singh Srail*

*Institute for Manufacturing, University of Cambridge, Cambridge, UK*

*Ekaterina Yatskovskaya*

*Institute for Manufacturing, University of Cambridge, Cambridge, UK*

*Tomás S. Harrington*

*Institute for Manufacturing, University of Cambridge, Cambridge, UK*

## Abstract

Generalisations or ‘rules of thumb’ are widely used in industry to make initial assessments on the sustainability impacts of products. This paper re-examines the principles underpinning these generalisations in assessing the environmental impacts of medicine manufacturing network configurations. Each principle is illustrated through the application of selected methods, tools, and data using an exemplar case of a major non-proprietary analgesic. The findings identify the kind of generalised knowledge claims that are possible using alternative approaches, and which methodological issues may arise.

**Keywords:** generalisations, sustainable supply chains, pharmaceuticals

## Introduction

Simplifications based on some degree of generalisation are commonly employed in an effort to capture the environmental impact of an entire supply chain without imposing daunting data collection requirements on each chain member (Pagell and Shevchenko, 2014). For example, the NHS England attributing 21% of its carbon footprint to the procurement of pharmaceuticals is partly based on generalised relationships between certain characteristics of synthesising an Active Pharmaceutical Ingredient (API) and the amount of greenhouse gases released (Penny *et al.*, 2015).

Generalisations across pharmaceutical product categories, if discovered, are believed to greatly simplify the otherwise challenging task of assessing environmental sustainability in medicine manufacturing networks – see, for example Hartmann (2015); Soete *et al.* (2014). While there is no single approach to generalisations discovery, these are typically proposed as ‘quicker’ and ‘simpler’ – and hence more desirable – alternatives to standardised, analytical approaches such as Life Cycle Assessment (LCA), which often provide complementary insights for the purposes of mathematical Supply Network Design – see Eskandarpour *et al.* (2015) for an overview.

*Table 1 – Views on, and exemplar of generalisations*

Views on generalisation (based on Metcalfe, 2005)		Example application to sustainability evaluation in pharmaceutical supply chains		
Epistemology	Research agenda	Methodology	Data collection method	Reference
Scientific	Pursue of general truth by induction and falsification	Statistical analysis of quantitative data	Observations; Experiments	Soete <i>et al.</i> (2014)
Systems thinking	See something in terms of boundaries, purpose, connectivity	Diagrammatic visualisation; simultaneous equations	Documents; Observations	Jiménez-González <i>et al.</i> (2010)
Interpretative	Generate critical heuristic and rules of thumb for future action	Survey; Case study research; Mixed methods	Interviews; Questionnaire	Hartmann (2015)

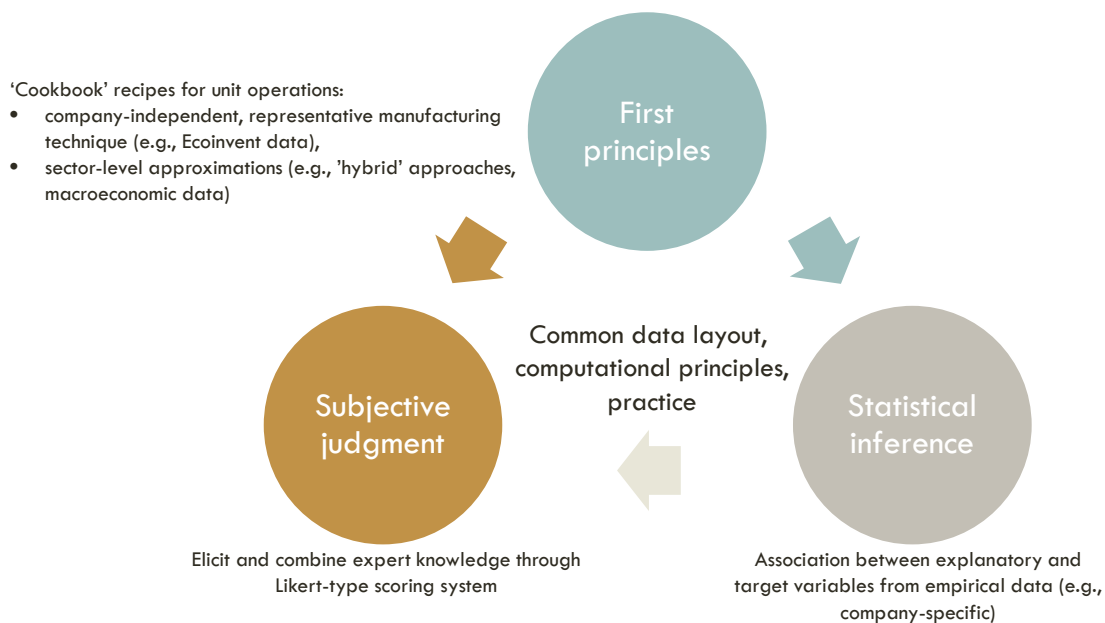
This paper aims to systematically address the emerging need for simplification in assessing the ‘end-to-end’ environmental repercussions of current and future medicine manufacturing network configurations. It does so by 1) identifying a set of generalisation principles underpinning the approaches more commonly used in the academic literature; and 2) illustrating each principle through the application of selected methods, tools, and data to the same exemplar case based on a large-selling non-proprietary analgesic. These points are elaborated, in the ‘materials and methods’, and ‘findings’ sections. The ‘discussion’ section outlines the key contribution of the paper, which is the simultaneous exploration of approaches that are often seen as mutually exclusive, to answer questions such as: what kind of generalised knowledge claim are supported? Which methodological tensions may arise? Which synergies can be exploited? The paper closes by pinpointing research limitations, and providing directions for future work.

