

Application du SPC pour le contrôle des METAR

Islam BOUSRI^{1*}, Abdelhadi SABBA², Sara CHIKHI¹,

Abstract

La Maîtrise statistique des procédés, ou le SPC est une application qui utilise des techniques statistiques pour contrôler un procédé ou une méthode de production. Les principaux outils et procédures utilisés peuvent nous aider à surveiller le comportement des processus, à découvrir les problèmes dans les systèmes internes et à trouver des solutions aux problèmes de production. Le présent travail représente une adaptation de cette technique pour contrôler la qualité des messages METAR, et SYNOP. Les données sous forme de paramètres mesurés sont obtenues en temps réel. Ces données sont ensuite reportées sur un graphique avec des limites de contrôle prédéterminées. Ces limites sont déterminées par la moyenne et la taille de l'échantillon.

Keywords

La méthode SPC, METAR, SYNOP, principe de Pareto, Carte de contrôle.

¹ *Prévision Numérique du Temps, Centre National de la météorologie, Office Nationale de La Météorologie, 1, Avenue Mohamed Khemisti BP 153 Dar el Beida-16011- Alger. Algérie*

² *Direction Météorologique de la Région Centre, Office Nationale de La Météorologie, 1, Avenue Mohamed Khemisti BP 153 Dar el Beida-16011- Alger. Algérie*

*Correspondant: bousri.islam@gmail.com

Contents

Introduction	1
1 Théorie et formules Mathématiques	2
1.1 Le principe de Pareto	2
1.2 Carte de contrôle	2
2 Résultats et discussions	3
3 Conclusion	4
References	4

Le SPC ou Statistical Process Control est défini comme un élément d'assurance qualité. Son but est de contrôler un processus évaluable, et de le traduire en graphique temporel basé sur des fondements statistiques. On dit qu'un processus est sous contrôle statistique lorsque toutes les causes spéciales de variation ont été éliminées et que seule la variation de cause naturelle (ou courante) subsiste. Un graphique de contrôle est illustré par des données qui se situent dans les limites de contrôle en l'absence de modèles ou de tendances non aléatoires. Les méthodes de contrôle des processus statistiques sont appliquées à pratiquement tout type d'organisation, y compris les organisations de service, de fabrication et d'enseignement. M.Pettersson a utilisé en 2004, la méthode SPC comme moyen d'amélioration de la précision de détection des taux de désabonnement croissants. Il a montré que les méthodes SPC peuvent fournir aux analystes de marchés un outil puissant pour suivre les mouvements et les désabonnements des clients. Un système d'alerte précoce (SAP), basé sur les mêmes idées que celles utilisées dans les industries de transformation, a été construit, il donnera une prévoyance et un temps plus long pour réagir contre le désabonnement, offrant ainsi un avan-

tage sur les concurrents [1]. F. MONFORT a réalisé une application de la carte de contrôle à la surveillance du taux en filature peignée en 2005. Il a discuté l'efficacité statistique du test de contrôle quotidien du contremaître en filature et la précision souhaitable dans le centrage initial du taux [2]. En 2006, B.Fournier et al, ont présenté une famille très flexible de distributions statistiques et ont proposé l'introduction dans la méthodologie SPC des distributions lambda généralisées (GLD). Ils ont démontré par cette étude que les limites de contrôle généralement considérées dans le RCP sont prévues avec précision lors de la modélisation des lois statistiques habituelles au moyen de ces distributions. Ils ont traité aussi des données expérimentales issues d'une ligne de production industrielle à l'aide d'un logiciel basé sur les règles SPC Western Electric [3]. S.L.Jeng a développé en 2007 un modèle de réseau neuronal avec analyse en composantes indépendantes (ICA) afin d'identifier la perturbation et reconnaître les changements dans les paramètres de processus corrélés. Ces modèles de réseaux conceptuels peuvent être utilisés pour surveiller et éliminer les paramètres du processus de fabrication lorsqu'une perturbation se produit dans le processus sous-jacent. Il a tracé le graphique de Shewhart traditionnel et le graphique à somme cumulée (CUSUM), pour l'ensemble des données simulées afin d'évaluer la capacité d'identification de l'approche proposée [4].

Le but de notre travail est d'appliquer la méthode de contrôle des processus statistiques au contrôle des messages METAR. Plus précisément; la carte de contrôle, et le principe de Pareto.

