Statistical Process Control techniques to monitor quality determinants in digital Voice-of-Customer

Barravecchia, F., Mastrogiacomo, L., Tavani, L. and Franceschini, F.

Politecnico di Torino, DIGEP (Department of Management and Production Engineering), Torino, Italy

STRUCTURED ABSTRACT

Purpose - Digital Voice-of-Customer (digital VoC) primarily consists of textual feedback posted by users of products or services on the web. Digital VoC may represent a valuable source of information for quality management, and its promising potential is also receiving a lot of attention in the new Quality 4.0 framework. However, manufacturers and service providers still lack operative approaches to fully exploit the value of digital VoC. This study tries to answer the following research question: *How Statistical Process Control (SPC) techniques can be used to monitor digital VoC over time*?

Design/methodology/approach - This article explores the applicability of SPC to support digital VoC analysis. Two types of control charts, for variables and attributes, were applied to a real case study concerning a product-service system (car-sharing).

Findings - SPC tools may represent an interesting alternative to traditional quality tracking approaches to analyze the evolution of quality determinants over time.

Originality/value – This study shows how Artificial intelligence algorithms and SPC tools may support product and service designers in implementing continuous improvement actions by analyzing digital VoC over time.

Keywords: Statistical Process Control, Control charts, Voice-of-Customer, Topic modelling, Quality 4.0

Paper type: Research paper

1. INTRODUCTION

The survival of any business in an increasingly competitive scenario is closely tied to having satisfied and loyal customers (Psarommatis et al., 2020). In this view, organisations need to monitor the performance of their products/services over time. To effectively manage quality, it is essential to understand which are the key influencing factors (Franceschini et al., 2009), also known as quality determinants, to allow organisations to measure, control, and improve the way customers perceive a product/service (Mukherjee, 2019).

Organisations employ quality tracking techniques to monitor the evolution of quality by directly auditing consumers (Kamsu-Foguem et al., 2013; Xu et al., 2021) and by analysing the so-called "Voice-of-Customer" (VoC) (Jach et al., 2021). Traditional quality tracking techniques collect VoC from interviews, questionnaires, or focus groups. The main limitations of these activities are that (i) they are expensive in terms of required time and resources, (ii) they can only reach a limited sample of customers, and (iii) they can be quite intrusive for the interviewed customers.

The digitalisation of consumption has enabled customers to release massive quantities of VoC on the web, i.e., digital VoC. Such data, also known as User-Generated Contents (UGC), primarily consists of customer reviews, i.e. unstructured textual records describing the customer's experience with a specific product or service (Elg et al., 2021; Mastrogiacomo et al., 2021). Many studies have already proven that digital VoC analysis can be of great value for quality management and design (Barravecchia et al., 2021; Barravecchia et al., 2020a, 2020b; Mastrogiacomo et al., 2020). One of the most effective techniques for analysing digital VoC is topic modelling, i.e. artificial intelligence algorithms that can extract latent topics discussed within collections of unstructured text documents. When applied to digital VoC collections, topic modelling algorithms allow the extraction of latent quality determinants of the product/service analysed (Mastrogiacomo et al., 2020). The output of topic modelling algorithms is rich in information and it can be of great value in the management and monitoring of quality.

To date, digital VoC analytics have not yet been used to explore the evolution of quality determinants. The current challenge is then to understand how to leverage on digital VoC analysis to monitor quality determinants over time. In this view, the objective of this study is to provide a preliminary analysis of how Statistical Process Control (SPC) may support digital VoC analysis. In detail, the study tries to answer the following research question: *How can Statistical Process Control (SPC) techniques be used to monitor digital VoC over time?*

The remainder of this paper is organised into three sections. Section 2 provides the conceptual background and the significant research that has been conducted on quality tracking and digital VoC. Section 3 describes two preliminary applications of SPC techniques for the analysis of digital VoC.

Finally, Section 4 summarises the main contributions of the work, its limitations, and possible future research directions.

2. CONCEPTUAL BACKGROUND

2.1. Customer Satisfaction

According to the international standard ISO/FDIS 10004 (2018), customer satisfaction is the customer perception of the degree to which expectations are met by a product (or service). There are multiple interpretations of the concept of Customer Satisfaction (CS) in the literature; however, all definitions share three common elements:

- consumer satisfaction is an emotional or cognitive response;
- the response relates to a particular focus (expectations, product, consumer experience, etc.);
- the response occurs at a particular time (after consumption, after choice, etc.).

CS is the resulting aggregate assessment of all latent dimensions characterising product or service quality (Matzler & Sauerwein, 2002). Tracking latent dimensions over time provides an insight into how and why customer satisfaction is changing.

2.2. Voice-of-Customer

To analyse customer satisfaction, it is crucial to take the "Voice-of-Customer" (VoC) into account (Aguwa et al., 2012, Lysenko-Ryba et al., 2022). By actively investigating VoC it is possible to anticipate customer future needs and better design new products (Aguwa et al., 2012). Product development can be positively influenced by data collected in post-sales, as evidenced by the systematic use of this data in many industries such as, for example, in the automotive sector (Szwejczewski et al., 2015). Through VoC analysis, it is possible to identify (Wang & Tseng, 2011):

- *Explicit requirements*: requirements that the customer states explicitly (for example: a customer explicitly states that wants a red car).
- *Implicit requirements*: requirements that the customer does not express explicitly, but wants or needs (for example, someone who buys a washing machine expects it to be able to wash clothes, but will not explicitly express this need. It is considered an intrinsic characteristic of the washing machine).
- *Latent requirements*: requirements that cannot be easily expressed by the customer. Meeting these requirements is critical to the success or failure of a product.

A variety of techniques are available to intercept and analyse VoC. Table 1 provides a summary of traditional ones.

Table 1 - Traditional techniques for VoC Analysis (Freeman & Radziwill, 2018).

Technique	Description
Survey	Surveys are a popular method for collecting easily quantifiable feedback. They use predefined questions in a variety of formats including fillable text-boxes or multiple choice. Researchers can conduct surveys easily in person, over the phone, through web forms, or video. Surveys are useful for assessing and monitoring customer preferences and satisfaction, as well as for evaluating the impact of changes to products or services.
Benchmarking	Benchmarking is a practice in which organisations study how other organisations satisfy their customers' needs. It is a means to study best practices and learn how to pinpoint weaknesses in processes and design workflows.
Gemba Visits	Gemba is a technique in which the researcher goes to the workplace to get direct information about what customers want and need. It can be an excellent technique for observing workers directly in their environment, which is particularly useful when customers and workers don't feel they have the freedom to complain openly
Focus group	Focus groups allow researchers to spend time with customers to solicit answers to specific questions or engage in wide-ranging brainstorming sessions. These events can be conducted in person or with collaborative technology
Metodo Delphi	The Delphi Method is an interviewing technique in which researchers present subject matter experts with multiple rounds of questionnaires. Respondents then deliberate during each round until they narrow down their responses and reach a consensus.
Warranty data	Collecting data during the servicing of warranty claims can provide valuable information about product failures and customer dissatisfaction, as well as how customers think the products fail to live up to their promised functionality.