### **Materials and methods**

The term ‘generalisations’ denotes knowledge claims, gathered under certain methodological conditions, which exhibit properties that make them applicable to more than one situation (Metcalfe, 2005). Table 1 shows how these properties vary depending on the epistemological stance taken, and provides specific examples from sustainable supply chain research with a focus on the pharmaceutical sector.

The number of approaches and tools available within each epistemology is potentially vast due the inter-disciplinary nature of investigating sustainability-related aspects in supply chain network design (Haskins, 2006). To keep the focus, the examples reported in Table 1 are limited to approaches and tools that are relevant at the interface between the following adjacent but distinct fields: Supply chain network design (SCND), simplified/streamlined Life Cycle Assessment (LCA), and Hot Spot Analysis (HSA). While the application of LCA to support SCND is widespread (see e.g., Eskandarpour *et al.*, 2015), the relationship between LCA and HSA is still debated (UNEP-SETAC, 2014), and the use of the latter in SCND is relatively unexplored (Liu and Srai, 2011).

Within the chosen scope, the analytical approaches to generalisation identified by synthesis of the relevant literature (not reported here in full due to space constraints), classified, and then put to use on an example centred on paracetamol (also known as acetaminophen and by its chemical name, N-(4- hydroxyphenyl) acetamide), one of the largest-selling non-proprietary drugs worldwide and the most widely used and prescribed first line analgesic in UK.



*Figure 1 – Generalisation principles*

Although without reference to the pharmaceutical sector, the application of different types of analysis to a single hypothetical example instead of discussing each model separately has been previously used to evaluate differences and similarities in a practical way (Bouman *et al.*, 2000).

## Findings

The most common generalisation principles found in the literature on sustainable supply chains of pharmaceutical products can be summarised as shown in Fig. 1, and as described below:

- First principles (FP): generalise the understanding of detailed causal mechanisms. A system thinking epistemological stance is typically taken. Representative approaches (off-the-shelf tools) include: LCA (IfU Hamburg Umberto NXT Universal, with Ecoinvent database); flowsheet simulation (Intelligen SuperPro Designer);
- Statistical inference (SI): identifies associative relationships based on a track record of evidence, taking a scientific epistemological stance. Representative approaches (tools) include: data visualisation and descriptive statistics (Tableau); multivariate regression and supervised segmentation (specialised packages for the statistical programming language *R* (R Development Core Team, 2008); Artificial Intelligence (EstiMol);
- Subjective judgment (SJ), based on individual knowledge elicitation, generalises the relative importance of aspects that are relevant for taking a certain decision. Assumes an interpretative epistemological stance. Representative approaches may vary, but mostly employ scoring systems whereby evidence is gathered from literature/case study (for an overview, see UNEP-SETAC, 2014).

