

A Hybrid Approach to Statistical Process Control

Massimiliano Giorgio^a, Michele Staiano^b

^aDipartimento di Ingegneria Aerospaziale e Meccanica, Seconda Università di Napoli
Real Casa dell'Annunziata 29, Via Roma - 81031 Aversa (Italy)
massimiliano.giorgio@unina2.it

^bDipartimento di Progettazione Aeronautica, Università di Napoli Federico II
80, P.le V. Tecchio - 80125 Napoli (Italy)
mstaiano@unina.it

Abstract

Successful implementation of statistical process control techniques requires for operational definitions and precise measurements. Nevertheless, very often analysts can dispose of process data available only by linguistic terms, that would be a waste to neglect just because of their intrinsic vagueness. Thus a hybrid approach, which integrates fuzzy set theory and common statistical tools, sounds useful in order to improve effectiveness of statistical process control in such a case. In this work, a fuzzy approach is adopted to manage linguistic information, and the use of a Chi-squared control chart is proposed to monitor process performance.

Key Words: Statistical Process Control; Fuzzy Set Theory; Generalized p Control Charts.

1. Introduction

Manufacturing was historically the cradle of statistical process control (SPC) and the use of control charts is nowadays widespread in every industrial sector. In recent years, in addition to well known techniques, many non-standard application of SPC charts have been proposed (for a review see: MacCarthy *et al.*, 2002). This evolution entails the same nature of data to be used for process monitoring and involves different methodologies to widen the range of problems which SPC charts could be applied.

Nevertheless, there is still a lack of methods enabling to directly use both fuzzy and crisp data. In fact, in these kind of situations it is very common that SPC techniques neglect vague information or, at least, neglect their vagueness. The first solution is inefficient because it implies a waste of information; on the contrary, the latter exposes to the risk of introducing into the analysis pieces of information that are not really available.

In this paper the above mentioned critical issue is tackled by adopting a hybrid approach which integrates fuzzy set theory and common statistical tools. Specifically, we formulated first of all a multiple input (crisp and fuzzy data) transfer function that provides, as single output, the quality characteristic of the process. Moreover, we introduced some simple SPC tools suitable to monitor and control this quality characteristic.

The whole proposal is presented by developing an example of application about an express courier service.

2. Notation

- l indicates the variable "loss of the package";
- u indicates the variable "resolution of complaint";
- t indicates the real variable "delivery delay";
- ξ indicates the fuzzy variable "pick-up punctuality";
- η indicates the fuzzy variable "package damage";
- ω indicates a fuzzy variable obtained in terms of ξ and η ;
- $\mu(\cdot)$ membership function;

Ξ, H, Ω_{η} generic fuzzy sets defined on the spaces of the variables ξ, η and ω : Ξ, H, Ω_{η} ;

$Z \sim N(0,1)$ is a standard Normal r.v.: $F_Z(z) = \int_{-\infty}^z \frac{1}{\sqrt{2\pi}} \exp\left(-\frac{x^2}{2}\right) dx$;

$\mathbf{Z} \sim N_r(\mathbf{0}, \mathbf{I})$ indicates that \mathbf{Z} is a Standard Multinormal r.v.;

\mathbf{I} indicates the identity matrix;

Σ indicates the variance-covariance matrix with elements $\sigma_{ii} = \text{Var}\{X_i\}$ and $\sigma_{ij} = \text{Cov}\{X_i, X_j\}$;

z_{γ} is the γ -100 quantile of the standard Normal r.v.: $F_Z(z_{\gamma}) = \gamma$;

$\chi^2_{v,\gamma}$ is the γ -100 quantile of Chi-squared r.v. with v degree of freedom;

$X \sim B(n, p)$ indicates that X is a Binomial random variable: $\Pr\{X = x\} = \binom{n}{x} p^x \cdot (1-p)^{n-x}$;

$\mathbf{X} \sim Mn_r(n, \boldsymbol{\pi})$ indicates that \mathbf{X} is a Multinomial randomvariable $\Pr\{\mathbf{X} = \mathbf{x}\} = n! \prod_{i=1}^{r+1} \frac{\pi_i^{x_i}}{x_i!}$;

$$\mathbf{X} \quad \mathbf{X} = (X_1, X_2, \dots, X_r)';$$

$$\mathbf{x} \quad \mathbf{x} = (x_1, x_2, \dots, x_r)';$$

$$\boldsymbol{\pi} \quad \boldsymbol{\pi} = (\pi_1, \pi_2, \dots, \pi_r)';$$

$$x_{r+1} \quad x_{r+1} = n - \sum_{i=1}^r x_i;$$

$$\pi_{r+1} \quad \pi_{r+1} = 1 - \sum_{i=1}^r \pi_i.$$

3. Quality evaluation and control

The aim of the present work is to propose a practical way to perform SPC, taking into account data regarding intangible and tangible - eventually vague - aspects of the process output.