2.3. Quality tracking

Companies implement many strategies to track the evolution of a product/service quality over time, identify anomalies or criticalities, and uncover potential areas for improvement. Three practical approaches are typically considered (Bandaru et al., 2015; Hallencreutz & Parmler, 2021):

- *post-purchase evaluation*: this approach is performed by asking a customer to evaluate a purchased product or service after its delivery or use (Kumar & Anjaly, 2017);
- *periodic quality survey*: this approach, which is generally based on the administration of questionnaires or structured interviews, provides snapshots of customer perceptions (Izogo & Ogba, 2015; Su & Hwang, 2020);
- *continuous monitoring of quality*: this approach involves the ongoing monitoring of quality characteristics over time (Chen et al., 2015; Gregorio & Cronemyr, 2011; McColl-Kennedy & Schneider, 2000).

2.4. Digital VoC analysis

Information released by users on the web comes in different forms (text, audio, photos, videos). They are primarily published on social networks, discussion forums, blogs, review aggregators, and e-commerce platforms. By their influence on determining the demand and sales of a product, digital VoC can be considered a new form of word-of-mouth, so that consumers often perceive it as more reliable than traditional promotional practices (Wang, 2015).

Digital VoC is mainly composed of unstructured textual UGC. Users of a product/service freely describe their experience without following standard patterns and without providing readily processable structured data. In addition, the digital VoC is often composed of a large number of records shared by the customers. Given these properties, digital tools are needed to automatically extract information from digital VoC. Text mining approaches, i.e., techniques for the automatic analysis of unstructured textual documents, are widely implemented to analyse digital VoC (Berry & Kogan, 2010; Carnerud, 2020). These techniques allow hidden relationships to be found within textual data (Berry & Kogan, 2010). One of the most widely used text mining techniques for analysing unstructured textual information is topic modelling. The term topic modelling refers to the family of statistical methodologies that allow the latent topics discussed within a collection of documents to be extracted (Blei, 2012). In other words, these algorithms do not require human coding or a preliminary classification of data, as they can "read" a collection of documents and automatically extract the most discussed topics. Each topic is distinguished by a set of keywords. Figure 1 represents the conceptual scheme of a topic modelling algorithm. Given an extensive collection of digital VoC records $\{1,2,3,\dots,J\}$, topic modelling algorithms deal with the following aspects (Blei et al., 2003):

- identifying a set of topics that describe the text corpus (i.e., the collection of digital VoC);
- associating a set of keywords to each topic ($topical\ content$: $TC_{w,d}$),
  where:
  - $w \in \{1,\dots,W\}$ are the keywords of the vocabulary related to the digital VoC collection;
  - W is the total number of words contained in the digital VoC vocabulary.
  - $d \in \{1,\dots,D\}$ are the topics identified by the topic modelling algorithm
  - D is the total number of identified topics
- defining a specific mixture of these topics for each digital record (*topical prevalence*: $TP_{j,d}$)
  where:
  - $j \in \{1,\dots,J\}$ are the analysed digital VoC records
  - J is the total number of analysed digital VoC records
  - $\sum_{d=1}^{D} TP_{j,d} = 1 \quad \forall j$

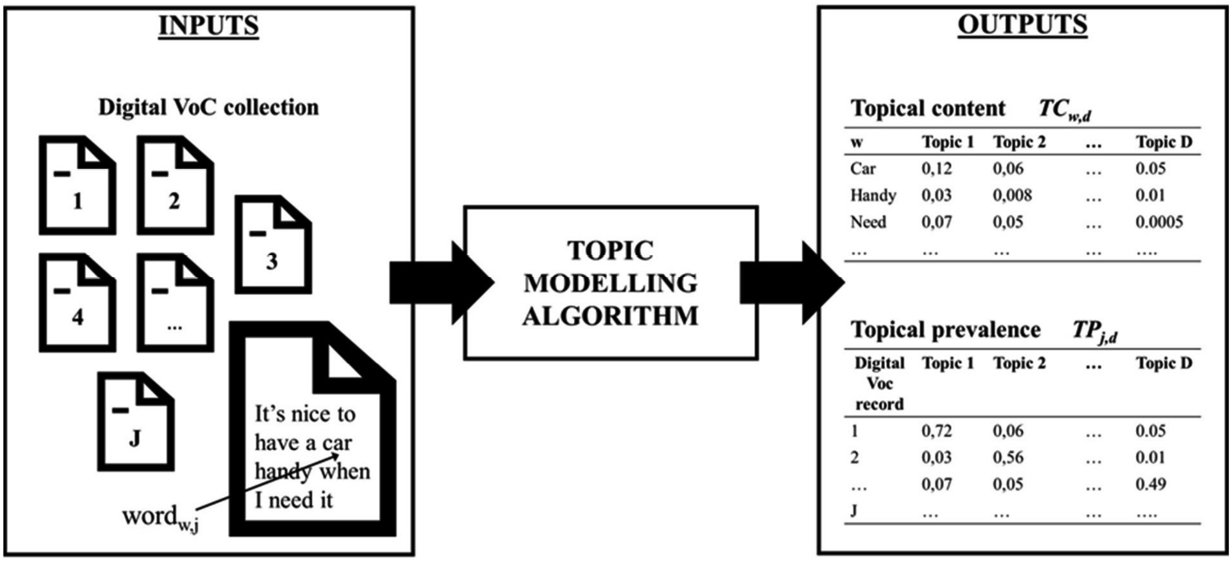

Figure 1 - Graphical representation of the functioning of topic modelling algorithms. The input consists of *J* textual documents (digital VoC records). The output consists of: $TC_{w,d}$ (Topical content matrix), i.e., the weights associated with each *w*-th keyword that characterises the *d*-th topic; $TP_{j,d}$ (Topical prevalence matrix), i.e., multinomial probability distribution that indicates the proportion of the *d*-th topic discussed within the *j*-th document.

2.5. Statistical Process Control and digital VoC Analysis

Collecting and analysing digital VoC records makes it possible to audit customer judgements more reliably than traditional techniques (Sweeney et al., 2008). The analysis of digital VoC can help organisations to overcome some of the limitations of conventional VoC collection techniques.

Topic modelling algorithms are able to extract the latent quality determinants (topics) of products or services, i.e., the feature that can positively or negatively influence the perceived quality (Barravecchia et al., 2020; Mastrogiacomo et al., 2021). It can be assumed that if a topic is discussed, then it is important for the customer, and therefore is critical to the perceived quality of the object under study.