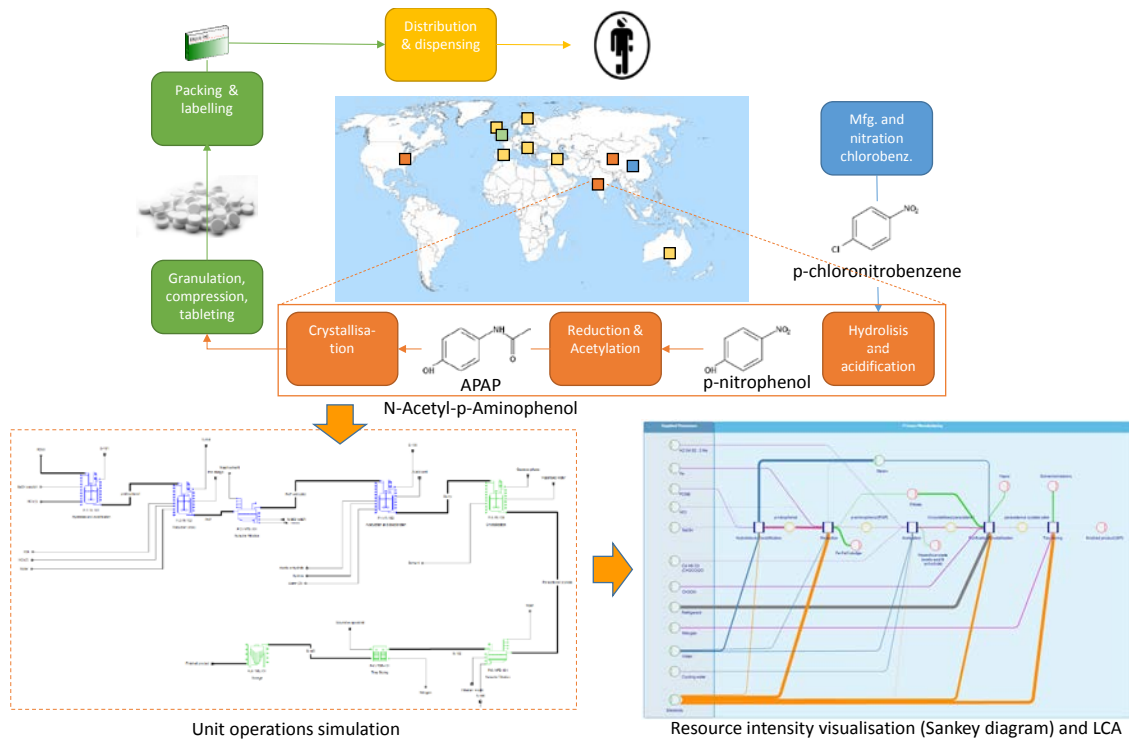
The results obtained from the application of the selected approaches are illustrated in dedicated sub-sections below.

### *Evaluation of the exemplar case using First Principles (FP)*

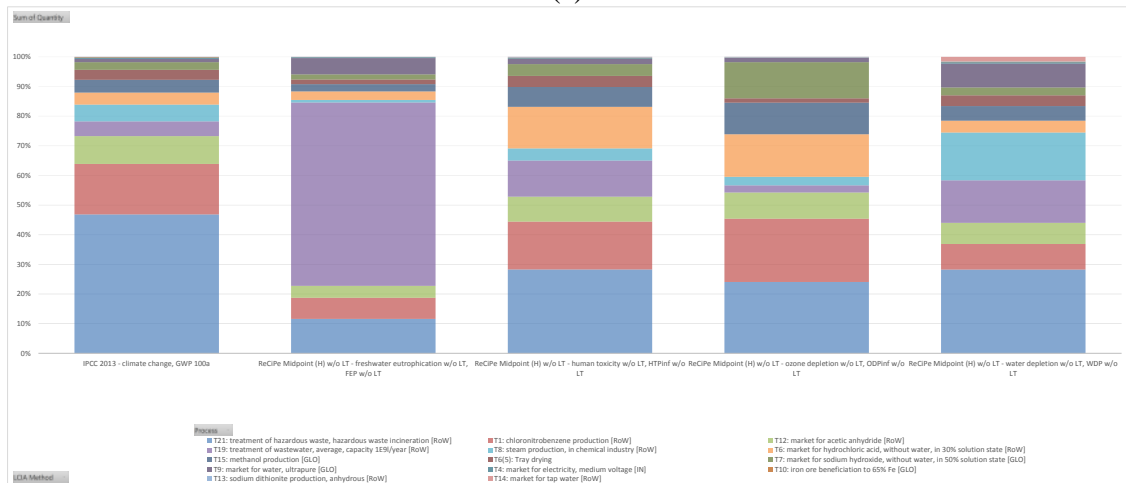
Although paracetamol is a well-known pharmaceutical product, the authors could not easily identify references in the public domain that provided a sufficiently transparent account of the material flows ‘end-to-end’, across the manufacturing system of interest. A hypothetical system of unit operations for the manufacture of paracetamol was therefore derived according to FP generalisations from data available in the public domain. The steps followed to this purpose are summarised below and illustrated schematically in Fig. 2 (more details are provided elsewhere, see Settanni, 2016):

