SKU classification: A literature review and conceptual framework

van Kampen, Tim J.; Akkerman, Renzo; van Donk, Dirk Pieter

Published in:
International Journal of Operations and Production Management

Link to article, DOI:
10.1108/01443571211250112

Publication date:
2012

Citation (APA):
SKU classification: A literature review and conceptual framework

Tim J. van Kampan
Department of Operations
Faculty of Economics and Business
University of Groningen
Groningen, The Netherlands
t.j.van.kampen@rug.nl

Renzo Akkerman
Department of Management Engineering
Technical University of Denmark
Kgs. Lyngby (Copenhagen), Denmark
renzo@dtu.dk

Dirk Pieter van Donk
Department of Operations
Faculty of Economics and Business
University of Groningen
Groningen, The Netherlands
d.p.van.donk@rug.nl

Published in
International Journal of Operations and Production Management

SKU classification: A literature review and conceptual framework

Abstract
Purpose – Stock Keeping Unit (SKU) classifications are widely used in the field of production and operations management. Although many theoretical and practical examples of classifications exist, there are no overviews of the current literature, and general guidelines are lacking with respect to method selection. The purpose of this paper is to systematically synthesise the earlier work in this area, and to conceptualise and discuss the factors that influence the choice of a specific SKU classification.

Design/methodology/approach – This paper structurally reviews existing contributions and synthesises these into a conceptual framework for SKU classification.

Findings – How SKUs are classified depends on the classification aim, the context and the method that is chosen. Three main production and operations management aims where found: inventory management, forecasting and production strategy. Within the method three decisions are identified to come to a classification: the characteristics, the classification technique and the operationalisation of the classes.

Research limitations/implications – Drawing on our literature survey, we conclude with a conceptual framework describing the factors that influence SKU classification. Further research could use this framework to develop guidelines for real-life applications.

Practical implications – Examples from a variety of industries and general directions are provided that managers could use to develop their own SKU classification.

Originality/value – This paper aims to advance the literature on SKU classification from the level of individual examples to a conceptual level and provides directions on how to develop a SKU classification.

Keywords SKU classification, demand classification, production strategy, forecasting, inventory management

Paper type Literature review

1. Introduction
In production and operations management, companies often have to deal with many different products, or Stock Keeping Units (SKUs). Here, SKUs refer to items of stock that are completely specific as to function, style, size, colour, and, usually, location (Silver et al., 1998, p. 32). The production and inventory policies of these different SKUs are influenced by the characteristics of the product. Differences in annual sales volume, predictability of demand, product value, or storage requirements might result in different production and inventory policies. As a consequence, companies that sell a wide variety of SKUs often struggle with the control of their production and inventory systems. Therefore, in real-life situations, it is generally seen as advantageous to distinguish a limited number of SKU classes based on the characteristics of these SKUs. This enables companies to make decisions on production strategy (e.g. make-to-stock or make-to-order), production and inventory management and customer service for entire SKU classes rather than for each product separately.

In order to create a SKU classification, two simple questions need to be answered: how many classes are used and how are the borders between the classes determined. Various approaches and techniques exist to classify SKUs. A well-known approach is the ABC analysis, which usually classifies product groups based on either demand value or demand volume. Another well-known approach is the FNS technique, which distinguishes product classes based on demand rate (Fast, Normal, and Slow). Empirical studies seem to use approaches inspired by the specific context, and it is often far from clear why a certain method was employed or whether other approaches could also have been used. Technical papers provide and develop analytical tools to classify SKUs, but it remains unclear under what circumstances or context they should be applied. It seems
that there is a lack of guidance as to which techniques should be used to classify SKUs and which characteristics should be included under specific circumstances.

In the absence of papers that provide an overview of contributions on SKU classification, combined with a lack of papers that structure the classification process there is no guidance for academics and practitioners on this topic. Syntetos et al. (2009) confirm that classification has not received sufficient academic attention given the implications of the decision-making in that area. Therefore, the aim of our paper is to structure the previous work on SKU classification in order to provide directions on how SKU classifications can be designed. Our review explores what factors drive choices in SKU classifications and what techniques are appropriate in different circumstances. We argue that much can be gained in research and practice by knowing these factors and their relationships.

The first step in our approach is to systematically review the existing SKU classifications and to identify the aims, techniques used, and SKU characteristics adopted. The insights gained are used to discuss how different factors influence a SKU classification. The outcome of this step is expressed in a conceptual framework that supports the design of SKU classifications and provides the basis for further theory building in this area. The outcomes of this study also have practical relevance as they might guide managers in selecting an appropriate method for classifying SKUs.

This paper is structured as follows. The next section further introduces SKU classification and elaborates on the main research questions. Subsequently, we describe the research approach, and present the results from the literature survey. In the final parts of the paper, the results are discussed and conclusions are drawn.

2. Motivation and research questions
The main aim of any SKU classification is to use the similarity of products with regards to different properties to systematically classify products. Krishnan and Ulrich (2001) identified four perspectives within the academic community from which product properties are studied: marketing, organizations, engineering design, and operations management. In this paper we focus on the classification of products from the production and operations management perspective.

Within production and operations management, inventory management and forecasting are fields where a variety of SKU classifications is traditionally used to support decision-making. One of the oldest and best-known classification approaches is the ABC analysis which is used in inventory management (see Silver et al. (1998) for the technique, Schomer (1965) for an early application or Zhang et al. (2001) for a spreadsheet extension). The aim of the ABC analysis is that, if one focuses on the relatively small number of products that represent a major part of the sales volume (i.e. the A products), relatively large reductions in inventory costs can be obtained. This builds heavily on the insights advanced by Pareto (1906). However, some authors argue that cost reductions mainly occur through the appropriate treatment of the C products (see Viswanathan and Bhatnagar, 2005; Teunter et al., 2010 for a discussion on this topic). Other characteristics than volume are also used in classifications for inventory management. For example, the XYZ technique differentiates, as with the ABC technique, between three categories of products, but this time based on variability in demand (see Schönsleben, 2003).These basic techniques are widely used and have been implemented in commonly used software tools, such as SAP’s ERP and APS software (Hoppe, 2006), to make it easier for practitioners to tailor production and inventory activities to the demand characteristics of their products.

SKU classification is also frequently used in forecasting. Selecting the proper forecasting method is important to be able to balance the costs of keeping inventory and
the risk of stock-outs. The latter aspect is especially important in controlling spare parts due to the impact the absence of a spare part can have (see Cavalieri et al. (2008) for an overview on the management of spare part inventories). Here, the demand is generated by the process requiring the spare parts, often leading to a situation where there is only an occasional need for a certain part. The reliance of the production process on the availability of the specific spare parts is an important consideration in managing spare part inventories where forecasting these low volumes is difficult. Therefore, the selection of the appropriate forecasting techniques for spare parts is an important decision that can be supported by a SKU classification (see Syntetos et al., 2005; Boylan and Syntetos, 2008).