The analysis of the outputs of topic modelling algorithms represents a possible source of information for assessing how consumer perceived quality varies over time. The variability in the "discussion level" (topical prevalence) of quality determinants may be investigated through tools typically employed in Statistical Process Control (SPC).

In the literature, there are few studies aimed at investigating how and to what extent the level of discussion of quality determinants varies over time. One of the first studies is dated 2008 (Lo, 2008). Lo applies a Support Vector Machine (SVM) algorithm to identify UGCs containing complaints by users of a website. After a preliminary classification of the complaints, their proportion is monitored through the use of control charts for attributes.

Ashton and Evangelopoulos (2012, 2014, 2015) proposed a model capable of exploring the evolution of different topics over time. More specifically, the authors identified the topics discussed in digital

VoC related to a large retail company using the Latent Semantic Analysis (LSA) algorithm and proposed an approach to monitor them over time. To keep track of changes in the proportions of the different quality factors, the authors used Exponential Moving Average Control Charts. Each control chart is specific to a single topic.

More recently, Liang and Wang (2019) proposed a monitoring methodology combining the analysis of the topics discussed in the digital VoC and the related sentiment expressed by customers. The proposed approach also tracks the distribution of customer sentiments, distinguishing positive from negative sentiments.

3. DIGITAL VOC CONTROL CHARTS

The analysis of the latent quality determinants in the digital VoC can identify anomalous behaviours. In general, two types of variability can be recognised:

- *natural variability*, which indicates the cumulative effect of a large number of small and uncontrollable causes;
- s*ystematic variability*, which indicates distortions in the process.

The identification of the *topical prevalence* of a latent determinant affected by a systematic variability can allow the detection of "out of control" situations of the analysed product/service. A control chart can signal the presence of a new source of variation, which can indicate an alteration in customer satisfaction due to specific assignable causes.

The application of control charts for the analysis of digital VoC needs to address the following aspects:

- *Which type of control chart to adopt*? Different types of variables can be considered, e.g., continuous variables (level of discussion of a quality determinant) or discrete variables (most discussed quality determinant in a digital VoC record).
- *How to manage quality determinants with trends*? It has been empirically observed that the quality determinants can exhibit different natural temporal trends in the topical prevalence: increasing, stationarity or decreasing trends (Barravecchia et al., 2020; Mastrogiacomo et al., 2021). The reasons behind these patterns may be different, including for example the evolution of customer needs or the learning of particular product/service aspects (Franceschini, 2002). Quality determinants characterised by specific temporal trends might require “focused” control charts capable of "following" these trends.
- *How to identify the sampling period*? The production rate of digital VoC by users may be subject to a variation over time. The choice of the sampling period (i.e., how often digital VoC records are grouped into a sample) can influence monitoring results. For example, a

long sampling period may mitigate the dynamicity of observations, but on the other hand, it may generate delays in the detection of potential out-of-controls.

- *How to manage the emergence of new quality determinants*? The number and type of quality determinants may change over time. The impact of the emergence or disappearance of some quality determinants needs to be investigated in detail.

In order to preliminary explore these issues, the following sub-sections propose an application of two different types of control charts to monitor a digital VoC database relating to a Product-Service System (car-sharing). The proposed applications concern the construction of the control charts.

3.1. Case study

The case study concerns the analysis of digital VoC regarding car-sharing services (Mastrogiacomo et al., 2021). The investigation is based on the application of the Structural Topic Model (STM) algorithm, which allows to include the metadata associated with the digital VoC for the definition of the topic model. The algorithm was implemented on the open-source R software using the STM package (M. E. Roberts et al., 2019). Analysed data are records retrieved in December 2019 from different databases: Yelp, Google, Trustpilot, Facebook, and Playstore. The time window of the analysis (i.e. the interval where the analysed samples are collected) is from January 2006 to December 2019 . Only English-language reviews were selected, with a total of almost 17,000 reviews from 22 car-sharing providers (Car2go, DriveNow, Maven, Zipcar, Goget), operating in 3 different countries (US, Canada, and UK).

Table 2 reports the keywords characterising the identified quality determinants and the corresponding assigned labels.

The comprehensive database of digital VoC was analysed in to identify quality determinants relevant to all the analysed car-sharing companies. Conversely, control charts developed in the following sections are related to a specific car-sharing provider.

Table 2 - Top keywords and related semantic labels of the quality determinants of car-sharing.

Quality determinant (d)	Keywords (w)	Quality determinant label
1	help, phone, call, person, office, answer, number	Customer service (physical office)
2	damage, report, accident, fault, member, enterprise, claim	Accident & damages management
3	sign, process, website, license, drive, driver, registration	Registration process
4	charge, fee, late, return, time, pay, hour	Charges & fees
5	park, lot, spot, find, ticket, street, space	Parking areas
6	app, work, update, book, map, reserve, time	App reliability
7	trip, end, time, make, actual, take, system	End trip issues
8	gas, dirty, rent, clean, tank, card, tire	Car condition
9	need, convenient, quick, recommend, awesome, clean, perfect	Convenience
10	hour, price, rate, cost, expense, mile, cheaper	Use rates
11	minute, reservation, walk, wait, home, time, away	Car proximity
12	car, available, location, vehicle, area, change, time	Car availability
13	use, time, now, far, user, review, star	Efficacy
14	city, year, insurance, member, gas, need, month	Sharing benefits
15	service, custom, issue, company, terrible, problem, experience	Customer service responsiveness
16	way, drive, little, take, get, town, bus	Intermodal transportation
17	time, start, location, turn, lock, pick, key	Car start-up issues
18	call, member, cancel, ask, rep, refund, manage	Customer service courtesy
19	account, card, email, credit, month, day, membership	Billing and membership
20	reservation, plan, time, need, book, cancel, advance	Car reservation

3.2. $\bar{x} - s$ control charts for car-sharing quality determinants

In this section, we present an application of $\bar{x} - s$ control charts for car-sharing quality determinants. The *topical prevalence* ($TP_{j,d}$) of quality determinants (topics) is the variables considered in the analysis. Two separate control charts are provided for each quality determinant. Due to the variability of the number of records in each sample, $\bar{x} - s$ control charts with variable sample size were considered.

In order to track the evolution of quality determinants over time, we introduce the concept of Interval Mean Topical Prevalence ($IMTP_{d,t}$) which represents the average topical prevalence in digital VoC for the *d*-th quality determinants related to the *t*-th sampling period:

$$IMTP_{d,t} = \frac{\sum_{j}^{R_t} TP_{j,d}}{|R_t|} \quad (1)$$

where R_t is the set of digital VoC records collected in the *t*-th sampling period, $|R_t|$ is the cardinality of the R_t set (sample size of VoC records).