- The boundaries of the foreground system span from the manufacture the Active Pharmaceutical Ingredient (API) to the manufacture of a solid dosage form.
- The scope of the analysis was limited to a subset of manufacturing operations, and the associated material and energy flows, although the operations associated with transportation and packaging, were excluded at this stage.
- In the absence of site-level data that are typically available for specific businesses, it was assumed that the manufacturing of the API takes place in India, where most suppliers of paracetamol into the UK are located, according to current site authorisation data (<http://eudragmdp.ema.europa.eu/>).
- Several routes for the synthesis of the API, and alternative technologies to manufacture the dosage forms (mainly batch vs flow chemistry) are available – see, respectively, Joncour (2014); Soete *et al.*, (2013). The gathered data refer to one of the routes most exploited commercially, which seems representative of the prevailing manufacturing technology in key manufacturing countries such as India.
- Molar mass balances for the reactive and non-reactive systems involved in most manufacturing steps were based on literature and patent search. Public domain data on typical batch size and manufacturing capacity at industrial scale are sparse.
- The amount and composition of the material flows between the manufacturing operations at the relevant scale were estimated from the data thus gathered, using a combination of manual computations and computer-aided process simulation.
- The estimated flows were imported in an LCA tool, where the foreground system was expanded, whenever possible, and the Life Cycle Inventory and Impact Assessment calculations were carried out. Expansion of the foreground system implies the retrieval, by database search, and addition of ‘generic’ unit operations which are assumed to be representative of all the activities taking place upstream.

A variety of impact categories and impact assessment methods are available, the discussion of which is beyond the scope of this paper. For illustrative purposes Fig. 2b shows the contribution of the unit processes mapped in Fig. 2a across five impact categories quantified according with two assessment methods. In the remainder of this paper, the focus is placed on the impact category ‘Global Warming Potential’ (GWP), expressed as kgCO<sub>2</sub>-eq. (here, computed according to the IPCC methodology using Umberto NXT). A focus on GWP is common in estimating a product ‘carbon footprint’. According the FP generalisation of the manufacturing system of interest here, the estimated carbon footprint associated with 1kg of API is circa 78 kgCO<sub>2</sub>-eq (or circa 68 kgCO<sub>2</sub>-eq per kg of finished product, coated tablets), 57% of which is attributed to the unit processes included in the primary manufacturing phase, whereas 43% represents ‘upstream’ impacts which are embedded in the procured materials. The breakdown in Fig. 2b reveals that the main contributors are the treatment of hazardous waste – namely iron oxide sludge generated during the reduction of *p*-nitrophenol into *p*-aminophenol – and the manufacture of the main starting material, *p*-chloronitrobenzene, both of which are characteristic features of the synthesis route assumed upfront for the API.



(a)



(b)

Figure 2 – (a) Hypothetical manufacturing network for a non-proprietary analgesic based on FP generalisations (b) Impact assessment across five categories

### Evaluation of the exemplar case using statistical inference (SI)

The use of statistical inference, or ‘parametric’ estimation is based on the assumption that the environmental impact of a product (e.g., its carbon footprint) is a dependant variable that has the propensity to be related statistically to a set of potentially informative attributes of such product. Hence, generalisations developed on the basis of the SI principle are typically established in the form of associative relationships between some dependant variable (or ‘target’ to be predicted) and some knowable attributes that relate with it (parameters), for which previous observations exist. Approaches that apply this principle typically favour the speed of results over good explanation facilities, as opposed to the FP case examined in the previous sub-section.