The proposal relies on the idea that one or more experts are able to define a *transfer function* from some input variables (critical aspects of the service/product) to the overall process quality (output). Moreover, we assumed that skilled employees are able to routinely provide the above mentioned inputs of the transfer function. Specifically, we supposed that the evaluation of the inputs can be made on the basis of little reference keys, without claiming for stronger scale properties to express scores. In such a way the skilled employees are not compelled to quantify what is intangible for them in order to realize this evaluation.

Different approaches could be adopted to formulate the transfer function. For instance, some authors refer to a *process quality model* for the assessment, improvement and quality control, depicting a basic framework and its component via logical diagrams (Wood, 1994; Beamon *et al.*, 1998). In this paper a hybrid approach will be adopted to manage both vague and crisp information and the transfer function will be obtained via fuzzy mathematics (Zadeh, 1968, 1973).

The approach guarantees flexibility since the transfer function is stated via "IF...THEN" rules and elementary operators, which can be easily retuned or updated to consider any evolution in the customer requirements.

Finally, the use of a Chi-squared control chart is proposed to monitor process performance and an easy-to-use procedure for successful implementation of the chart is presented.

4. A practical example: an express courier service

To exemplify the use of the proposed approach let us consider the case of an express courier service. This choice was inspired by the fact that the Winner of the 1990 Malcolm Baldrige National Quality Award for Service was Federal Express Corporation (henceforth: FedEx).

To achieve excellence FedEx adopted a "key service measurement index" composed of the following twelve weighted critical categories (Hude, 1991):

- Delivery late on the right day - 1 point;
- Delivery on the wrong day - 5 points;
- Invoice adjustments requested by the customer - 1 point;
- Traces - 1 point;
- Missed pick-ups - 10 points;
- Complaints rehandled - 5 points;
- Lost packages - 10 points;
- Damaged packages - 10 points;
- Overgoods - 5 points;
- Calls abandoned by customers - 1 point;
- Missing proof of delivery information - 1 point;
- an International Service Quality Indicator - 1 point.

In order to control the process on the basis of these categories a demerit control chart could be implemented following one of the well known approaches (Jones *et al.*, 1999). These kind of charts imply that all variables related to the above categories are intended to be binary. Nevertheless, for some of the above categories additional information is often available and it would be a waste to neglect.

The alternative approach proposed in this paper explicitly aims in involving all the available information into the analysis, even considering data that could be intrinsically vague.

For the sake of simplicity, let us focus our attention on the following five variables:

- Loss of the package;
 - Resolution of complaint;
 - Delivery delay;
 - Pick-up punctuality;
 - Package damage;
-

related to five of the most critical categories considered by FedEx.

In order to implement SPC we assumed that:

- "loss of the package" is a binary variable, l - that assumes the value 1 if the package is lost, 0 otherwise;
- "resolution of complaint" is a binary variable, u - that assumes the value 1 if there is a complaint AND its resolution is unsatisfactory (so it has to be rehandled);
- "delivery delay" is assumed to be a real non negative variable, t , measured from the desired delivery time;
- the last two qualitative variables: "pick-up punctuality" and "package damage", are intrinsically vague so the fuzzy approach appears suitable to cope with them.

For instance a judgment about "pick-up punctuality", ξ , can be expressed on the following rating scale (from worst to best judgment):

{Late, Acceptable, Timely}

which is implemented by suitably chosen fuzzy sets defined in the *basic linguistic term set*. In Figure 1 the fuzzy sets corresponding to the three judgments are depicted by membership functions, $\mu_i(\xi)$ - $i \in \{L, A, T\}$, defined on the conventional interval $[0, 100]$.

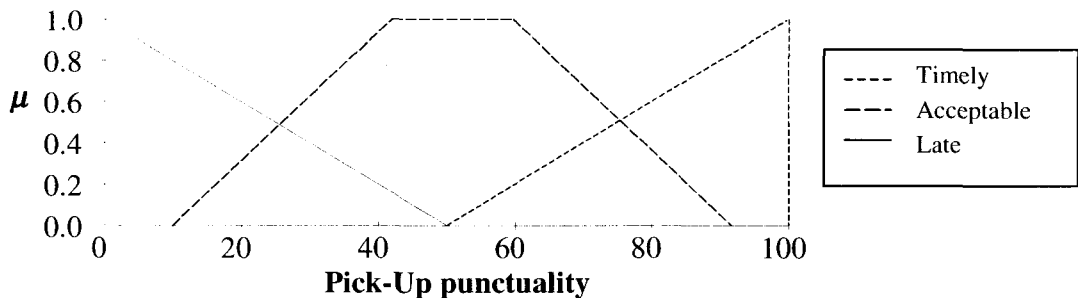


Figure 1. Fuzzy sets for the evaluation of Pick-Up punctuality on the basic linguistic term set