For each *t*-th sampling period, the sum of the $IMTP_{d,t}$ related to all the identified quality determinants is equal to 1:

$$\sum_{d=1}^{D} IMTP_{d,t} = 1 \qquad \forall\, t \in (1, \dots, T) \quad (2)$$

where *D* is the number of identified topics, and *T* is the total number of sampling period considered.

Table 3 shows an example of the calculation of $IMTP_{d,t}$ for three quality determinants $d = A, B, C$, and three sampling period (January, February and March 2021).

Table 3 - Example of the calculation of $IMTP_{d,t}$ for 3 time periods (January, February and March 2021) and three quality determinants (A,B,C)

Digital VoC record (j)	**Date**	**Sampling period (t)**	**Topical Prevalence ($TP_{j,d}$)**			**Interval Mean Topical Prevalence ($IMTP_{d,t}$)**		
			Quality determinant A (d=1)	Quality determinant B (d=2)	Quality determinant C (d=3)	Quality determinant A (d=1)	Quality determinant B (d=2)	Quality determinant C (d=3)
1	January 2021	1	0.8	0.15	0.05			
2	January 2021	1	0.1	0.7	0.2	$IMTP_{1,1} = \frac{0.8+0.1+0.8}{3} =$ 0.57	$IMTP_{2,1} =$ 0.33	$IMTP_{3,1} =$ 0.1
3	January 2021	1	0.8	0.15	0.05			
4	February 2021	2	0.25	0.7	0.05			
5	February 2021	2	0.45	0.15	0.4	$IMTP_{1,2} = \frac{0.25+0.45+0.35}{3} =$ 0.35	$IMTP_{2,2} =$ 0.32	$IMTP_{3,3} =$ 0.33
6	February 2021	2	0.35	0.1	0.55			
7	March 2021	2	0.15	0.65	0.2			
8	March 2021	2	0.2	0.1	0.7	$IMTP_{1,3} = \frac{0.15+0.2+0.1}{3} =$ 0.15	$IMTP_{2,3} =$ 0.35	$IMTP_{3,3} =$ 0.5
9	March 2021	2	0.1	0.3	0.6			

At the *t*-th sampling period, for the *d*-th quality determinant, the standard deviation of topical prevalence can be calculated as (Montgomery, 2020):

$$s_{d,t}^2 = \frac{\sum_j^{R_t}\left(TP_{j,d} - IMTP_{d,t}\right)^2}{|R_t| - 1} \tag{3}$$

With these assumptions, the central line (CL) and the control limits (UCL and LCL) for the s control charts are respectively (Montgomery, 2020):

$$CL_d = \bar{s}_d = \left[\frac{\sum_{t=1}^{T}(|R_t| - 1)s_{d,t}^2}{\sum_{t=1}^{T}|R_t| - T}\right]^{\frac{1}{2}} \tag{4}$$

$$UCL_{d,t} = B_4\,\bar{s}_d \tag{5}$$

$$LCL_{d,t} = B_3\bar{s}_d \tag{6}$$

Where T is the total number of considered time periods, and B_3 and B_4 are constants tabulated for various values of sample size ($|R_t|$) (Montgomery, 2020).

The central line and variable control limits for the corresponding $\bar{x}$ control chart can be calculated as follows:

$$CL_d = \overline{\overline{x_d}} = \frac{\sum_{t=1}^{T}\sum_{j}^{R_t} TP_{j,d}}{\sum_{t=1}^{T}|R_t|} \tag{7}$$

$$\mathrm{UCL}_{d,t} = \overline{\overline{x_d}} + A_3\bar{s}_d \tag{8}$$

$$LCL_{d,t} = \overline{\overline{x_d}} - A_3\bar{s}_d \tag{9}$$

A_3 is a constant depending on the sample size ($|R_t|$) (Montgomery, 2020).

As an example, Table 4 and Table 5 show respectively the values obtained for the construction of the $\bar{x}$ and s control charts, for a six-month sampling period, related to the "Charges and fees" ($d = 4$) quality determinant.

Table 4 - Values obtained for the construction of the *s* control chart for the "Charges and fees" (d=4) quality determinant. Sampling period: 1 semester. S1 and S2 indicate respectively the first and the second semester of each year.

Sampling period (t)	$\|R_t\|$	$s^2_{4,t}$	CL_4	**B3**	**B4**	$LCL_{4,t}$	$UCL_{4,t}$
S1 2006	26	0,0352	0,0771	0,5715	1,4285	0,0441	0,1101
S2 2006	28	0,0439	0,0771	0,5880	1,4120	0,0453	0,1089
S1 2007	46	0,0990	0,0771	0,6820	1,3180	0,0526	0,1016
S2 2007	54	0,0766	0,0771	0,7072	1,2928	0,0545	0,0997
S1 2008	65	0,0583	0,0771	0,7338	1,2662	0,0566	0,0976
S2 2008	51	0,0692	0,0771	0,6985	1,3015	0,0539	0,1003
S1 2009	59	0,0584	0,0771	0,7203	1,2797	0,0555	0,0987
S2 2009	80	0,0518	0,0771	0,7606	1,2394	0,0586	0,0956
S1 2010	66	0,1058	0,0771	0,7359	1,2641	0,0567	0,0975
S2 2010	80	0,0913	0,0771	0,7606	1,2394	0,0586	0,0956
S1 2011	98	0,0894	0,0771	0,7841	1,2159	0,0604	0,0937
S2 2011	79	0,0987	0,0771	0,7590	1,2410	0,0585	0,0957
S1 2012	86	0,0667	0,0771	0,7692	1,2308	0,0593	0,0949
S2 2012	68	0,0830	0,0771	0,7399	1,2601	0,0570	0,0972
S1 2013	72	0,0713	0,0771	0,7474	1,2526	0,0576	0,0966
S2 2013	80	0,0630	0,0771	0,7606	1,2394	0,0586	0,0956
S1 2014	87	0,0722	0,0771	0,7706	1,2294	0,0594	0,0948
S2 2014	94	0,0803	0,0771	0,7794	1,2206	0,0601	0,0941
S1 2015	100	0,0871	0,0771	0,7863	1,2137	0,0606	0,0936
S2 2015	102	0,0839	0,0771	0,7884	1,2116	0,0608	0,0934
S1 2016	102	0,0698	0,0771	0,7884	1,2116	0,0608	0,0934
S2 2016	136	0,0687	0,0771	0,8171	1,1829	0,0630	0,0912
S1 2017	110	0,0716	0,0771	0,7963	1,2037	0,0614	0,0928
S2 2017	82	0,0744	0,0771	0,7636	1,2364	0,0589	0,0953
S1 2018	101	0,0688	0,0771	0,7873	1,2127	0,0607	0,0935
S2 2018	104	0,0776	0,0771	0,7905	1,2095	0,0609	0,0932
S1 2019	60	0,0634	0,0771	0,7227	1,2773	0,0557	0,0985
S2 2019	54	0,1076	0,0771	0,7072	1,2928	0,0545	0,0997

Table 5 - Values obtained for the construction of the $\overline{x}$ control chart, for the "Charges and fees" (d=4) quality determinant. Sampling period: 1 semester. S1 and S2 indicate respectively the first and the second semester of each year.