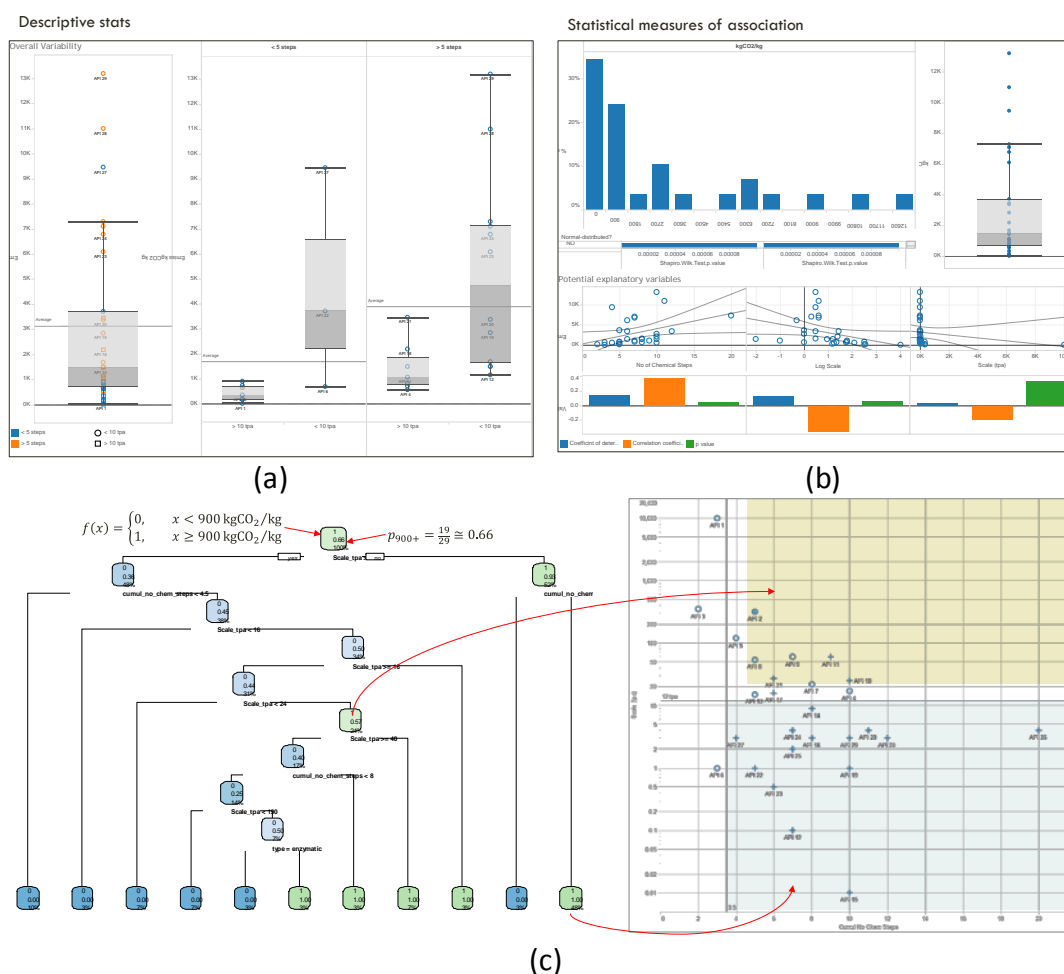


Figure 3 – Analysis of carbon footprint data and product attributed across for multiple APIs

Specific applications of SI to pharmaceutical products use linear and non-linear regression to establish generalisations in the form of associative relationships between variables that may be chosen differently across studies. Example include the estimate of an APIs' carbon footprint from such parameters as the amount of organic solvent used, as well as process parameters like molar efficiency and synthesis time (Soete *et al.*, 2014). Industry-wide initiatives include a tool that apply SI for the quick approximation of the footprint of pharmaceutical products developed by the Association of British Pharmaceutical Industry (ABPI) in collaboration with the Carbon Trust, based on data contributed by major industrial players (Carbon Trust, 2013).

To apply the SI principle to the case exemplar considered here, an excerpt was obtained under confidentiality agreement from a broader dataset underpinning the ABPI tool. Due to the nature of the data excerpt, only the primary manufacture of the paracetamol API could be assessed. The data was preliminary assessed through descriptive statistics to characterise the values recorded for each variable – three continuous, two categorical. By segmenting the carbon footprint values reported across the anonymised products it appeared that larger scales exhibit less variability and systematically lower emissions (Fig. 3a). Each continuous variable was tested for normality, and the strength and direction of its linear association with the other variables was assessed. For the sake of example, Fig. 3b shows the normality test, the coefficient of correlation, coefficient of determination ( $R^2$ ), and  $p$ -value of the test of significance of the overall regression

between the variable ‘carbon footprint’, the log of the production scale, and the number of synthesis steps).

Since no clear-cut linear association between the variables emerged from the data, a supervised segmentation approach was applied to automate the process of ranking the explanatory variables according with their ability to predict the value of a modified target – whether the carbon footprint would be less than 900 kgCO<sub>2</sub>-eq per kg of API. The main difference being that, unlike regression, segmentation aims to predict whether something will happen rather than ‘to what extent’ it will happen (Provost and Fawcett, 2013). A tree-structure model was generated using the *rpart* package for the statistical software *R* (shown on the leftmost side of Fig. 3c). The model demonstrates that most APIs with a carbon footprint above 900 kgCO<sub>2</sub>-eq per kg of product (crosses on the rightmost part of Fig. 3c) are characterised by manufacturing capacity of less than 12 ton per annum (tpa) and a number of steps above 3.5. In principle, the exemplar case of paracetamol API modelled according to FP falls into this category, since a small scale was hypothesised for the sake of simplicity. However, the result obtained from a FP environmental impact assessment of the GWP associated with the API was significantly below the value expected applying SI based on past observations.