Similarly, the evaluation about "package damage", η , can be expressed on the following rating scale (see Figure 2):

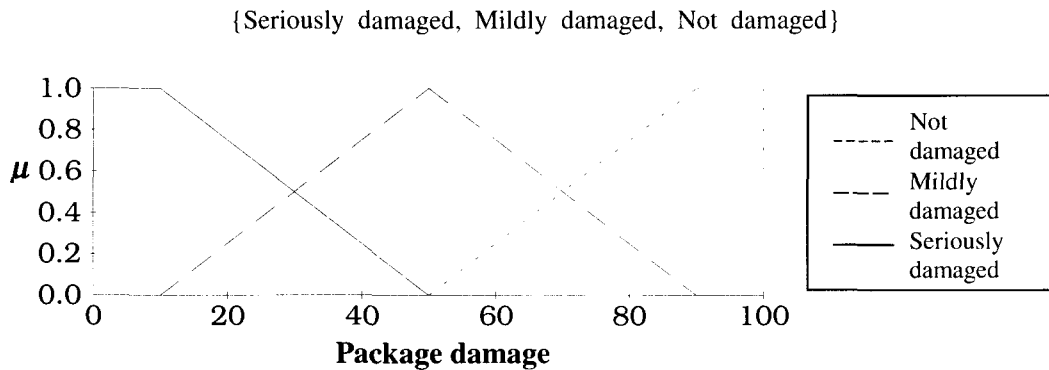


Figure 2. Fuzzy sets for the evaluation of Package damage chosen on the basic linguistic term set

As anticipated, the evaluation of all the five variables related to a specific delivery has to be expressed by appropriately skilled employees. These inputs are used to obtain an overall quality evaluation of process performance (the quality characteristic, QC) expressed on a $(r+1)$ level ordinal scale. The transfer function that provides the QC value, given the nature of the inputs, is obtained via a fuzzy approach. Obviously, this function has to take into account the peculiarities of the service. For this reason we assumed that it has to be compiled by an expert who carefully listened to the *Voice Of Customer*.

For instance, assuming $r=3$, the expert could provide the following transfer function:

IF $0 \leq q < q_1$ THEN QC = "no quality";

IF $q_1 \leq q < q_2$ THEN QC = "low quality";

IF $q_2 \leq q < q_3$ THEN QC = "acceptable quality";

IF $q_3 \leq q < q_{max}$ THEN QC = "high quality";

where

$0 < q_1 < q_2 < q_3 < q_{max}$ (in §3.1 q_{max} will be conventionally set to 4)

and

$$q = \begin{cases} (1-l) \cdot (1-u) \cdot \left[1 - \left(t/t_{max}\right)^2\right] \cdot \omega_i & t \leq t_{max} \\ 0 & t > t_{max} \end{cases}$$

The above formulation of q is consistent with these conceivable assumptions:

- q equals to 0 if there is the loss of the package ($l=1$);
- q equals to 0 if there is the unsatisfactory resolution of complaint ($u=1$);
- q equals to 0 if $(t > t_{max})$ - since the expert knows that the courier guarantees money refund if a maximum tolerable delivery delay, $t > t_{max}$, is exceeded;

- q equals to: $\left[1 - (t/t_{\max})^2\right] \cdot \omega_c$ if $[(l=0) \text{ AND } (u=0) \text{ AND } (t \leq t_{\max})]$ - where the (Taguchi loss) function $\left[1 - (t/t_{\max})^2\right]$ (see Figure 3) quantifies the (reverse) contribution of delivery delay t and ω_c (see § 3.1) expresses the contribution of the fuzzy inputs ξ and η .

Thus each delivery results in a $(r+1)$ -dimensional vector, with only one component equal to one and the others equal to zero.

Obviously, if the process is stable, on the basis of a sample of n deliveries it is possible to define a Multinomial r.v., \mathbf{X} , that can be used to perform the process control.

The corresponding control chart is a natural extension (Marcucci, 1985) of the classical p -chart, that will be presented into the essential details in the following section.

Both stability and quality drift can be diagnosed by means of the control chart setting the upper control limit to $\chi^2_{r,0.99730}$ and by computing the usual Pearson Chi-squared goodness of fit statistic.