Sampling period (t)	$\|R_t\|$	$IMTP_{4,t}$	CL_4	A_3	$LCL_{4,t}$	$UCL_{4,t}$
S1 2006	26	0,0374	0,0759	0,5942	0,0301	0,1218
S2 2006	28	0,0417	0,0759	0,5722	0,0318	0,1201
S1 2007	46	0,0592	0,0759	0,4448	0,0417	0,1102
S2 2007	54	0,0600	0,0759	0,4102	0,0443	0,1076
S1 2008	65	0,0628	0,0759	0,3736	0,0471	0,1047
S2 2008	51	0,0556	0,0759	0,4222	0,0434	0,1085
S1 2009	59	0,0581	0,0759	0,3923	0,0457	0,1062
S2 2009	80	0,0612	0,0759	0,3365	0,0500	0,1019
S1 2010	66	0,0915	0,0759	0,3707	0,0474	0,1045
S2 2010	80	0,0833	0,0759	0,3365	0,0500	0,1019
S1 2011	98	0,0879	0,0759	0,3038	0,0525	0,0994
S2 2011	79	0,0818	0,0759	0,3386	0,0498	0,1020
S1 2012	86	0,0702	0,0759	0,3244	0,0509	0,1010
S2 2012	68	0,0713	0,0759	0,3652	0,0478	0,1041
S1 2013	72	0,0764	0,0759	0,3548	0,0486	0,1033
S2 2013	80	0,0778	0,0759	0,3365	0,0500	0,1019
S1 2014	87	0,0707	0,0759	0,3226	0,0511	0,1008
S2 2014	94	0,0965	0,0759	0,3103	0,0520	0,0999
S1 2015	100	0,0841	0,0759	0,3008	0,0528	0,0991
S2 2015	102	0,0807	0,0759	0,2978	0,0530	0,0989
S1 2016	102	0,0751	0,0759	0,2978	0,0530	0,0989
S2 2016	136	0,0758	0,0759	0,2577	0,0561	0,0958
S1 2017	110	0,0794	0,0759	0,2867	0,0538	0,0980
S2 2017	82	0,0704	0,0759	0,3323	0,0503	0,1016
S1 2018	101	0,0840	0,0759	0,2993	0,0529	0,0990
S2 2018	104	0,0829	0,0759	0,2949	0,0532	0,0987
S1 2019	60	0,0682	0,0759	0,3889	0,0460	0,1059
S2 2019	54	0,1023	0,0759	0,4102	0,0443	0,1076

Figure 2 shows the graphical representation of the $\bar{x}$ control chart for the "Charges and fees" (d=4) quality determinant.

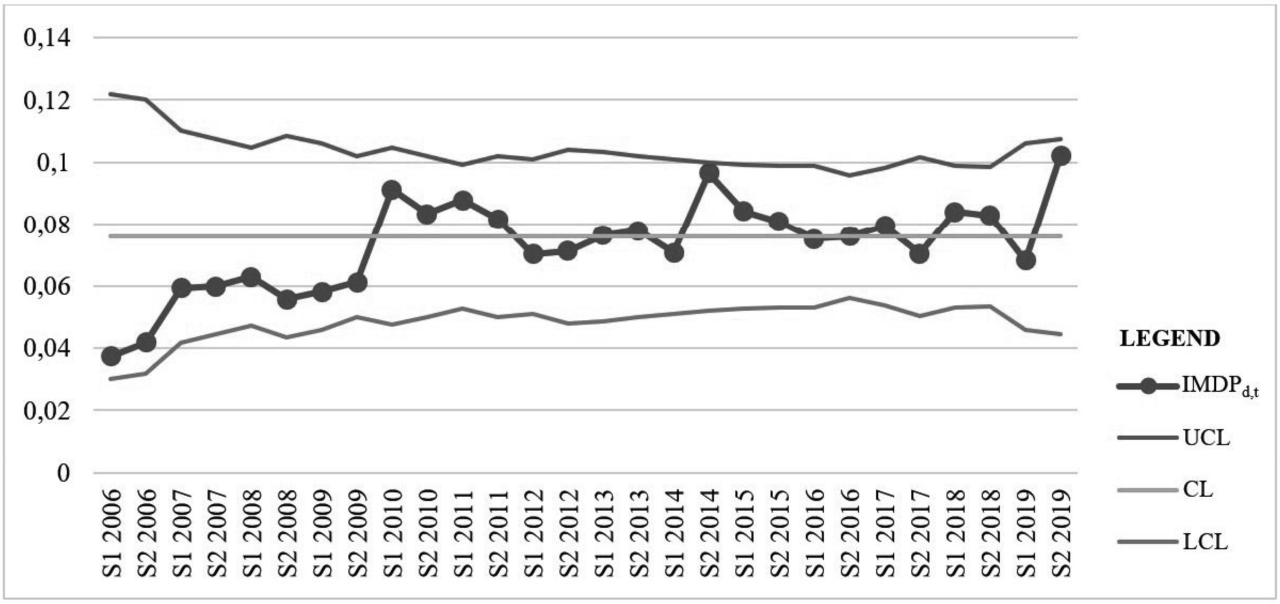

Figure 2 - $\bar{x}$-control charts ($IMTP_{4,t}$) for the "Charges and fees" (d=4) quality determinant. Sampling period: 1 semester. S1 and S2 indicate respectively the first and the second semester of each year.

We observe that none of the points falls outside the control lines. However, according to the Western Electric rules, an out-of-control is observed, due to a high number of consecutive points (from S1 2006 to S2 2009), falling on the same side of the central line (Montgomery, 2020).

Figure 3 shows the s-control chart for the "Charges and fees" (d=4) quality determinant.