1. Théorie et formules Mathématiques

Le principe de Pareto

Le principe de Pareto est également appelé la loi des 80-20, ou le principe de la rareté des facteurs [5][6]. Il indique que pour de nombreux événements, environ 80% des effets proviennent de 20% des causes. Il tient son nom de l'économiste italien Vilfredo Pareto qui, en 1906, a remarqué que 80% des terres en Italie appartenaient à 20% de la population. Il a confirmé le principe lorsqu'il a également remarqué que 20% des gousses de pois dans son jardin produisaient 80% des pois [6]. Il a ensuite mené des enquêtes sur une variété d'autres pays et a découvert que le principe s'appliquait en matière de propriété foncière par rapport à la population. Une observation attentive de nombreux domaines permettrait de réaliser que ce principe est en fait plus répandu. En affaires par exemple, une règle de base est que 80% de vos ventes proviennent de 20% de vos clients. Bien qu'il ne soit pas surprenant que les 20% les plus aisés de la population mondiale contrôlent 82,70% de leurs revenus [7], il est intéressant de savoir que la règle s'applique également aux sous-ensembles de la fourchette de revenus. Un exemple est que les 3 personnes les plus riches du monde en possèdent autant que les 7 suivantes réunies [7]. Dans les lignes qui suivent la loi de Pareto est présentée sous un angle pratique et très simplifié qui est généralement employé dans les domaines de la qualité et de la gestion. Néanmoins cette loi est définie de manière très précise sur le plan mathématique. Soient a et b deux nombres réels strictement positifs. La densité de probabilité d'une loi de Pareto de paramètres a et b est définie pour $x \in [a; +\infty[$ par la formule :

$$f_{a,b} = \frac{b \cdot a^b}{x^{b+1}} \quad (1)$$

Et la fonction de répartition est :

$$F_{a,b}(x) = 1 - \left(\frac{a}{x}\right)^b \quad (2)$$

$$F_{a,b}^{-1}(P) = \frac{a}{(1-P)^{\frac{1}{b}}} \quad (3)$$

avec $0 \leq P < 1$

Carte de contrôle

Le concept de la carte de contrôle statistique a été développé pour la première fois dans les années 1920 par Shewhart [8] afin d'améliorer la fiabilité des systèmes de transmission téléphonique. Le concept résulte de l'observation que les opérateurs réagissent souvent de manière excessive et apportent des changements inappropriés dans les paramètres en réponse à des variations d'indicateurs qui sont simplement aléatoires [9]. Des connaissances préalables sont requises pour la conception de la carte p . Lors de la création

d'un diagramme p , la taille de l'échantillon et la fréquence d'échantillonnage doivent d'abord être spécifiées avant de tracer la proportion observée d'événements indésirables dans chaque échantillon successif [10]. La carte p , où p représente la proportion, est utile pour la surveillance de routine d'un résultat binaire, comme la survenue d'un événement indésirable (par exemple, suivre la proportion de produits non conformes). Pour X une variable aléatoire qui à un METAR reçu $X = 1$ si le message est juste sinon $X = 0$. La proportion p de messages erronés dans l'ensemble de la réception correspond à la probabilité pour que X prenne la valeur 1. X suit une loi de Bernoulli de paramètre $B(p)$. Lorsqu'un échantillon de n message est tiré dans l'ensemble des messages reçus, on sait que la variable aléatoire $\bar{X} = \frac{1}{n} \sum_{i=1}^n X_i$ qui à un tel échantillon associe le nombre de messages erronés dans l'échantillon comme suit :

- La loi hypergéométrique $H(N, pN, n)$, N étant le nombre total de messages reçus, lorsque le tirage de l'échantillon est un tirage sans remise.
- La loi binomiale $B(n, p)$ lorsque le tirage de l'échantillon est un tirage avec remise, ou considéré comme tel si la taille de l'échantillon est inférieure à 10% de la taille de la population.

Dans notre cas, nous nous placerons dans le second cas.

Alors, la variable aléatoire $\bar{X} = \frac{1}{n} \sum_{i=1}^n X_i$

, qui à chaque échantillon de taille n associe la proportion de messages erronés dans cet échantillon \bar{p} pour moyenne : p , et pour écart type : $\frac{\sqrt{p(1-p)}}{\sqrt{n}}$

La proportion théorique de messages erronés étant p , les limites de la carte p de Shewhart sont, pour un échantillon de taille n_i :

$$p_i \in \left[p - \frac{\sqrt{p(1-p)}}{\sqrt{n_i}}, p + \frac{\sqrt{p(1-p)}}{\sqrt{n_i}} \right]$$

$K=1,2$ Voici comment se présente une carte p :

- $p - 2 \frac{\sqrt{p(1-p)}}{\sqrt{n_i}}, p + 2 \frac{\sqrt{p(1-p)}}{\sqrt{n_i}}$: Sont nommées respectivement limite de contrôle inférieure (LCI) et limite de contrôle supérieure (LCS) de la carte.
- $p - \frac{\sqrt{p(1-p)}}{\sqrt{n_i}}, p + \frac{\sqrt{p(1-p)}}{\sqrt{n_i}}$: Sont nommées respectivement limite de surveillance inférieure (LSI) et limite de surveillance supérieure (LSS) de la carte.