SKU classification is not limited to inventory management and forecasting. They are also used to determine the production strategy. Several contributions have been made in this respect. For instance, Hoekstra and Romme (1992) classify SKUs to decide whether to make them to stock or to order. The related issue of finding the right level of postponement for different product classes was studied by Pagh and Cooper (1998). In the same decade, Fuller et al. (1993) discussed the tailoring of logistics, and Fisher (1997) discussed the appropriate supply chain for a specific product. Numerous authors followed these seminal works, e.g. by refining the classification methods presented (Stavrulaki and Davis, 2010), by developing industry-specific frameworks (Soman et al., 2004), or by demonstrating the value of using classification methods (Christopher et al., 2009). Between the above mentioned works, there are obviously differences in focus and in the characteristics that are used, but all are based on some kind of classification of SKUs and they all provide insights in relation to production strategy.

The characteristics that are used to classify SKUs are numerous. Examples of characteristics that are used in different approaches are volume and variability (D'Alessandro and Baveja, 2000), different types of variability (Talluri et al., 2004), unit cost, dollar value, criticality and lead time (Ramanathan, 2006), duration of life cycle, time window for delivery, volume, variety, and variability (DWV^3) (Childerhouse et al., 2002; Christopher et al., 2009).

Syntetos et al. (2009) study spare part management and state that stock classification has been overlooked. They remark that the issue of classification has not received as much academic attention as the implications of the relevant decision-making in that area would require. We would argue that this is not only the case for spare parts but for SKUs in general. Even though many applications can be found, no overview exists of the applications or techniques that can be used. As a consequence, there is a lack of guidance for practitioners who want to use a SKU classification within productions and operations management. At the same time the existing applications are often based on, or inspired by, a certain production environment (e.g. D'Alessandro and Baveja (2000) use a specific batch sizes to distinguish classes), and it is not always clear if an approach has wider applicability.

In the absence of structural guidance on SKU classification and the scattered applications found in the literature we argue that it is appropriate to synthesize the existing work and strive towards a conceptual foundation for SKU classification. A systematic review of the literature on applications of SKU classification will provide the ingredients to build such a foundation. Further, we aim to provide guidance on how SKUs can be classified. From the above discussion four main research questions emerge:

RQ1: What are the aims in SKU classification?
RQ2: Which characteristics are used to classify the SKUs?
RQ3: Which classification techniques are used?
RQ4: How is the classification influenced by the context?

The answers to the above questions will provide the building blocks for a conceptual framework along with a basis for theory building. Meredith (1993) provides two necessary conditions for external validity of conceptual frameworks. The first one (it should be based on earlier studies) is rather straightforward as we conduct a systematic literature review. The second one (it should be based on real world descriptions) is taken into account by mainly considering descriptions of applications of SKU classification in the literature.

3. Research method and data analysis

There are numerous situations in which SKU classifications are used. In our review, we focus on contributions to the production and operations management literature. In many papers, SKU classification is not an aim in itself but an approach adopted to achieve another aim (e.g. a SKU classification is used to minimize the inventory value). As such, it is a challenge to find papers that classify SKUs, and also to cover the entire scope in which classification studies might be found. To address this challenge, a broad, structured literature review was conducted. Using the ISI Web of Science database (with subject areas ‘management’ and ‘operations research & management science’) enabled us to cover not only influential journals in the field of production and operations management but also journals in adjacent fields. Since SKU classification is often not the main topic of a paper we searched for combinations of keywords in a single sentence to find potentially relevant papers. We used primary keywords related to the object to be classified (e.g. demand, product, ABC, SKU) and secondary keywords related to the classification process (e.g. classification, characterisation, category). The secondary keyword ‘analysis’ was only used in combination with ‘ABC’ as the primary keyword, as coupling it with other primary keywords mainly lead to inappropriate papers such as ones on product analysis. Table 1 lists all the keywords used. Our initial search resulted in 479 papers in 85 journals (search conducted October 2008).

<table>
<thead>
<tr>
<th>Primary keywords</th>
<th>Secondary keywords</th>
</tr>
</thead>
<tbody>
<tr>
<td>SKU, Product, Products, Demand, ABC</td>
<td>Classification, Classifying, Categorization, Categorisation, Categorizing, Categorising, Characterization, Characterisation, Characterizing, Characterising, Category, Categories, Segregation, Segregating, Classes</td>
</tr>
<tr>
<td>ABC</td>
<td>Analysis</td>
</tr>
</tbody>
</table>

In this initial selection, the primary and secondary keywords appeared in a single sentence in either the abstract or the title. In a further filtering, our main aim was to check whether these papers actually dealt with the classification of SKUs, or whether these words just happened to appear together. In other words, we checked whether the secondary keyword actually related to the primary keyword. As a result, the majority of papers were rejected, and only 91 papers were retained for possible inclusion in this review. Many of these papers did use various SKU classifications (for example, papers on inventory rationing would use a customer’s price setting as their basis) but did not discuss how they came to these classes, and were therefore excluded. As a consequence, this phase reduced our initial selection to 20 papers. Due to the fact that older publications are not always fully indexed in the ISI database, the literature discussions in these 20 papers were investigated to find references to other studies on SKU classification. In our search for additional papers, we focused on contributions that outlined and applied SKU classification techniques. An additional 54 papers were thus
considered, of which 25 were selected for inclusion after a further check of the papers. Therefore, our final selection amounted to 45 papers. The structured literature review is schematically summarised in Figure 1.

**Phase 1 Literature search**
Primary keywords related to the object to be classified and secondary keywords related to the classification process.

**Phase 2 Abstract selection**
Whether or not the primary and secondary keyword related to each other.

**Phase 3 Paper selection**
Whether or not the paper actually dealt with the classification of demand.

**Phase 4 Backward literature search**
Literature discussions in the selected papers searched to find additional papers on the subject of demand classification.

**Phase 5 Additional paper selection**
Whether or not the paper actually dealt with the classification of demand.

**Final selection**

479 abstracts found for abstract review

91 papers selected for full paper review

20 papers selected

54 additional papers found for full paper review

25 additional papers selected

45 papers

*Figure 1 – Systematic selection of papers on SKU classification*

The papers that were included were read by the authors and a number of details were distilled to answer the research questions. These where: the aim behind the SKU classification (RQ1), the characteristics upon which the classification was based (RQ2), the classification technique used (RQ3), and finally the industry in which the classification was performed and related context-specific aspects, when present (RQ4). Whenever there were doubts regarding one of these attributes this was discussed between the authors.