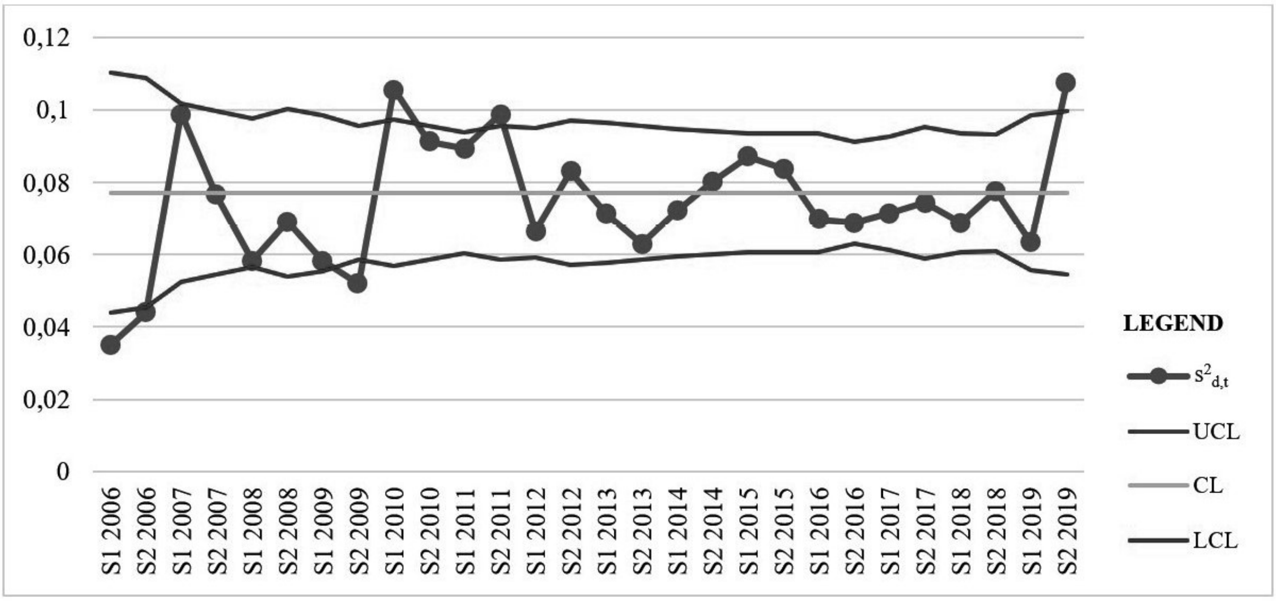

Figure 3 – $\boldsymbol{s}$-control charts for the "Charges and fees" (d=4) quality determinant. Sampling rate: 1 semester. S1 and S2 indicate respectively the first and the second semester of each year.

In this case, several observations fall outside the control limits, highlighting an anomalous variability in the topical prevalence of the "Charges and fees" (d=4) quality determinants. The causes of out-of-control must be analyzed in detail to avoid their recurrence. As an example, let us consider the last point of the control charts, the increase in s^2 is caused by failures in the application of charges and fees. A detailed analysis of the process may allow to identify specific causes.

3.3. p-control charts for car-sharing quality determinants

In this section, we provide a second way to analyze topical prevalence. As an example, let's try to monitor the fraction of "winnings" of a particular quality determinant over the others. By a "winning" quality determinant is meant a quality determinant whose *topical prevalence*, within a record (customer review), prevails over the others. We may consider the "winning" quality determinant as the most representative of the digital VoC record. Each review can be associated with only one "winning" determinant. The topical prevalence of the winning quality determinant can be set to the value 1 through a binary transformation. Conversely, the topical prevalence of the remaining $(D-1)$ quality determinants can be set to the value 0.

For each j-th digital VoC record, the "winning" quality determinant is the one that shows the highest topical prevalence:

$$V_{j,d} = \begin{cases} 1 & if \;\; TP_{j,d} = Max_j(TP_{j,d}) \\ 0 & if \;\; TP_{j,d} \neq Max_j(TP_{j,d}) \end{cases} \tag{10}$$

Table 6 shows an example of this binary transformation.

Table 6 - Example of binary transformation for the development of a p-control chart for quality determinants

Quality determinants	A	B	C	D
$TP_{j,d}$	0.8	0.10	0.07	0.03
$V_{j,d}$	1	0	0	0

As before, the sum of the topical prevalences associated with a j-th digital VoC record is equal to 1:

$$\sum_{d=1}^{D} V_{j,d} = 1 \qquad \forall\, j \in (1, \dots, J) \tag{11}$$

Once the binary transformation of all reviews has been performed, it is possible to compute the values of $\bar{p}_{d,t}$, i.e., the "winning" percentage of the d-th quality determinant in the t-th sampling period:

$$p_{d,t} = \frac{\sum_{j}^{R_t} V_{j,d}}{|R_t|} \qquad \forall\, t \in (1, \dots, T) \tag{12}$$

where R_t is the set of digital VoC collected in the *i*-th sampling period, $|R_t|$ is the cardinality of the R_t set.

The central line (CL) and the control limits (UCL, LCL) of the p-control chart are defined as follows (Montgomery, 2020):

$$CL_d = \overline{p_d} = \frac{\sum_{t=1}^{T}|R_t| \cdot \bar{p}_{d,t}}{\sum_{t=1}^{T}|R_t|} \tag{13}$$

$$UCL_{d,t} = \overline{p_d} + 3\sqrt{\frac{\overline{p_d}\,(1-\overline{p_d})}{|R_t|}} \tag{14}$$

$$LCL_{d,t} = \overline{p_d} - 3\sqrt{\frac{\overline{p_d}(1-\overline{p_d})}{|R_t|}} \tag{15}$$

As an example, Table 7 shows the values obtained for the development of the p-control chart for a six-month sampling period, related to the "Charges and fees" (d=4) quality determinant.

Table 7 - Values obtained for the construction of a p-control chart for the "Charges and fees" (d=4) quality determinant. Sampling rate: 1 semester. S1 and S2 indicate respectively the first and the second semester of each year.

Sampling period (t)	$\|R_t\|$	$p_{d,t}$	CL_4	$LCL_{4,t}$	$UCL_{4,t}$
S1 2006	26	0	0.0995	0	0,2756
S2 2006	28	0,0357	0.0995	0	0,2692
S1 2007	46	0,0652	0.0995	0	0,2319
S2 2007	54	0,1111	0.0995	0	0,2217
S1 2008	65	0,0615	0.0995	0	0,2109
S2 2008	51	0,0588	0.0995	0	0,2253
S1 2009	59	0,0338	0.0995	0	0,2164
S2 2009	80	0,05	0.0995	0	0,1999
S1 2010	66	0,2121	0.0995	0	0,2100
S2 2010	80	0,1750	0.0995	0	0,1999
S1 2011	98	0,1326	0.0995	0,0088	0,1902
S2 2011	79	0,1265	0.0995	0	0,2005
S1 2012	86	0,127	0.0995	0,0026	0,1963
S2 2012	68	0,0294	0.0995	0	0,2084
S1 2013	72	0,0833	0.0995	0	0,2053
S2 2013	80	0,1125	0.0995	0	0,1999
S1 2014	87	0,0689	0.0995	0,0032	0,1958
S2 2014	94	0,1276	0.0995	0,0069	0,1921
S1 2015	100	0,14	0.0995	0,0097	0,1893
S2 2015	102	0,1176	0.0995	0,0106	0,1884
S1 2016	102	0,1078	0.0995	0,0106	0,1884
S2 2016	136	0,1176	0.0995	0,0225	0,1765
S1 2017	110	0,0909	0.0995	0,0139	0,1851
S2 2017	82	0,0487	0.0995	0,0003	0,1987
S1 2018	101	0,0792	0.0995	0,0101	0,1889
S2 2018	104	0,125	0.0995	0,0114	0,1876
S1 2019	60	0,06667	0.0995	0	0,2154
S2 2019	54	0,07407	0.0995	0	0,2217

A graphical representation of the p-control charts for the "Charges and fees" (d=4) quality determinant is reported in Figure 4.