While regression requires process and product-related parameters as an input, ‘black box’ tools such as EstiMoL explore relationships between the molecular structure of the API of interest and target variables such as GWP (Wernet *et al.*, 2010). In the case of paracetamol, the EstiMol tool produces an estimate of 5.965 kgCO<sub>2</sub>-eq ( $\pm$  6.491).

#### *Evaluation of the exemplar case using subjective judgment (SJ)*

An alternative to estimating the ‘magnitude’ of the environmental impact of manufacturing pharmaceutical products is to express the relative importance of the contribution to that impact from the relevant elements of the supply chain as scores (ordinal data), typically obtained by eliciting expert judgment. For example, Liu and Srari (2011) elicit subjective judgment to evaluate the maturity of an organisation’s ‘green’ capabilities in product design, procurement, production, logistics, and reverse logistics, thus discovering possible generalisations based on a number of business cases. This approach has been applied to multiple case studies in the pharmaceutical supply chain (Hartmann, 2015).

Similarly, the Wuppertal Institute’s Sustainability Hot Spot Analysis (UNEP-SETAC, 2014) also employs evidence gathered in the form of ordinal data, either from literature or through experts interviews, to score the resource intensities in a two-step process: first, within each life-cycle stage, across groups of resources; then between each life-cycle stage an evaluation of each stage of a product life cycle in relation to the other stages; finally, multiplying the scores obtained in the previous steps. However, the scoring in the second step is not carried out as a pairwise comparison, and the final step does not provide explicit computational devices to reduce subjective bias.

To address these limitations, a variation of Wuppertal’s three-step approach is suggested, based on an application of the principles of the Analytical Hierarchy Process - AHP (Saaty, 1990), and applied to the paracetamol example as follows (Fig. 4):

- Step 1: four features (direct emissions and emissions embedded in resources, energy, and waste), are scored against each other in terms of their relative importance. This is accomplished in a pairwise criterion comparison matrix;
- Step 2: three aggregated life cycle stages of the pharmaceutical product (API and, dosage form manufacture; packaging) are scored against each other with respect to each feature. This is accomplished in a number of pairwise options comparison matrices.

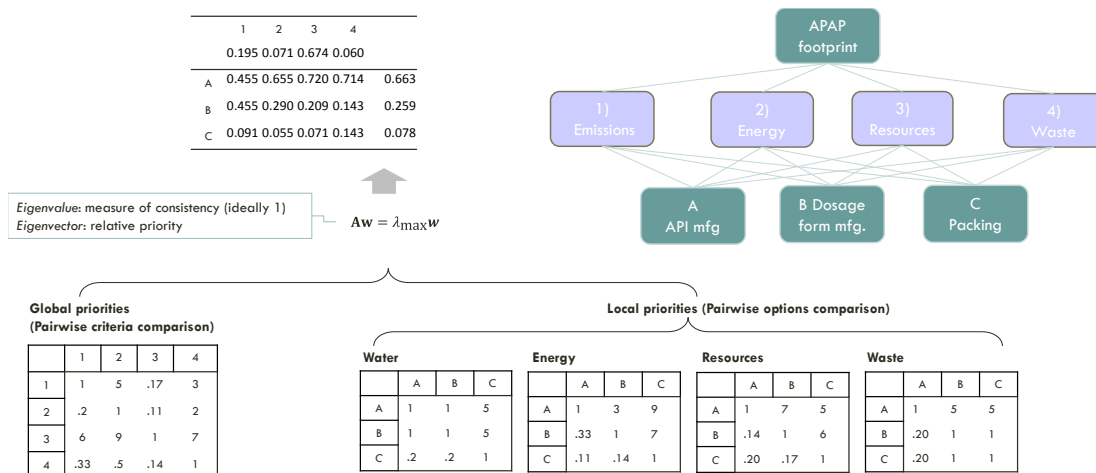


Figure 4 – Application of AHP in the subjective judgment-based Hotspot Analysis of paracetamol

- Step 3: The relative importance (ranking) of each life cycle stage considering all the resource groups simultaneously is produced by means of AHP matrix algebra

The scores in the example are for illustrative purposes only, and based on evidence from limited literature available on paracetamol (Hartmann, 2015; Soete *et al.*, 2013), as well as the researcher’s experience. The eigenvalues and eigenvectors required to obtain the ranking according to the matrix equation shown in Fig. 4 were computed using MATLAB®. The API manufacturing stage ranks highest – which aligns with the results of the assessment from FP previously carried out.

## Discussion

In the previous section, three generalisation principles were identified (FP, SI and SJ), and illustrated by applying representative approaches and tools to the same streamlined example based on the manufacture of paracetamol API, and a solid dosage form. With the exception of SI, the application of each approach was carried out using, whenever possible, publicly available data related to the hypothetical case. The findings from such an application are summarised in Tab. 2, and are discussed below.