In the introduction and motivation sections three aims (RQ1) were already identified for classifying SKUs: inventory management, forecasting and production strategy. In the analysis of the papers these aims are used as a starting point to discuss the reasons to classify SKUs.

The characteristics that were collected for RQ2 showed a great diversity. In order to compare the different approaches and to structure the outcomes we used a number of categories. Existing literature provided different categories to characterize SKUs in different situations. Fuller et al. (1993) aim to tailor logistic processes to the wishes of customers. They presented eight questions/dimensions to analyse whether a product is shipped according to a logistically distinct method (Fuller et al., 1993, p. 93). Bartezzaghi et al. (1999) studied how demand lumpiness is generated by different market characteristics and provide five classification dimensions. Christopher and Towill (2000) came up with five characteristics that influence the design of value stream delivery strategies. In the presence of this variety of structures to classify products from various perspectives we tried to come up with a general structure. As SKU characteristics mainly result from customer demand and the characteristics of the product, we based our categories on the concept of a customer order for a product. We define the order as a demanded amount of a product by a customer at a moment in time. From this, we identified four main characteristics in SKU classification: volume,
product, customer and timing. The abovementioned characteristics used by Fuller et al. (1993), Bartezzaghi et al. (1999) and Christopher and Towill (2000) can be grouped by our four main characteristics (see Table 2). By using these four main characteristics we expect to be able to frame most of the characteristics used for a SKU classification from a production and operations management perspective.

<table>
<thead>
<tr>
<th></th>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>Volume</td>
<td>Sales volume, order size</td>
<td>Variety of each customer’s request (CoV of demand)</td>
<td>Volume, variability</td>
</tr>
<tr>
<td>Product</td>
<td>Profit margin, relations to other products, services included with delivery, handling and storage requirements, substitutability</td>
<td>Numerousness of customers, heterogeneity of customers, correlation between customers behaviour</td>
<td>Variety, duration of life cycle</td>
</tr>
<tr>
<td>Customer</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Timing</td>
<td>Delivery speed/window/frequency</td>
<td>Frequency of order placing</td>
<td>Time window for delivery</td>
</tr>
</tbody>
</table>

A large variety of techniques (RQ3) can be used to come to a classification. To structure these techniques we looked at the type of data needed for each approach. In relation to forecasting, Amstrong (2001, p. 9) identified the special character of judgemental knowledge sources as opposed to statistical knowledge sources. Assuming that such a general distinction would also be relevant for inventory management and production strategy, we use this distinction to categorize the techniques found in the papers in being based on either (i) judgemental or (ii) statistical data sources.

For RQ4 we tried to identify how the context influences a classification of SKUs. However, given the broad explorative nature of this question, no direct rules could be established upfront. Therefore, we decided to list what is reported in the papers related to the specific context and tried to structure and discuss the emerging findings in the results and discussion section. To give an initial indication of context, we listed the industry in which the approach was applied.

4. Results
The purpose of classification schemes is to determine the number of classes and the borders between the classes. This is done through the specification of the classification parameters and their cut-off values. Opposed to the seemingly simple nature of this purpose a multitude of alternatives exist to do so. Table 3 lists the papers that were selected in our search process, including information on the aim of the SKU classification, the industry in which the technique was applied, and details relating to the four characteristic categories we have identified (volume, product, customer and timing). In line with the traditional ABC approach, quite a few papers use the (annual) demand value which is a combination of two characteristics from different categories: volume and product. For clarity reasons we therefore split this into demand volume and unit cost. To create uniformity in the overview, we sometimes slightly adapted the terminology used in the papers. For example, annual demand rate (Gelders and Van Looy, 1978), annual sales (Huiskonen et al., 2005), demand volume (Partovi and Hopton, 1994), monthly demand (Porras and Dekker, 2008) have been made uniform by using demand volume where we put the specific period between brackets. In the following subsections, we will discuss the results following the four research questions posed in Section 2.
4.1. Classification aims
Various reasons for classifying SKUs can be found in the papers in our sample. Table 3 shows that most of the work is applied in inventory management or, to a lesser extent, in forecasting. Few papers have a wider scope and use SKU classification to support decision-making on an appropriate production strategy.

The inventory management contributions mainly set out to determine order/production quantities, reorder points, safety stock, etc. for different SKU classes. The characteristics on which the classes are determined vary, and will be discussed in the next section. In many examples, SKU classes are used to reduce inventory levels by focusing on the fast moving stocks (similar to the ABC analysis). However, when all the products are slow movers (as with spare parts), SKU class selection is influenced by other characteristics (Flores and Whybark, 1987; Williams, 1984; Eaves and Kingsman, 2004). Studies that apply specifically to slow moving spare parts are indicated in Table 3 by ‘(spare parts)’ alongside the industry description. Here, the category customer does often not relate to external customers but to the internal production process which in most situations only has an occasional need for the spare parts.

The second largest classification aim is related to forecasting methods. Here, the classification of SKUs facilitates the selection of an appropriate forecasting method for the determined product classes (Syntetos et al., 2005). An important aspect in these studies is the demand pattern over time. Whether the demand pattern is smooth, sporadic or lumpy greatly influences the performance of the different forecasting methods. A specific situation is the forecasting of slow moving products, which often have an intermittent and erratic demand pattern (Boylan and Syntetos 2008, p. 484).

In relation to production strategy, several issues are addressed. Aitken et al. (2003) focus on product lifecycles and the resulting differences in supply chain strategy for associated products. D’Alessandro and Baveja (2000) use a product classification to choose between different distribution channels, including customer prioritisation and make-to-order or make-to-stock decisions. Fisher (1997) aims to determine the best supply chain for a product, largely based on demand predictability. His main idea is to use a physically efficient, lean, make-to-stock supply chain for predictable demand, whereas unpredictable demand should be handled within a market-responsive, agile, make-to-order supply chain. Similar guidelines are provided by Li and O’Brien (2001) and Vonderembse et al. (2006).
Table 3 – Summary of studies on SKU classification