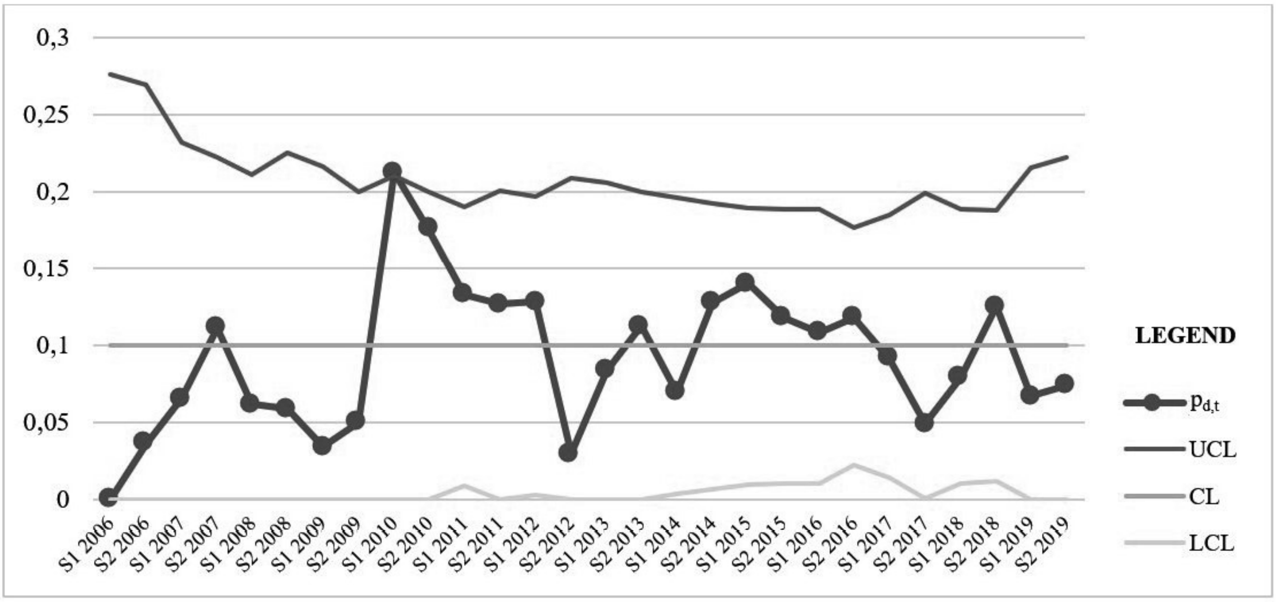

Figure 4 - **p-**control chart for the quality determinant “Charges and fees” (d=4).

In this case, the p-control chart shows a points outside the control limits. This occurs in the first semester of 2010 (S1 2010). Again, the causes of the out-of-control point must be analyzed in detail to avoid its recurrence. In this case the point does not represent an anomalous situation.

4. DISCUSSION AND CONCLUSIONS

The purpose of the current study was to investigate the applicability of Statistical Process Control tools for digital VoC analysis. Specifically, two different types of control charts, $\bar{x} - s$ and p, were considered to monitor the evolution of quality determinants over time. Control charts were developed based on the outputs generated by a topic modelling algorithm applied to a digital VoC database. A real case study concerning the car-sharing companies supported the explanation of the proposed approaches.

The results of this preliminary investigation show that control charts can provide valuable support in monitoring the quality determinants of products and services over time. The two types of control charts applied in this study are based on different principles, and their use has shown many differences. $\bar{x} - s$ control charts receive as input the average topical prevalence values generated by the topic modelling algorithm. In contrast, p control charts require the definition of the most discussed quality determinant for each VoC record. The introduction of the binary transformation may present some critical issues since it involves a pre-elaboration of the considered topical prevalences. In p-control charts, each quality determinant is mutually exclusive within a single digital VoC record; on the contrary, $\bar{x} - s$ control charts can consider several quality determinants that could be relevant within a single digital VoC record. This second option seems to be more useful in practical contexts. Overall, on the basis of this preliminary study, it appears that control charts for variables can provide more effective support in identifying out-of-control conditions of quality determinants, taking into account all information generated by the topic modelling algorithm.

SPC tools applied to digital VoC can support designers of products and services by providing a clear understanding of the evolution of quality determinants, allowing them to directly intervene in the design process in order to correct anomalies or unforeseen situations highlighted by the control charts. However, much more work will need to be done to successfully implement SPC in digital VoC analysis. The main limitation of this study is that the control charts proposed are suitable for monitoring stationary processes. However, some quality determinants show trends in the topical prevalence (e.g. a constant increase or decrease of $IMTP_{d,t}$ over time). Regression control charts or Moving Average control charts may represent valid tools to analyze these different behaviors.
Moreover, other metadata associated with the digital VoC (for example, record ratings) may be considered in the analysis. Further research will be directed on how to integrate this information to support designers to detect anomalous situations.

5. ACKNOWLEDGEMENTS

This work has been partially supported by the "Ministero dell'Istruzione, dell'Università e della Ricerca" Award "TESUN-83486178370409 finanziamento dipartimenti di eccellenza CAP. 1694 TIT. 232 ART. 6".

6. REFERENCES

Aguwa, C. C., Monplaisir, L., & Turgut, O. (2012). Voice of the customer: Customer satisfaction ratio based analysis. *Expert Systems with Applications*, *39*(11), 10112–10119.

Ashton, T., & Evangelopoulos, N. (2012). Control charts for customer comments: a case study and a research agenda. *Proceeding of the Southwest Decision Sciences Institute*, 661–669.

Ashton, T., Evangelopoulos, N., & Prybutok, V. (2014). Extending monitoring methods to textual data: a research agenda. *Quality & Quantity*, *48*(4), 2277–2294.

Ashton, T., Evangelopoulos, N., & Prybutok, V. R. (2015). Quantitative quality control from qualitative data: control charts with latent semantic analysis. *Quality & Quantity*, *49*(3), 1081–1099.