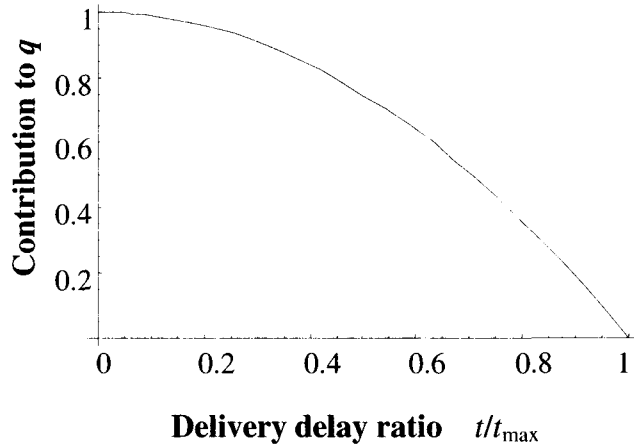


Figure 3. Reverse contribution of Delivery delay

The Fuzzy transfer function

In the following the algorithm to derive ω_c is described step by step.

The first step consists of defining the fuzzy rules:

$$\text{IF } \xi = \Xi_i \text{ AND } \eta = H_j \text{ THEN } \omega = \Omega_{ij}$$

that are summarized in the following figure.

Package Pick up	Seriously damaged	Mildly damaged	Not damaged
Timely	TS	TM	TN
Acceptable	AS	AM	AN
Late	LS	LM	LN

Figure 4. Rules used to define the fuzzy transfer function

Each rule corresponds to a fuzzy set defined on a conventional score range from zero to four (i.e.: $q_{max}=4$), as sketched in Figure 5.

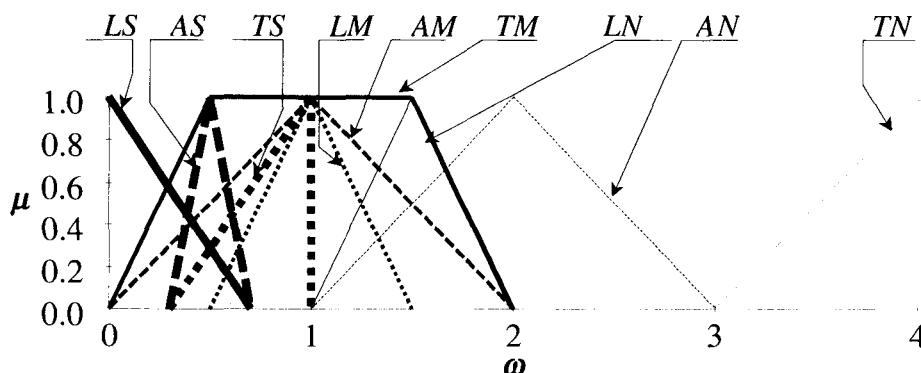


Figure 5. Fuzzy sets for the variable w on a conventional score scale ($q_{max}=4$)

To obtain the output value ω_c (see Figure 6 and Figure 7) the following additional steps are performed:

- the membership value of the two inputs (ξ , η) to each linguistic variable in the related term set (i.e. $i \in \{L, A, T\}$, $j \in \{S, M, N\}$.) is determined: $\mu_{\Xi}(\xi)$, $\mu_{H_j}(\eta)$
- the composition of antecedents Ξ_i, H_j is put into effect with the rule of the minimum:

$$\mu_{\Xi_i \cap H_j}(\xi, \eta) = \min[\mu_{\Xi_i}(\xi), \mu_{H_j}(\eta)]$$

- the alignment of the involved rules is realized by limiting the membership of the consequence Ω_{ij} to that one of the antecedents $\Xi_i \cap H_j$:

$$\mu'_{\Omega_{ij}}(\omega) = \min[\mu_{\Omega_{ij}}(\omega), \mu_{\Xi_i \cap H_j}(\xi, \eta)] \quad \omega \in E_{\Omega}$$

- the composition of the consequences Ω_{ij} into the output fuzzy set $C = \bigcup_{ij} \Omega_{ij}$ is obtained by the rule of the maximum:

$$\mu_c(\omega) = \max_{i,j} [\mu'_{\Omega_{ij}}(\omega)] \quad \omega \in E_\Omega$$

- a crisp representative value of the consequent part C , finally, results by means of the centroid:

$$\omega_c = \frac{\int_c \omega \mu_c(\omega) d\omega}{\int_c \mu_c(\omega) d\omega}$$

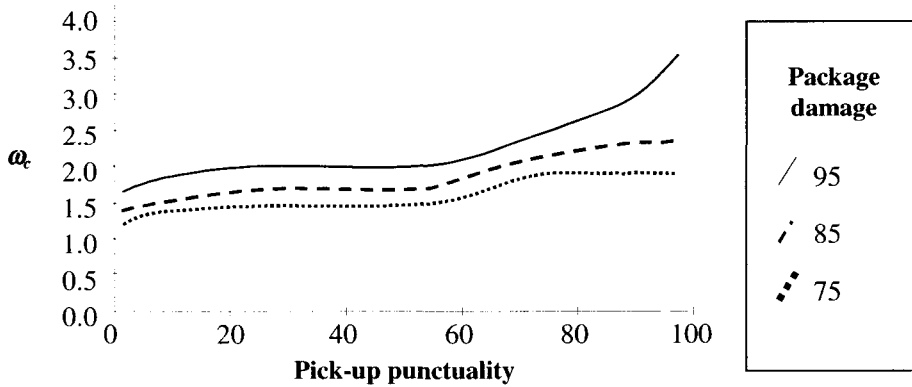


Figure 6. Fuzzy transfer function output value ω_c versus Pick-up punctuality for different levels of Package damage (see legend and refer to the value in abscissa in Figure 1 and 2).