<table>
<thead>
<tr>
<th>Study</th>
<th>Aim</th>
<th>Industry</th>
<th>Characteristics</th>
</tr>
</thead>
<tbody>
<tr>
<td>Aitken et al. (2003)</td>
<td>PS</td>
<td>Lighting</td>
<td>Demand volume, Product volume, market variety, order winners,</td>
</tr>
<tr>
<td>Bhattacharya et al. (2007)</td>
<td>IM</td>
<td>Pharmaceutical industry</td>
<td>Demand volume (daily), Unit costs, lead time, perishability, storage costs</td>
</tr>
<tr>
<td>Boylan et al. (2008)</td>
<td>IM/ FOR</td>
<td>Automotive, aerospace, chemical</td>
<td>Demand volume (mean + Coefficient of Variation (CoV))</td>
</tr>
<tr>
<td>Canen and Galvao (1980)</td>
<td>IM</td>
<td>Manufacturing</td>
<td>Demand volume* (annual), Unit cost*</td>
</tr>
<tr>
<td>Canetta et al. (2005)</td>
<td>IM</td>
<td>Electronics - component inventory</td>
<td>Demand volume (monthly (mean + CoV))</td>
</tr>
<tr>
<td>Cavalieri et al. (2008)</td>
<td>IM</td>
<td>Process industry (spare parts)</td>
<td>Demand volume, Unit cost, Criticality, number of installations</td>
</tr>
<tr>
<td>Chakravarty (1981)</td>
<td>IM</td>
<td>General</td>
<td>Demand volume, Unit cost</td>
</tr>
<tr>
<td>Chen et al. (2008)</td>
<td>IM</td>
<td>General</td>
<td>Demand volume* (annual), Unit cost*, criticality, lead time</td>
</tr>
<tr>
<td>Chrisman (1985)</td>
<td>IM</td>
<td>Cylinder parts</td>
<td>Demand volume* (annual), Unit cost*</td>
</tr>
<tr>
<td>D’Alessandro and Baveja (2000)</td>
<td>PS</td>
<td>Chemical</td>
<td>Demand volume (weekly, mean + CoV), Unit cost*</td>
</tr>
<tr>
<td>Duchessi et al. (1988)</td>
<td>IM</td>
<td>Spare parts</td>
<td>Demand volume* (annual), Unit cost*</td>
</tr>
<tr>
<td>Eaves and Kingsman (2004)</td>
<td>FOR</td>
<td>Air Force</td>
<td>Demand size variability, Lead time variability, Unit cost*, Criticality</td>
</tr>
<tr>
<td>Ernst and Cohen (1990)</td>
<td>IM</td>
<td>Automotive (spare parts)</td>
<td>Demand volume (monthly, mean + CoV), Returns volume (annual), Criticality</td>
</tr>
<tr>
<td>Fisher (1997)</td>
<td>PS</td>
<td>General</td>
<td>Demand predictability, Unit cost*, lead time, Criticality</td>
</tr>
<tr>
<td>Flores and Whybark (1986)</td>
<td>IM</td>
<td>Manufacturing</td>
<td>Demand volume*, Unit cost*, lead time</td>
</tr>
<tr>
<td>Flores and Whybark (1987)</td>
<td>IM</td>
<td>Manufacturing and service firm (spare parts)</td>
<td>Demand volume*, Unit cost*</td>
</tr>
<tr>
<td>Flores et al. (1992)</td>
<td>IM</td>
<td>General</td>
<td>Demand volume* (annual), Unit cost*, unit cost (mean), lead time</td>
</tr>
<tr>
<td>Gajpal et al. (1994)</td>
<td>IM</td>
<td>Manufacturing (spare parts)</td>
<td>-</td>
</tr>
<tr>
<td>Gardner (1990)</td>
<td>FOR</td>
<td>Military (spare parts)</td>
<td>Demand volume, Criticality (impact)</td>
</tr>
<tr>
<td>Gelders and van Looy (1978)</td>
<td>IM</td>
<td>Petrochemical industry</td>
<td>Demand volume (annual), Unit cost</td>
</tr>
<tr>
<td>Ghobbar and Friend (2002)</td>
<td>FOR</td>
<td>Aviation (spare parts)</td>
<td>Demand size (squared CoV), Mean inter-demand interval</td>
</tr>
</tbody>
</table>

* indicates annual or monthly data, CoV indicates Coefficient of Variation.
<table>
<thead>
<tr>
<th>Study</th>
<th>Application/Purpose</th>
<th>Demand Volume</th>
<th>Lead Time</th>
<th>Criticality</th>
<th>Inter-Demand Interval Predictability</th>
<th>Number of Orders/Lead Times</th>
<th>Notes</th>
</tr>
</thead>
<tbody>
<tr>
<td>Güvenir and Erel (1998)</td>
<td>IM</td>
<td>Demand volume* (annual)</td>
<td>Unit cost*, lead time</td>
<td>-</td>
<td>-</td>
<td>Number of orders for the item in a year</td>
<td></td>
</tr>
<tr>
<td>Harhalakis et al. (1989)</td>
<td>IM</td>
<td>Demand volume* (monthly)</td>
<td>Unit Cost*, lead time</td>
<td>-</td>
<td>-</td>
<td>Number of orders for the item in a year</td>
<td></td>
</tr>
<tr>
<td>Huiskonen (2001)</td>
<td>IM</td>
<td>Demand volume</td>
<td>Specificity, unit cost</td>
<td>-</td>
<td>-</td>
<td>Inter-demand interval predictability</td>
<td></td>
</tr>
<tr>
<td>Huiskonen et al. (2005)</td>
<td>IM</td>
<td>Demand volume (annual)</td>
<td>Annual sales of A/B products in the same order as the C-product</td>
<td>Proportion of C product sales to customer types</td>
<td>-</td>
<td>Number of orders (annual)</td>
<td></td>
</tr>
<tr>
<td>Kobbacy and Liang (1999)</td>
<td>IM</td>
<td>Demand volume (mean + variance), randomness</td>
<td>Lead time (mean + variance)</td>
<td>-</td>
<td>-</td>
<td>Seasonal patterns, trend</td>
<td></td>
</tr>
<tr>
<td>Mukhopadhyay et al. (2003)</td>
<td>IM</td>
<td>Demand volume (annual)*, during replenish lead time</td>
<td>Unit cost*</td>
<td>Criticality to production process</td>
<td>-</td>
<td>-</td>
<td></td>
</tr>
<tr>
<td>Onwuhuolu and Dube (2006)</td>
<td>IM</td>
<td>Demand volume* (annual)</td>
<td>Unit cost*</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td></td>
</tr>
<tr>
<td>Partovi and Anandarajan (2002)</td>
<td>IM</td>
<td>Demand volume (annual)</td>
<td>Unit cost, ordering cost, lead time</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td></td>
</tr>
<tr>
<td>Partovi and Burton (1993)</td>
<td>IM</td>
<td>Demand volume (annual)</td>
<td>Unit cost, lead time</td>
<td>Criticality</td>
<td>-</td>
<td>-</td>
<td></td>
</tr>
<tr>
<td>Partovi and Hopton (1994)</td>
<td>IM</td>
<td>Demand volume (annual)</td>
<td>Unit cost, lead time</td>
<td>Criticality</td>
<td>-</td>
<td>-</td>
<td></td>
</tr>
<tr>
<td>Porras and Dekker (2008)</td>
<td>IM</td>
<td>Demand volume (monthly)</td>
<td>Unit cost</td>
<td>Criticality</td>
<td>-</td>
<td>-</td>
<td></td>
</tr>
<tr>
<td>Portougal (2002)</td>
<td>FOR</td>
<td>Demand volume</td>
<td>Profit margin</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td></td>
</tr>
<tr>
<td>Ramanathan (2006)</td>
<td>IM</td>
<td>Demand volume* (annual)</td>
<td>Unit cost*, criticality, lead time</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td></td>
</tr>
<tr>
<td>Reid (1987)</td>
<td>IM</td>
<td>Demand volume (annual)</td>
<td>Unit cost</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td></td>
</tr>
<tr>
<td>Ritchie and Kingsman (1985)</td>
<td>IM</td>
<td>Demand volume (weekly, empirical distribution)</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td></td>
</tr>
<tr>
<td>Sani and Kingsman (1997)</td>
<td>FOR</td>
<td>Demand volume (annual)</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td></td>
</tr>
<tr>
<td>Stanford and Martin (2007)</td>
<td>IM</td>
<td>Demand volume (annual)</td>
<td>Unit cost</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td></td>
</tr>
<tr>
<td>Syntetos et al. (2005)</td>
<td>FOR</td>
<td>Demand size (squared CoV)</td>
<td>-</td>
<td>-</td>
<td>Mean inter-demand interval</td>
<td>-</td>
<td></td>
</tr>
<tr>
<td>Williams (1984)</td>
<td>IM/ FOR</td>
<td>Demand (lumpiness)</td>
<td>-</td>
<td>-</td>
<td>Mean number of lead times between demands, variance of lead time</td>
<td>-</td>
<td></td>
</tr>
<tr>
<td>Wu et al. (2006)</td>
<td>FOR</td>
<td>Demand pattern (lifecycle)</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td></td>
</tr>
<tr>
<td>Zhou and Fan (2007)</td>
<td>IM</td>
<td>Demand volume* (annual)</td>
<td>Unit cost* (mean), lead time</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td></td>
</tr>
</tbody>
</table>