Generalisations based on FP mostly concern individual unit operations, and provide the most detailed insights in terms of environmental burden attributed to individual unit operations/product life cycle phases. However, in the absence of real-life industry data, the results generated by simulation can be very sensitive to the assumptions made upfront. For example, in the simplified case considered here, the unit operations model is not optimally designed to recover waste flows such as wastewater (e.g., filtrate) and the iron sludge generated in the reduction stage. While the presence of significant effluents at various stages of API manufacturing is in line with some industry outlooks, the presence of non-optimised water flows at the plant level may raise issues, especially in water stressed areas such as India. Also, the estimates obtained using this approach appear to be very sensitive to the assumptions regarding the amount of fresh and recirculated heat exchange media such as the refrigerant assumedly used during crystallisation. Finally, most materials used in the secondary manufacturing were difficult to retrieve while expanding the model, and were therefore cut-off.

Approaches based on the SI and SJ generalisation principles are typically meant to overcome the complexities of FP by simplification. SI, in particular, aims to generalise



Table 2 – Summary of findings with regards to the GWP impact category

	Generalisation principle applied					
	FP		SI		SJ	
	Value	Unit*	Value	Unit*	Value	Unit
Raw materials**	32.9	kg/CO <sub>2</sub> -e	NA		NA	
API manufacture	44.5	kg/CO <sub>2</sub> -e	>900	kg/CO <sub>2</sub> -e	0.663	-
Secondary manufacture	0.06	kg/CO <sub>2</sub> -e	NA		0.259	-
Packing***	0.91	kg/CO <sub>2</sub> -e	NA		0.078	-

\*Relative to 1 kg of crude paracetamol (API); \*\*For the manufacture of the API; \*\*\*No end-of-life scenario.

the relationship between some informative attributes of a manufacturing system and some ‘target’ variable. However, as demonstrated by the application to the exemplar case, a sufficiently track record of products assessed ‘end-to-end’, most likely from FP, is needed to provide the necessary data for SI to apply. Even in when data and free tools are available (for example, by the pharmaceutical industry associations as in the UK), there seems to be a large variation between the values obtained from FP and the values inferred. Since the underpinning data (e.g., carbon footprint of a product) are the result of computations rather than direct observation/measurement, legitimate concerns arise regarding whether and to what extent each data point contributed by different firms and then employed for the purposes of SI is computed consistently – e.g., following the same principles, and using similar tools and computational procedures.

Finally, approaches based on the SJ principle, such as scoring, may be appealing when the emphasis is on providing heuristics for taking action along a supply chain rather than estimating a metric. In this paper, SJ was applied to synthesise evidence from the literature. Often, such evidence is limited for specific pharmaceutical products, and difficult to align, thus making a meaningful comparison difficult. It was also observed that current SJ approaches may benefit from techniques that help reduce subjective bias by using rigorous computational devices to combine scores – as, for example, AHP.

## Conclusion

This paper presented a first attempt to identify a set of generalisation principles underpinning the approaches more commonly used to make knowledge claims about the environmental repercussions of current and future medicine manufacturing network configurations. Each principle was illustrated by application of selected methods, tools, and data to the same exemplar case based on a large-selling non-proprietary analgesic. While most works treat these approaches and the underpinning principles as mutually exclusive, the application of different types of analysis to a single hypothetical example has proven useful to evaluate differences and similarities in a practical way.

This paper has several limitations, including assumptions used in building a simulation model of the manufacturing operations from first principles in the absence of industry-specific information. Also, only a limited number of approaches and tools available have been considered. Despite its limitations, the findings of this research help to address, both conceptually and through example applications, the challenge to capture the environmental impact of the entire supply chain, without imposing daunting data collection campaigns on each chain members.

To endorse a truly ‘end-to-end’ perspective, it is necessary to consider the pharmaceutical product ‘in use’. A more appropriate definition of the functional unit, than one unit mass of API or dosage form is necessary to this purpose. Future work should expand the boundaries of the analysis downstream to capture the complexities of the

medicine distribution and utilisation systems, to support the assessment of more patient-centric and distributed medicine networks.

### Funding acknowledgment

Support from the Advanced Manufacturing Supply Chain Initiative (Grant No. 35708-233529, ReMediES – Reconfiguring Medicines End-to-end supply), and the EPSRC (Grant No. EP/I033459/1, Centre for Innovative Manufacturing in Continuous Manufacturing and Crystallisation; and Grant No. EP/K02888X/1, Engineering Driven Sustainable Supply Networks-A UK/India Collaborative Study) is gratefully acknowledged.