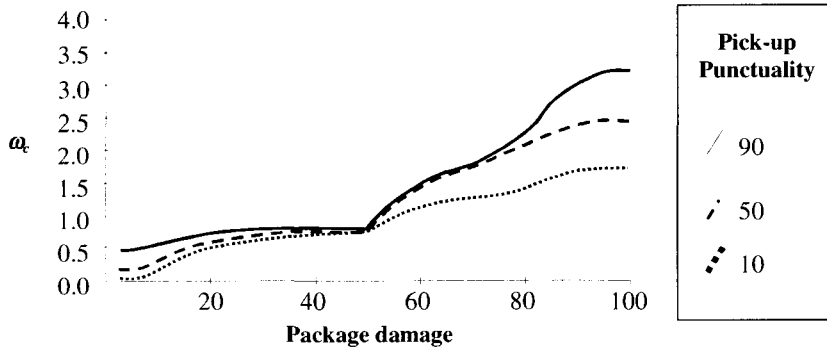


Figure 7. Fuzzy transfer function output value ω_c versus Package damage for different levels of Pick-Up punctuality (see legend and refer to the value in abscissa in Figure 1 and 2).

5. The generalized p -chart

Let us consider a Multinomial random variable: $\mathbf{X} \sim Mn_r(n, \boldsymbol{\pi})$. It is well known that:

$$E\{\mathbf{X}\} = n \cdot \boldsymbol{\pi},$$

$$\sigma_{ii} = \text{Var}\{X_i\} = n \cdot \pi_i \cdot (1 - \pi_i);$$

$$\sigma_{ij} = \text{Cov}\{X_i, X_j\} = -n \cdot \pi_i \cdot \pi_j \quad \forall i \neq j.$$

Accordingly to the *Central Limit Theorem*:

$$\Sigma^{-1/2} \cdot (\mathbf{X} - n\boldsymbol{\pi}) \stackrel{n \rightarrow \infty}{\sim} N_r(\mathbf{0}, \mathbf{I})$$

and also:

$$(\mathbf{X} - n\boldsymbol{\pi})' \cdot \Sigma^{-1} \cdot (\mathbf{X} - n\boldsymbol{\pi}) \stackrel{n \rightarrow \infty}{\sim} \chi_r^2$$

which give suitable approximations for large n .

The last asymptotic equality provides a natural generalization of the p -chart (see Appendix A) that can be used when the quality characteristic can be observed only as an attribute and the outcome of the inspection is described in more than two distinct ways (i.e. $r+1$ where $r > 1$).

In the first phase, on the basis of k samples of size n_i , it is simple to estimate the

components of the vector $\boldsymbol{\pi}$ using the formula $\pi_i = \pi_i = \frac{1}{n_{TOT}} \sum_{j=1}^k n_j \cdot \hat{\pi}_{ij}$, where $\hat{\pi}_{ij} = \frac{x_{ij}}{n_j}$.

The second phase is therefore straightforwardly implemented by setting the Upper Control Limit (UCL) to: $\text{UCL} = \chi_{r, 0.99730}^2$

and by computing:

$$(\mathbf{x}_i - n_i \boldsymbol{\pi})' \cdot \Sigma^{-1} \cdot (\mathbf{x}_i - n_i \boldsymbol{\pi}) = \sum_{j=1}^{r+1} (\hat{\pi}_{ij} - \pi_j)^2 / (n_i \pi_j)$$

that is the standard Pearson Chi-squared goodness of fit statistic. Obviously with Chi-squared chart the "out-of-control signal" is not directional, i.e. one cannot recognize if the change is a quality drift or an improvement. This situation calls for a proper *a posteriori* analysis.

Capability analysis and interpretation of out-of-control signals

At the end of the first phase, a so called *capability analysis* is appropriate, in order to check if the quality level is satisfactory given the actual market scenario and the defined quality goals.

Thus the aim of this analysis is to assure that the process is *suitable*, with respect to the constraints derived from the above mentioned considerations, and also *profitable*, i.e. the process is expected to guarantee an economical return.