Aim = The dominant application purpose: Inventory Management (IM), Forecasting (FOR) or Production Strategy (PS)
* The study uses (annual) demand value; we have converted this to demand volume and unit cost
4.2. Characteristics used for SKU classification

As explained in Section 3, we distinguish four main categories for SKU characteristics: volume, product, customer and timing. In terms of volume, most authors include the demand volume over a certain period in their classification, often in combination with unit cost to calculate the demand value. Especially for applications in inventory management, demand value often reflects inventory investment and it is argued that products with high values warrant special attention. However, according to Flores and Whybark (1987), very little specific guidance has been given on how to actually pay ‘special attention’ and improve performance. Alongside the absolute volume, a number of authors (e.g. D'Alessandro and Baveja, 2000; Ernst and Cohen, 1990) also include the variability in volume, mostly by calculating a Coefficient of Variation (CoV) over several demand periods (e.g. weekly, monthly). Other authors suggest analysing the volume of individual orders (e.g. Ghobbar and Friend, 2002; Kobbacy and Liang, 1999; Syntetos et al., 2005). Our overview shows a limited number of such papers, and data on individual orders seems to be used mostly in relation to forecasting. However, this does not imply that studies in forecasting only use data on individual orders. Finally, some alternative approaches have been proposed within the category volume. For example, Wu et al. (2006) try to identify demand patterns over a product’s lifecycle, to improve forecasting for other products. Here, the focus is thus not only on absolute volumes but also on how these volumes evolve over a product’s lifecycle.

The second category, product, is found in most papers. Related to our earlier remark on the frequent use of demand value, the product’s unit cost is one of the most common characteristics used. However, we did find a large range of other characteristics in this category, such as lead times related to production or supply. Further, context-specific characteristics such as product perishability, commonality and substitutability have also been used.

The third category, customer, is not used often. Huiskonen et al. (2005) provide an example where the importance of the customer is used. In their approach C products (in the ABC classification) become more important to meet customer requirements if they are sold to important customers or have a relation to A products. Further, the use of customer characteristics seems limited to the classification of spare parts. This reflects the importance of that part to the customer, where it should be reiterated that the customer of a spare part is often the internal production process. Criticality reflects the effects and financial consequences of not being able to deliver a spare part within the required lead time. The criticality may be determined informally by the insight of an expert (e.g. the VED classification, which labels products as vital, essential or desirable) or by more formal methods such as failure mode effects and criticality analysis (FMECA, see Boylan and Syntetos, 2008) or the analytical hierarchy process (AHP, see Gajpal et al., 1994)).

The final category, timing, seems relatively neglected in literature. The most notable measure used is the inter-demand interval. This measure gives an insight into the frequency of orders, and can be used to estimate when a next order for a product can be expected. It is therefore not surprising that the studies including such timing aspects tend to be those focused on forecasting. Johnston and Boylan (1996) were the first who formally established the importance of the inter-demand interval as a classification parameter. A few authors have investigated other timing related characteristics. Examples are SKU classes based on seasonality or trends (e.g. Ernst and Cohen, 1990; Kobbacy and Liang, 1999).
4.3. Techniques used for SKU classification

The papers we studied show a large variety in techniques to come to a classification (see Table 4). As introduced in Section 3, we distinguish between two types of knowledge sources: (i) judgemental and (ii) statistical. Techniques based on expert judgement are ways to capture the opinions of managers. Statistical knowledge sources are based on data of a number of SKU characteristics. Within the statistical techniques there is a wide variety in the complexity of the technique and in the number of characteristics used. In Table 4 they range from simple guidelines based on a limited number of SKU characteristics to advanced mathematical models that can more easily deal with a large number of SKU characteristics.

Table 4 – Summary of SKU classification techniques used

<table>
<thead>
<tr>
<th>Knowledge source</th>
<th>Technique</th>
<th>Study</th>
</tr>
</thead>
<tbody>
<tr>
<td>Judgemental</td>
<td>VED</td>
<td>Cavalieri et al. (2008), Mukhopadhyay et al. (2003)</td>
</tr>
<tr>
<td></td>
<td>AHP</td>
<td>Flores et al. (1992), Gajpal et al. (1994), Partovi and Burton (1993), Partovi and Hopton (1994)</td>
</tr>
<tr>
<td></td>
<td>TOPSIS</td>
<td>Bhattacharya et al. (2007)</td>
</tr>
<tr>
<td></td>
<td>Distance modelling</td>
<td>Chen et al. (2008)</td>
</tr>
<tr>
<td></td>
<td>FSN/FNS</td>
<td>Cavalieri et al. (2008), Gelders and Van Looy (1978), Mukhopadhyay et al. (2003)</td>
</tr>
<tr>
<td></td>
<td>Bi-criteria ABC</td>
<td>Cavalieri et al. (2008), Flores and Whybark (1986), Flores and Whybark (1987), Harhalakis et al. (1989)</td>
</tr>
<tr>
<td></td>
<td>Cluster analysis</td>
<td>Canetta et al. (2005), Duchessi et al. (1988), Ernst and Cohen (1990), Wu et al. (2006)</td>
</tr>
<tr>
<td></td>
<td>Neural networks</td>
<td>Huiskonen et al. (2005), Partovi and Anandarajan (2002)</td>
</tr>
<tr>
<td></td>
<td>Genetic Algorithm</td>
<td>Güvenir and Erel (1998)</td>
</tr>
</tbody>
</table>