### References

- Bouman, M., Heijungs, R., van der Voet, Ester, van den Bergh, Jeroen C. J. M and Huppes, G. (2000), “Material flows and economic models: an analytical comparison of SFA, LCA and partial equilibrium models”, *Ecological Economics*, Vol. 32 No. 2, pp. 195–216.
- Carbon Trust (2013), *New carbon footprint tool launched to calculate the carbon footprint of pharmaceutical products* [<http://www.europeanpharmaceuticalreview.com/20994/news/>]
- Eskandarpour, M., Dejax, P., Miemczyk, J. and Péton, O. (2015), “Sustainable supply chain network design: An optimization-oriented review”, *Omega*, Vol. 54 No. 0, pp. 11–32.
- Hartmann, C. (2015), “Impact of environmental sustainability on the configuration of supply chains in the pharmaceutical industry”, Master of Philosophy, Institute for Manufacturing, University of Cambridge, Cambridge, UK, 2015.
- Haskins, C. (2006), “Application of Systems Engineering and Other Disciplines to the Design of a Sustainable Supply Chain”, 4th Annual Conference on System Engineering, April 7-8, Los Angeles.
- Jiménez-González, C., Ponder, C.S., Hannah, R.E. and Hagan, J.R. (2010), “Designing a Sustainable Pharmaceutical Industry: The Role of Chemical Engineers”, in am Ende, David J. (Ed.), *Chemical Engineering in the Pharmaceutical Industry*, John Wiley & Sons, Inc, pp. 57–65.
- Joncour, R., Duguet, N., Metay, E., Ferreira, A. and Lemaire, M. (2014), “Amidation of phenol derivatives: a direct synthesis of paracetamol (acetaminophen) from hydroquinone”, *Green Chem*, Vol. 16 No. 6, pp. 2997–3002.
- Liu, Y. and Srari, J.S. (2011), “Sustainable Supply Chain Configuration: Proposing a Hot Spot Analysis Methodology”, Proceedings of the International Conference on Materials Science, Metal & Manufacturing, 12-13 December, Singapore.
- Metcalf, M. (2005), “Generalisation: Learning Across Epistemologies”, *Forum Qualitative Social Research*, Vol. 6 No. 1.
- Pagell, M. and Shevchenko, A. (2014), “Why Research in Sustainable Supply Chain Management Should Have no Future”, *Journal of Supply Chain Management*, Vol. 50 No. 1, pp. 44–55.
- Penny, T., Collins, M., Whiting, A. and Aumônier, S. (2015), *Care Pathways: Guidance on Appraising Sustainability*. [<http://www.sduhealth.org.uk>]
- Provost, F. and Fawcett, T. (2013), *Data science for business*, O'Reilly, Sebastopol, CA.
- R Development Core Team (2008), *R: A Language and Environment for Statistical Computing*, Vienna, Austria, available at: <http://www.R-project.org>.
- Saaty, T.L. (1990), “Decision making by the analytic hierarchy process: Theory and applications - How to make a decision: The analytic hierarchy process”, *European Journal of Operational Research*, Vol. 48 No. 1, pp. 9–26.
- Settanni, E. (2016), *Estimating material balances in the manufacture of crude paracetamol*. Unpublished report. Institute for Manufacturing, University of Cambridge.
- Soete, W. de, Debaveye, S., Meester, S. de, van der Vorst, G., Aelterman, W., Heirman, B., Cappuyns, P. and Dewulf, J. (2014), “Environmental sustainability assessments of pharmaceuticals: an emerging need for simplification in life cycle assessments”, *Environmental science & technology*, Vol. 48 No. 20, pp. 12247–12255.
- Soete, W. de, Dewulf, J., Cappuyns, P., van der Vorst, G., Heirman, B., Aelterman, W., Schoeters, K. and van Langenhove, H. (2013), “Exergetic sustainability assessment of batch versus continuous wet granulation based pharmaceutical tablet manufacturing: a cohesive analysis at three different levels”, *Green Chem*, Vol. 15 No. 11, pp. 3039–3048.
- UNEP-SETAC (2014), *Hotspots Analysis: mapping of existing methodologies, tools and guidance and initial recommendations for the development of global guidance*. [<http://www.lifecycleinitiative.org/>]
- Wernet, G., Papadokonstantakis, S., Hellweg, S. and Hungerbühler, K. (2009), “Bridging data gaps in environmental assessments: Modeling impacts of fine and basic chemical production”, *Green Chem*, Vol. 11 No. 11, pp. 1826–1831.