The analysis can be performed comparing the quality profile (i.e.: $\pi_{01}, \pi_{02}, \pi_{03}, \pi_{04}$) shown by the process at the end of the first phase, to the *target profile* ($\pi_{C1}, \pi_{C2}, \pi_{C3}, \pi_{C4}$), that have to be defined by the management.

When a suitable *loss function* is available (i.e. it is possible to define the loss values a_1, a_2, a_3, a_4 corresponding to each of the four levels) the evaluation can be directly performed comparing the expected losses. Specifically it should be verified that:

$$\sum_{i=1}^4 a_i \pi_{0i} \leq \sum_{i=1}^4 a_i \pi_{Ci}.$$

Unfortunately very often there is not enough information to fully specify a loss function. This prevents one from performing a quantitative comparison between the quality profiles. Nevertheless, in many cases it is still possible to state whether the process is capable on the basis of logical considerations regarding the form of the loss function (i.e.: the lower the quality, the higher the loss - $a_1 > a_2 > a_3 > a_4$). Specifically, the analysis can be performed just comparing the cumulated stable profile (i.e.: $\pi_{01}, \pi_{01} + \pi_{02}, \pi_{01} + \pi_{02} + \pi_{03}, 1$) to the cumulated target profile (i.e.: $\pi_{C1}, \pi_{C1} + \pi_{C2}, \pi_{C1} + \pi_{C2} + \pi_{C3}, 1$) as depicted in Figure 8 and decisions about process capability can be assumed accordingly to the following rules (see Appendix B for details):

- if each point of the first profile is below the corresponding point of the latter profile (i.e. fall in the clear area of the bars) the process is considered capable;
- if all the points of the first profile are above the points of the latter profile the process is not capable and has to be improved or changed;
- if the profiles cross, an additional investigation is required in order to check if the stable profile is more profitable than the defined target profile - a conservative choice consists of acting as if the process was not capable.

As soon as the process has demonstrated to be capable, the second phase starts and the estimated probabilities π_{0i} are used to set up the generalized p -chart. As underlined in the previous section this kind of chart is not directional, i.e. a signal does not enable to detect directly whether the potential change is a quality drift or an improvement. Thus, it results that, when an out-of-control signal is detected, current process data ought to be examined and compared to past ones, in order to recognize the nature of the change.

Also in this second phase, thus, the above graphical tool could be used to perform this analysis. Specifically, the current cumulated quality profile ($\pi_1, \pi_1 + \pi_2, \pi_1 + \pi_2 + \pi_3, 1$) has to

be compared to the cumulated stable profile (i.e.: π_{01} , π_{02} , π_{03} , π_{04}).

Obviously, when it is not possible to diagnose a positive shift, the cumulated quality profile it ought to be compared to the target profile in order to timely recognize if the process is still capable.

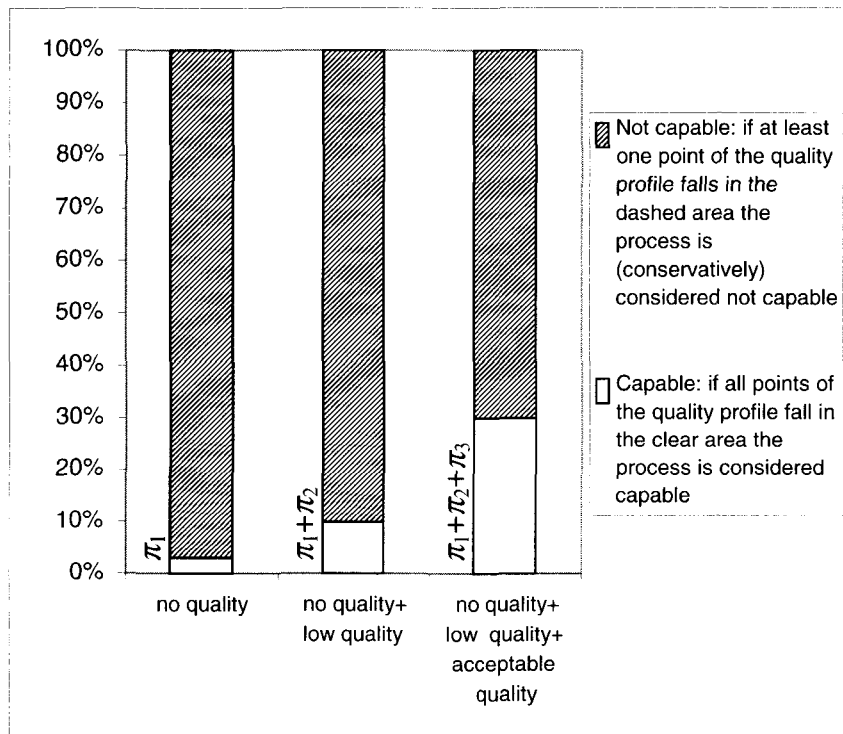


Figure 8. Diagram to visually control the quality profile in the capability analysis.