The main idea of the judgemental techniques is to extract the sometimes tacit knowledge held by managers. Such techniques are used to determine the criticality of a product (as in the VED technique) or to rank different characteristics using pair-wise comparisons in the AHP or TOPSIS technique (Technique for Order Preference by Similarity to Ideal Solution). The results from pair-wise comparisons can subsequently be used as inputs for mathematical models. For instance, the AHP technique used by Flores et al. (1992) starts with pair-wise comparisons of both the importance of the SKU characteristics and the performance of products in terms of these characteristics. These results are subsequently converted to numerical values to come to an overall score that integrates all these characteristics. Saaty (1980, 1994) provides a more detailed explanation of the AHP methodology.

Another way to process expert opinions is referred to as case-based distance modelling (Chen et al., 2008). The idea is to calculate a product’s distance to a predefined reference point (such as the largest volume and the highest criticality factor) for all important characteristics, leading to a classification with A, B and C categories. Even though there is a reasonable amount of modelling involved, the authors stress that the intuitive distance concept is easily understood by decision-makers.

A wide range of techniques can be found which rely on statistics. Some of these approaches classify SKUs on only one criterion whereas others incorporate a large number of characteristics. The traditional ABC approach and the related FSN/FNS
approach are examples that mostly sort products on a single characteristic. In the FSN (Fast, Slow and Non-moving) and the FNS (Fast, Normal and Slow moving) techniques, demand volume in a period is used to determine the product class. In the traditional ABC approach, the demand volume is generally multiplied by the unit price and then sorting is based on the single criterion demand value. For the ABC approaches, a dataset gathered by Reid (1987) is often used as a benchmark to test and compare techniques (Flores et al., 1992; Ramanathan, 2006; Ng, 2007; Zhou and Fan, 2007; Chen et al., 2008).

Other statistical techniques use more than one characteristic. When considering a pair of characteristics, researchers use tables, matrices or graphical techniques to illustrate their classification. For instance, D’Alessandro and Baveja (2000) plot all products on a graph with mean weekly demand volume along one axis, and the associated coefficient of variance on the other. For each quadrant in this graph, a production strategy is determined. Syntetos et al. (2005) distinguish four quadrants based on the mean inter-demand interval and the squared coefficient of variation of the demand sizes (when demand occurs). The cut-off values for their quadrants are based on a comparison of theoretical MSEs (mean squared errors) of different forecasting methods.

Another interesting technique is the decision tree. Here, the classification is performed in a stepwise fashion, one characteristic at a time. For instance, Porras and Dekker (2008) first look at the criticality of the product, then at the demand volume, and finally at price. For each combination, a specific inventory management procedure is developed. Kobbacy and Liang (1999) included statistical tests for each step in a decision tree to determine, for example, whether there is a trend (e.g. seasonality) in the demand pattern.

Finally, we found several more advanced statistical approaches for SKU classification that can easily deal with a large number of characteristics. Quite a number of authors present optimisation models to extend the basic ABC methodology by using multiple characteristics. For instance, Ramanathan (2006) considers annual demand volume, unit cost, product criticality and product lead time, and uses weighted linear programming to come to a classification. Zhou and Fan (2007) extend Ramanathan’s methodology by comparing a SKU’s most favourable and least favourable scores for the various SKU characteristics. Ng (2007) presents an alternative to Ramanathan’s optimisation model. His paper also includes a simple mechanism for calculating the classification score in a spreadsheet package rather than in specialized optimization software. Stanford and Martin (2007) integrate inventory control rules and traditional ABC classes based on demand value characteristics to optimally determine the number of classes and the borders between classes. Essentially, they model the cost performance of an inventory system with a given set of product classes and, with the product class set-up as a decision variable, they minimise the integrated cost.

4.4. Context in which the SKU classification is used

There is a wide variety of industries in which SKU classification is used. Table 3 shows examples of petrochemical and pharmaceutical industries (process manufacturing), automotive and lighting industries (discrete manufacturing), as well as high tech and low tech industries. The specific industry or context of the study was found to influence the choices made in the SKU classification. These contextual factors can be related to the product, the production process or the life cycle of the product. Güvenir and Erel (1998) provide a particular example of a specific characteristic related to the product. Their classification uses ‘stockability of the product’ since stocking explosive products
for the mining industry is not always possible. D'Alessandro and Baveja (2000) provide an example from the process industry where an operational characteristic of the production process (typical emulsion batch size) is used to distinguish between SKU classes. Applying the same approach in another context (with different batch sizes) would lead to different SKU class borders. In addition to aspects on the product or process level influencing SKU classification, the product life cycle of a product can also influence the classification. Wu et al. (2006) give an example from the high-tech industry, where the typically short product lifecycle has a major influence on the demand pattern.

5. Discussion: The relation between the factors in a SKU classification

The previous section shows the great diversity in papers and approaches to classify SKUs. In this section, we aim to extend the literature on SKU classification from the level of individual examples to the conceptual level by not only discussing the aspects presented in Section 4 but also their relationships.

The first observation that can be made from our study is that the aim of the classification, the characteristics, the technique, and the context are interrelated. Together they determine the specific SKU classification and therefore they should not be considered in isolation. The interrelatedness of the important aspects of a SKU classification is shown in a mind map (see Figure 2). Mind maps can be used for pre-analytic idea jostles (for more details see Eppler, 2006). Here, we sketch and use it as an intermediate step to explore the various relationships. In other words, it is a first step towards building a conceptual framework for SKU classification. Therefore in sections 5.1 to 5.4 we discuss each element presented in the mind map and explore possible relations between the aspects. Based on this discussion, Section 5.5 presents the conceptual framework in which we summarize and visualise how the elements relate to each other.