Bandaru, S., Gaur, A., Deb, K., Khare, V., Chougule, R., & Bandyopadhyay, P. (2015). Development, analysis and applications of a quantitative methodology for assessing customer satisfaction using evolutionary optimization. *Applied Soft Computing*, *30*, 265–278.

Barravecchia, F., Mastrogiacomo, L., & Fiorenzo, F. (2020). Categorizing quality determinants in mining user-generated contents. *Sustainability (Switzerland)*, *12*(23), 1–12, 9944.

Barravecchia, F., Mastrogiacomo, L., & Franceschini, F. (2020b). Identifying car-sharing quality determinants: A data-driven approach to improve engineering design. *International Conference on Quality Engineering and Management*, 125–140.

Barravecchia, F., Mastrogiacomo, L., & Franceschini, F. (2021). Digital voice-of-customer processing by topic modelling algorithms: insights to validate empirical results. *International Journal of Quality & Reliability Management*, ahead-of-print (doi: 10.1108/IJQRM-07-2021-0217).

Berry, M. W., & Kogan, J. (2010). *Text mining: applications and theory*. John Wiley & Sons.

Blei, D. M. (2012). Probabilistic topic models. *Communications of the ACM*, *55*(4), 77–84.

Blei, D. M., Ng, A. Y., & Jordan, M. I. (2003). Latent dirichlet allocation. *Journal of Machine Learning Research*, *3*, 993–1022.

Carnerud, D. (2020). The quality movement's three operational paradigms: a text mining venture. *The TQM Journal*, *32*(6), 1577–1598.

Chen, K.-S., Chang, T.-C., Wang, K.-J., & Huang, C.-T. (2015). Developing control charts in monitoring service quality based on the number of customer complaints. *Total Quality Management & Business Excellence*, *26*(5–6), 675–689.

Elg, M., Birch-Jensen, A., Gremyr, I., Martin, J., & Melin, U. (2021). Digitalisation and quality management: problems and prospects. *Production Planning & Control*, *32*(12), 990–1003.

Franceschini, F. (2002). Learning curves and p-charts for a preliminary estimation of asymptotic performances of a manufacturing process. *Total Quality Management*, *13*(1), 5–12.

Franceschini, F., Galetto, M., & Turina, E. (2009). Service quality monitoring by performance indicators: A proposal for a structured methodology. *International Journal of Services and Operations Management*, *5*(2), 251–273.

Freeman, G., & Radziwill, N. (2018). Voice of the Customer (VoC): A Review of Techniques to Reveal and Prioritize Requirements for Quality. *Journal of Quality Management Systems*, *17*, 1–29.

Gregorio, R., & Cronemyr, P. (2011). From expectations and needs of service customers to control chart specification limits. *The TQM Journal*, *23*(2), 164–178.

Hallencreutz, J., & Parmler, J. (2021). Important drivers for customer satisfaction – from product focus to image and service quality. *Total Quality Management & Business Excellence*, *32*(5–6), 501–510.

Izogo, E. E., & Ogba, I.-E. (2015). Service quality, customer satisfaction and loyalty in automobile repair services sector. *International Journal of Quality & Reliability Management*, *32*(3), 250–269.

Jach, P., Antony, J., Thomson, S. P., Cudney, B., & Furterer, S. (2021). Voice of the customer as a tool for service quality analysis in public transport. *The TQM Journal*, ahead-of-print(doi:10.1108/TQM-05-2021-0134).

Kamsu-Foguem, B., Rigal, F., & Mauget, F. (2013). Mining association rules for the quality improvement of the production process. *Expert Systems with Applications*, *40*(4), 1034–1045.

Kumar, A., & Anjaly, B. (2017). How to measure post-purchase customer experience in online retailing? A scale development study. *International Journal of Retail & Distribution Management*, *45*(12), 1277–1297.

Liang, Q., & Wang, K. (2019). Monitoring of user-generated reviews via a sequential reverse joint sentiment-topic model. *Quality and Reliability Engineering International*, *35*, 1180–1199.

Lo, S. (2008). Web service quality control based on text mining using support vector machine. *Expert Systems with Applications*, *34*(1), 603–610.

Lysenko-Ryba, K., Zimon, D., Madzík, P. and Šírová, E. (2022). Examination of aspects of pro-consumer refund system. *The TQM Journal*, Ahead-of-print. DOI:10.1108/TQM-06-2021-0182

Mastrogiacomo, L., Barravecchia, F., Franceschini, F., & Marimon, F. (2021). Mining quality determinants of Product-Service Systems from unstructured User-Generated Contents: the case of car-sharing. *Quality Engineering*, *33*(3), 425–442.

Matzler, K., & Sauerwein, E. (2002). The factor structure of customer satisfaction: An empirical test of the importance grid and the penalty-reward-contrast analysis. *International Journal of Service Industry Management*, *13*(4), 314–332.

McColl-Kennedy, J., & Schneider, U. (2000). Measuring customer satisfaction: Why, what and how. *Total Quality Management*, *11*(7), 883–896.

Montgomery, D. C. (2020). *Introduction to statistical quality control*. 8th edition, John Wiley & Sons, New York.

Mukherjee, S. P. (2019). *Quality, Domains and Dimensions*. Springer.

Psarommatis, F., Prouvost, S., May, G., & Kiritsis, D. (2020). Product quality improvement policies in industry 4.0: characteristics, enabling factors, barriers, and evolution toward zero defect manufacturing. *Frontiers in Computer Science*, *2*, 26.

Roberts, M. E., Stewart, B. M., & Tingley, D. (2019). STM: R package for structural topic models. *Journal of Statistical Software*, *91*(2), 1–40.

Su, Y., & Hwang, J.-S. (2020). Integration of customer satisfaction and sustained use of a product for value assessment. *Total Quality Management & Business Excellence*, *31*(15–16), 1760–1773.

Sweeney, J. C., Soutar, G. N., & Mazzarol, T. (2008). Factors influencing word of mouth effectiveness: receiver perspectives. *European Journal of Marketing*, *42*(3/4), 344–364.

Szwejczewski, M., Goffin, K., & Anagnostopoulos, Z. (2015). Product service systems, after-sales service and new product development. *International Journal of Production Research*, *53*(17), 5334–5353.

Wang, P. (2015). Exploring the influence of electronic word-of-mouth on tourists' visit intention: A dual process approach. *Journal of Systems and Information Technology*, *17*(5), 381–395.

Wang, Y., & Tseng, M. M. (2011). Integrating comprehensive customer requirements into product design. *CIRP Annals - Manufacturing Technology*, *60*(1), 175–178.

Xu, W., Guo, C., Guo, S., Wang, L., & Li, X. (2021). A novel quality comprehensive evaluation method based on product gene for solving the manufacturing quality tracking problem of large equipment. *Computers & Industrial Engineering*, *152*, 107032.