References

1. Beamon, B. M., Ware, T. M. (1998) "A process quality model for the analysis, improvement and control of supply chain systems", *Logistics Information Management*, Vol. 11, No. 2, pp. 105-113.
2. Ghobadian, A., Speller, S., Jones, M. (1994) "Service Quality - Concepts and Models", *International Journal of Quality & Reliability Management*, Vol. 11, No. 9, pp. 43-66.
3. Hude, H. (1991) "Quality in the Service Sector", *Making Statistics More Effective in*

Schools of Business, 6th Annual Conference, The Wharton School, University of Pennsylvania, June 7-8.

4. Kotz, S., Johnson, N. L. (1985), *Encyclopedia of Statistical Sciences*, Vol. 5, John Wiley & Sons, New York, pp. 659-664.
 5. Krzanowski, W., Marriott F. H. C. (1994), *Multivariate Analysis Part 1*, Arnold, 1st ed, London
 6. Laviolette, M., Seaman, J.W. Jr., Barrett, J.D., Woodall, W.H. (1995), "A probabilistic and statistical view of fuzzy methods" (with discussion), *Technometrics*, Vol. 37, No. 3, pp. 249-292.
 7. MacCarthy, B.L., Wasusri, T. (2002), "A review of non-standard applications of statistical process control (SPC) charts", *International Journal of Quality & Reliability Management*, Vol. (19 No. 3, pp. 295-320.
 8. Marcucci, M. (1985), "Monitoring Multinomial Processes", *Journal of Quality Technology*, Vol. 17, No. 2, pp. 86-91.
 9. Montgomery, D.C., Woodall, W.H. (1997), "A Discussion on Statistically-Based process Monitoring and Control" (with discussion), *Journal of Quality Technology*, Vol. 29, No. 2, pp. 121-211.
 10. Montgomery, D.C., (2001), *Introduction to Statistical Quality Control*, John Wiley & S., 4th ed., New York
 11. Wood, M. (1994), "Statistical Methods for Monitoring Service Processes", *International Journal of Service Industry Management*, Vol. 5, No. 4, pp. 53-68.
 12. Jones, L.A., Woodal, W. Conerly, M. D. (1999), "Exact Properties of Demerit Control Charts", *Journal of Quality Technology*, Vol. 31, No. 2, pp. 207-216.
 13. Johnson, N. L., Kotz, S. (1972), *Continuous multivariate distributions*, John Wiley & Sons, New York.
 14. Zadeh, L.A. (1968), "Fuzzy algorithms", *Information and Control*, Vol. 12, pp. 94-102.
 15. Zadeh, L.A. (1973), "Outline of a new approach to the analysis of complex systems and decision processes", *IEEE Transactions on Systems, Man, and Cybernetics*, Vol. 3, No. 1, pp. 28-44.
-

Appendix A

The p -chart

Let X be a Binomial r.v.:

$$X \sim B(n, p)$$

Then, obviously:

$$E\{X\} = n \cdot p$$

$$\text{Var}\{X\} = n \cdot p \cdot (1 - p)$$

Therefore the 3-sigma Upper Control Limits (UCL) Upper Control Limits (LCL) for the p -chart are defined as:

$$\text{UCL}_i = p + 3 \cdot \sqrt{[p \cdot (1 - p)] / n_i}$$

$$\text{LCL}_i = p - 3 \cdot \sqrt{[p \cdot (1 - p)] / n_i}$$

and the observed \hat{p}_i are plotted on the chart.

The chart is used throughout three main stages.

In the first, usually referred to as phase 1 analysis, k samples of size n_i are collected and

the trial control limits are evaluated, setting $p = p = \frac{1}{n_{TOT}} \sum_{i=1}^k n_i \cdot \hat{p}_i$ where $\hat{p}_i = x_i / n_i$ and

$$n_{TOT} = \sum_{i=1}^k n_i$$

If all the \hat{p}_i , $i=1, 2, \dots, k$, plot inside the control limits then one concludes that the process was in control in the past and trial control limits are suitable to control current or future production.

The "process capability analysis", reduces only to verify that the observed fraction of non-conformities \hat{p} does not exceed the specified maximum tolerable level p_T .

Then the phase 2 of control chart usage can be started: in this phase independent random samples are collected in different monitoring periods. Points \hat{p}_i are plotted (sample by sample) on the chart in order to verify that the process remains in control. The limits of the chart are calculated using for p the estimate, \hat{p} , obtained in the phase 1.