![Figure 2 – Mind map of SKU classification](image)

5.1. Aim of the SKU classification
Classifying SKUs is often not an aim in itself. Most studies in the area of inventory management aim to reduce the money or the space tied up in inventories and therefore use volume and product characteristics (e.g. space needs or unit cost). Often, classifications are based on the multiplication of a volume and a product characteristic (as in the ABC approach). However, can this really result in the best outcome in all
inventory management situations? We have three good reasons to believe that this is not always the case. Firstly, it is clear that inventory management for spare parts differs from that for regular products (Kennedy et al. (2002) and Boylan and Syntetos (2008) describe these differences). The focus in managing spare parts inventories is generally less on the money value or space needs of the parts but more on the consequences non-availability of parts for the customer – especially when this could stop an entire production system. For this reason, studies on the management of spare parts inventories often use customer characteristics, such as customer criticality, rather than product characteristics. Secondly, a recent contribution by Teunter et al. (2010) challenges the fundamental approach of multiplying demand volume and cost characteristics. They argue that, in order to optimise inventory, product categories should be based on the demand volume divided by the unit holding costs rather than being multiplied by the unit cost (they also take shortage costs and order quantities into account). Their rationale is that a better overall delivery performance can be achieved at a lower overall holding cost when a relatively high delivery performance (through higher inventory levels) is achieved for products with a low holding cost. Thirdly, classifying individual products ignores possible relationships between products. Shipping to a customer might only be possible or sensible if all the products on an order are available. Another example of a relation between SKUs is the similarity of products. Production planning might depend on clustering products on recipe or packaging format to reduce set-up costs. In designing a SKU classification system for inventory management, such issues should be considered.

Studies that have their aim in the area of forecasting more often consider timing characteristics than studies in other areas. This is probably related to the fact that the selection of a forecasting method is influenced by the variability of the demand. Variability not only relates to the volume (e.g. demand size variability, demand lumpiness) but also to the timing of the orders (e.g. mean inter-demand interval, intermittence - see Williams, 1984; Eaves and Kingsman, 2004; Syntetos et al., 2005).

Studies related to production strategy all use characteristics related to volume. The use of the total demand volume reflects the impact products have on the organisation. Fisher (1997) also stressed the differences in the predictability of demands for functional and innovative products as the driving force for different supply chain policies and practices.

Our synthesis of previous studies provides some evidence that studies with the same aim have characteristics which are commonly perceived to be appropriate to use. Therefore we argue that the selection of characteristics is influenced by the aim of the study. However, the fact that many studies include a certain characteristic does not necessarily mean that this characteristic should always be considered. Therefore, an interesting direction for further research is to further investigate how the aim influences the characteristics used. One particular direction could be to study the use of criticality in classifications for inventory management of spare parts. Most studies use criticality but is the use of criticality always necessary for inventory management of spare parts? Or are there contingencies when this is not the case? Exploring this dependency would be an interesting topic.

5.2. Characteristics in a SKU classification
We observe that virtually all the studies (44 out of 45 studies in our sample) used a characteristic related to volume where the level of aggregation depends on the aim of the study (next to other factors such as data availability, periodic reviews, industry norms): ranging from individual orders to aggregation on a daily, weekly, monthly or annual
basis. As noted previously, product characteristics such as unit price tend to be used in inventory management studies, and timing characteristics are mainly used in forecasting studies. Studies on spare parts often take customer characteristics (where the customer can be the production process), such as criticality, into account. Characteristics that are very specific for a setting are sometimes included. In our literature review, we found studies using ten characteristics, but we did not find clear arguments for the number of characteristics selected. Intuitively, one might expect a trade-off between the additional effort of acquiring more information on SKU characteristics and the gain in outcome quality. One avenue for further research would therefore be to investigate in which situations the use of a larger number of characteristics is beneficial. Particularly, we expect a difference in the number of characteristics used based on the level of automation in a production setting. In highly automated production settings we expect that SKU characteristics could be more easily retrieved due to lower costs of acquiring data which will result in a more refined classification based on a higher number of SKU characteristics.

5.3. SKU classification technique
The number of characteristics and the nature of the characteristics do influence technique selection. Some simple statistical approaches restrict the number of characteristics whereas, in general, the more complex statistical techniques can easily deal with a larger number of characteristics. The qualitative nature of some characteristics (e.g. criticality being defined as high, medium or low) can be used in some expert judgement approaches but cannot easily be used in mathematical approaches. In the latter, some authors explicitly exclude qualitative characteristics from their classification (e.g. Zhou and Fan, 2007; Ng, 2007), because qualitative characteristics are believed not to fit to the optimisation model. In selecting a technique for a specific situation, one has to assess the additional benefits of techniques that require a significant amount of modelling or data collection over other, simpler, techniques. In further research one could try to come up with guidelines or rules of thumb to come to this decision.

5.4. Context of the SKU classification
Section 4.4 shows that one should carefully consider whether contextual factors should be incorporated in the SKU classification. Examples are given for when the context influences which characteristics are included in the classification and for when the operationalisation of the classes is influenced by contextual settings. However, the importance of contextual factors in a number of studies does not mean that such factors are relevant in all situations. We observe a number of papers in which general demand classification techniques are presented (Chakravarty, 1981; Flores et al., 1992; Huiskonen, 2001; Ramanathan, 2006; Zhou and Fan, 2007; Ng, 2007; Chen et al., 2008). We also see a number of examples where identical ABC approaches are applied in different industries. This raises the question as to when the context, in which the SKUs are classified, is sufficiently different to warrant including contextual factors in the classification method. In other words, are some methods more general in their applicability than others? Investigating when it is desirable to include contextual factors is an interesting direction for further research. Guidance on which factors to include can possibly be found in literature that studied fundamental differences between industries (e.g. Taylor et al., 1981), within industries (e.g. Fransoo and Rutten, 1994) or provided characteristics of a specific industry (e.g. Akkerman and Van Donk, 2009). A particular direction could for instance be to study the effect of sequence-dependent set-ups in the
process industry. The set-up costs of a recipe in the process industry result in the clustering of demand for end products based on the recipe. Production intervals (e.g. cyclical plans or campaigns) of a recipe therefore influence the production interval of an end product. We therefore expect these set-up costs and the related production intervals to influence the SKU classification.

5.5. Conceptual framework

In all papers, classifying SKUs is about identifying a number of SKU classes and drawing borders between these classes. Together we call this the operationalisation of SKU classes. Next to the decision how to operationalise the SKU classes, decisions are made which characteristics to include and which technique to use. These three interrelated decisions made are labelled together as the method. Figure 3 visualises the interrelationships.

Before we construct our conceptual model, we first discuss the operationalisation of SKU classes, as the basic decisions on the number of classes and their borders are made in every classification. Here, we include possible relations to the aspects described in the previous sections.

The number of classes employed is usually between three and twelve (Stanford and Martin, 2007), but there is no guidance on how to determine the optimum. One could argue that some of the more popular techniques have three classes (e.g. ABC, FSN, VED) and in that situation the operationalisation of the classes is influenced by the technique. However, examples exist where these techniques are used with more classes (for example, Sani and Kingsman (1997) discuss an ABC application using 11 classes).

Different methods exist to define class boundaries. In the ABC approach, one often defines class borders based on percentages of products (e.g. 10% of products are A, 40% B, and 50% C). Other methods use visual inspection of data, descriptive statistics (e.g. quartiles, median) or operational characteristics (e.g. batch size). Companies with similar aims and characteristics may well set different class borders in their SKU classification. Eaves and Kingsman (2004) confirm that idea by stating “what is classed as a smooth demand pattern in military terms may well be considered intermittent in other industries”. D’Alessandro and Baveja (2000) contend that the choice of boundaries between classes may not even have any intrinsic meaning.

The number of classes (Sani and Kingsman, 1997) and the boundaries between them (Eaves and Kingsman, 2004) are essentially management decisions. However, how can
or should managers take such decisions, and what could be leading in such decisions? These questions and the many examples of ABC applications with different numbers of SKU classes, suggests an interesting direction for further research. Namely, whether it is only organisational or managerial considerations that influence the number of SKU classes and class boundaries, or whether there is some logic which explains how companies decide on the number of SKU classes and class boundaries. We would expect a trade-off between performance and complexity. While the best performance could theoretically be expected to be achieved by creating different classes for each product this will come at the expense of complexity. On the other hand, using only one class will result in a relatively poor performance. Again, different approaches have been followed and presented in the literature, but little foundation is offered for individual choices. Further exploration of the number of classes used to balance between performance and complexity, thus seems another challenging area for further research.

At the start of the discussion section, four interrelated areas were mentioned: the aim of the classification, the technique, the characteristics, and the context. Together these areas influence the central classification decision of how the SKU classes are operationalised. Based on the previous sub-sections, we feel confident to further refine the nature of the relationships of the factors as follows:

1. The aim influences the characteristics chosen (see Sections 5.1 and 5.2).
2. The context influences the characteristics chosen (see Sections 5.2 and 5.4).
3. The characteristics chosen influence the technique (see Section 5.3).
4. The technique chosen influences the operationalisation of the classes (see Section 5.5).
5. The context influences the operationalisation of the classes (see sections 5.4 and 5.5).

These five relations are graphically represented in Figure 4.

Two remarks need to be made regarding Figure 4. Firstly, the existence of a relationship between two areas does not mean that it influences the classification in all situations. For example, similar ABC classifications are used in different contextual settings. However, the relation indicates that the literature provides examples of studies in which these relationships exist and therefore should be considered. Secondly, the method of the SKU classification in Figure 4 might be influenced by the strategic aim of the company as well. A company that aims for high service levels might include more
characteristics, select more labour/capital intensive classification methods or use other class borders than a company aiming at low costs. However, in the absence of guidance on this topic we would argue that the above conceptual framework is a good starting point for further research as well as for further specifying the various relationships.

6. Conclusion and further research

This paper provides a systematic analysis of the literature on SKU classification resulting in an overview of aims, techniques and characteristics used to classify SKUs in various contexts. By synthesising and structuring the existing studies in this field, the lack of guidance on how to classify SKUs became apparent leading to detecting several important unanswered questions.

In addition to reviewing previous work on SKU classification, this study contributes to the literature by (i) distinguishing four main characteristics used for SKU classification (volume, timing, product and customer), (ii) discussing the main factors influencing SKU classification (Figure 2), (iii) revealing three key decisions that are made in each SKU classification method (Figure 3) and (iv) proposing a conceptual framework for SKU classification (Figure 4). Managers in practice can benefit from our findings as they provide an overview of studies conducted in a variety of industries. Managers in related industries can learn from these experiences. Furthermore, the paper highlights which decisions need to be taken to come to an appropriate SKU classification and as such offer practical guidance.

SKU classification is a widely applied concept in production and operations management which has, so far, received mainly context-specific and fragmented attention in the literature. As a consequence it is therefore difficult to assess and compare the performance of different approaches. The conceptual framework and the discussion in this paper contribute to the development of production and operations management theory on SKU classification by synthesizing previous work. This study provides the groundwork for theory building with respect to SKU classification. Related to the framework a number of directions for further research can be indentified.

One of the main aspects to study is the dependency on context (e.g. Whetten, 1989). What makes a specific industry or company sufficiently different from others to require the inclusion of specific contextual factors in the SKU classification method? Our study has shown some examples where the classification characteristics are influenced by specific industry characteristics. To be able to assess the performance of classification methods, context-specific factors need to be taken into account. Guidance for how to include such specific factors can possibly be found in literature that studied fundamental differences between industries (e.g. Taylor et al., 1981), within industries (e.g. Fransoo and Rutten, 1994) or provided examples from a specific industry (e.g. Akkerman and Van Donk, 2009). A particular direction could be to study the inclusion of set-up costs and the related cyclical production plans in process industries as the production interval on a recipe level influences the production interval on a SKU level. Additionally, a broader survey or case study research over a range of companies might reveal which and to what extent contextual factors should be taken into account when classifying SKUs, and if and how performance is influenced.

Another direction for further research is to identify how the aim of the study influences the selected characteristics. This study provides a number of examples of commonly used characteristics in studies with a common aim (e.g. the use of criticality for inventory management of spare parts). But is the use of criticality always necessary for inventory management of spare parts? Or are there contingencies when this is not the case? Exploring the dependency of the chosen characteristics on the aim of the study...
is an interesting direction for research. A possibility would be to conduct a review or a multiple case study on this topic. Evaluating the performance of a number of classifications with and without certain characteristics might provide such insight.

Some more specific directions for further research related to the classification method can also be identified. We have observed that recent contributions have applied new techniques such as distance modelling and neural networks in developing SKU classes. More studies are needed to clarify when such techniques, which require a reasonable amount of modelling, are preferable over other, simpler techniques. Comparing the performance of a range of classification techniques and the efforts needed to apply these techniques on a number of datasets might provide such insights. Similar directions for future research would be to evaluate the decisions on the number of characteristics, the number of classes and class borders. We expect the level of automation to influence the data collection efforts and therefore the decision on the number of characteristics. Further, we expect the use of the classification technique to influence the number of classes. Having a large number of classes could be useful in a highly automated production setting where it might be difficult to handle in a low automated production setting due to human limitations.

This paper is a first step to unravel whether some deeper logic can be found to explain how the different SKU classification decisions are made or should be made. Ultimately the aim for further research would be to construct a decision framework on how to determine an appropriate SKU classification.

Acknowledgements
The authors would like to thank the anonymous reviewers for their comments which greatly helped us to improve the quality of the paper.
References


Mukhopadhyay, S. K., Pathak, K., & Guddu, K. (2003), "Development of decision support system for stock control at area level in mines", *Institution of Engineers (India) Journal-Mining*, vol. 84, no. 1, pp. 11-16.


Portougal, V. (2002), "Demand forecast for a catalog retailing company", *Production and Inventory Management Journal*, vol. 43, no. 1-2, pp. 29-34.