The χ^2 chart as alternative to the p -chart

Let us consider that:

$$\begin{aligned} & \Pr\left\{p - z_{1-\alpha/2} \cdot \sqrt{\left[p \cdot (1-p)\right]/n_i} \leq \hat{p} \leq p + z_{1-\alpha/2} \cdot \sqrt{\left[p \cdot (1-p)\right]/n_i}\right\} = \\ & = \Pr\left\{-z_{1-\alpha/2} \leq (\hat{p} - p) / \sqrt{\left[p \cdot (1-p)\right]/n_i} \leq z_{1-\alpha/2}\right\} = \\ & = \Pr\left\{\left[(\hat{p} - p) / \sqrt{\left[p \cdot (1-p)\right]/n_i}\right]^2 \leq \chi_{1,1-\alpha}^2\right\} \end{aligned}$$

so a control chart almost equivalent (i.e. it is not directional) to the classical 3 sigma p -chart could be implemented by setting the limit:

$$\text{UCL} = \chi_{1,0.99730}^2 = 9 \quad (\chi_{1,1-\alpha}^2 = z_{1-\alpha/2}^2 \Rightarrow \chi_{1,0.99730}^2 = z_{0.99865}^2 = 3^2 = 9)$$

and plotting on the chart the values $\frac{n_i(\hat{p}_i - p)^2}{p(1-p)}$, that (let $p_1 = p$, $p_2 = 1 - p$, $x_{i1} = x$, $x_{i2} = n_i - x_i$) can also be calculated as:

$$\sum_{j=1}^2 (x_{ij} - n_i \cdot p_j)^2 / (n_i \cdot p_j)$$

Appendix B

Say L the expected loss evaluated from the probabilities π_i and the loss values a_i related to each of the four quality categories. The estimate, L_t , and the reference one, L_0 are, respectively:

$$L_t = a_1 \cdot \pi_{t1} + a_2 \cdot \pi_{t2} + a_3 \cdot \pi_{t3} + a_4 \cdot \pi_{t4} = \sum_{i=1}^4 a_i \cdot \pi_{ti}$$

$$L_0 = a_1 \cdot \pi_{01} + a_2 \cdot \pi_{02} + a_3 \cdot \pi_{03} + a_4 \cdot \pi_{04} = \sum_{i=1}^4 a_i \cdot \pi_{0i}$$

and the expected loss change ΔL can be simply evaluated by means of the relation:

$$\Delta L = L_t - L_0 = \sum_{i=1}^4 a_i \cdot (\pi_{ti} - \pi_{0i}) = \sum_{i=1}^4 a_i \cdot \Delta \pi_i,$$

where:

$$\Delta \pi_i = \pi_{ti} - \pi_{0i}; \quad i = 1, 2, 3, 4.$$

Obviously, being:

$$\sum_{i=1}^4 \Delta \pi_i = \sum_{i=1}^4 (\pi_{ti} - \pi_{0i}) = \sum_{i=1}^4 \pi_{ti} - \sum_{i=1}^4 \pi_{0i} = 0$$

it is easy to verify that:

$$\begin{aligned} \Delta L &= \sum_{i=1}^4 a_i \cdot \Delta \pi_i = \\ &= (a_1 - a_2) \cdot \Delta \pi_1 + (a_2 - a_3) \cdot (\Delta \pi_1 + \Delta \pi_2) + (a_3 - a_4) \cdot (\Delta \pi_1 + \Delta \pi_2 + \Delta \pi_3) + a_4 \cdot (\Delta \pi_1 + \Delta \pi_2 + \Delta \pi_3 + \Delta \pi_4) = \\ &= \sum_{i=1}^3 b_i \cdot \left(\sum_{j=1}^i \Delta \pi_j \right) \end{aligned}$$

where the coefficients b_i 's satisfy the conditions:

$$b_i = a_i - a_{i+1} > 0; \quad i = 1, 2, 3.$$

Since a positive value of ΔL implies a drift whereas a negative value implies an improvement, given that the b_i 's are all positive it results that:

$$\begin{aligned} \sum_{j=1}^i \Delta \pi_j \geq 0; \forall i \quad \text{and} \quad \exists i: \sum_{j=1}^i \Delta \pi_j > 0 &\Rightarrow \sum_{i=1}^3 b_i \cdot \left(\sum_{j=1}^i \Delta \pi_j \right) = \Delta L > 0 \\ \sum_{j=1}^i \Delta \pi_j \leq 0; \forall i \quad \text{and} \quad \exists i: \sum_{j=1}^i \Delta \pi_j < 0 &\Rightarrow \sum_{i=1}^3 b_i \cdot \left(\sum_{j=1}^i \Delta \pi_j \right) = \Delta L < 0 \end{aligned} \quad i \in \{1, 2, 3\}$$

In other situations it is not possible to diagnose the presence of drift/improvement without making additional hypotheses about the values of the a_i 's and/or b_i 's.