On remplace systématiquement LIC ou de LIS par 0 si le calcul de LIC ou de LIS donne un résultat négatif. On remarquera que les limites de contrôle et de surveillance varient en fonction de la taille de l'échantillon. Pour éviter cette variation, il est fréquent d'imposer que les échantillons aient tous la même taille. Pour chaque échantillon prélevé de n_i messages, on calcule la proportion d'erreur :

$$p_i = \frac{D_i}{n_i} \quad (4)$$

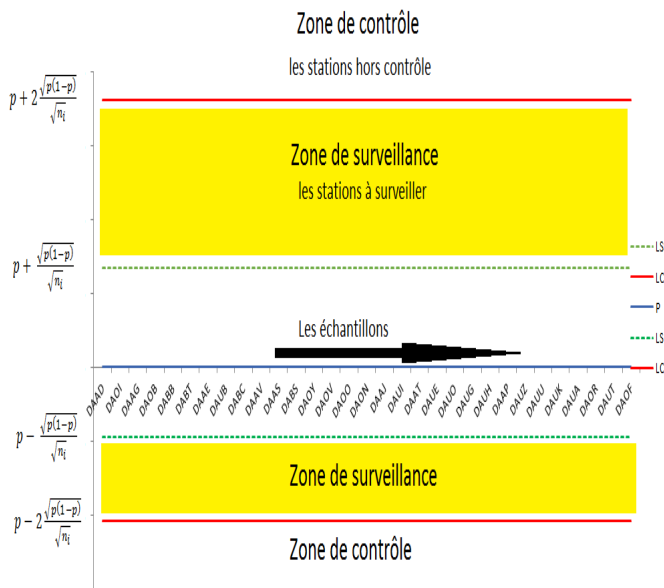


Figure 1. Représentation générale d'une carte de contrôle.

faisant le rapport du nombre de message erroné dans l'échantillon: D_i , par le nombre de message reçu dans l'échantillon $:n_i$. Sur la carte de contrôle, on porte la ligne centrale, les limites de contrôle et les points M_i de coordonnées (i, P_i) . Si tous les points M_i sont situés entre les limites de contrôle, le processus est déclaré maîtrisé ; dans le cas contraire, le processus est déclaré non maîtrisé.

2. Résultats et discussions

Appliquant la carte P, durant le mois de juillet 2019, nous avons calculé la proportion d'erreur pour chaque station en fonction du nombre de messages reçus erronés respectivement (Tableau 01), afin de calculer la moyenne p , les limites de contrôle (LCS, LCI) et les limites de surveillance (LSS, LSI) représentés dans le Tableau 02.

La figure suivante représente une carte P, des proportions d'erreurs des messages METAR du le mois de juillet 2019. La ligne rouge représente la limite de contrôle supérieur, en orange, la limite de surveillance supérieure, en bleu la proportion P , alors que les limites de contrôle et de surveillance inférieures ne sont pas représentées sur la carte, car elles comportent des valeurs négatives, ce qui implique une valeur nulle pour ces dernières.

Nous avons remarqué qu'il n'y a pas une grande variation de la moyenne et que tous les points se trouvent entre les deux limites (contrôle et surveillance), sauf pour la station de Ghardaïa qui a dépassé légèrement le seuil de surveillance. Cette analyse montre que le processus de production des METAR est stable et est sous contrôle. Aussi, la station de Ghardaïa est une station à surveiller, ce qui nécessite une action corrective à envisager, après détermination de la cause des erreurs commises. Appliquant le principe de Pareto, nous avons constaté que la source principale

Nom de Station	Réçu	Nombre d' Erreur	Taux d' Erreur
BOU-SAADA	402	0	0,00%
CHLEF	744	1	0,13%
DAR ELBEIDA	1487	0	0,00%
TIARET	744	0	0,00%
ANNABA	1486	2	0,13%
BATNA	742	0	0,00%
BEJAIA SOUMMAM	744	1	0,13%
BISKRA	743	0	0,00%
CONSTANTINE	1485	3	0,20%
JIJEL-ACHOUAT	743	0	0,00%
SETIF-AIN ARNET	742	0	0,00%
TEBESSA	1487	0	0,00%
EL-BAYADH	744	2	0,27%
MASACRA GHRISS	744	0	0,00%
ORAN-SENIA	1485	4	0,27%
TLEMCEM-ZENATA	1487	0	0,00%
DJANET	742	0	0,00%
IN-SALAH	741	2	0,27%
TAMANRASSET	1487	2	0,13%
OUARGLA	743	0	0,00%
TOUGGOURT	744	0	0,00%
ADRAR	740	0	0,00%
BECHAR	743	0	0,00%
TIMIMOUN	743	0	0,00%
TINDOUF	744	0	0,00%
EL-GOLEA	743	3	0,40%
EL-OUED	744	0	0,00%
GHARDAIA	744	5	0,67%
HASSI MESSAOUD	1487	1	0,07%
ILLIZI	744	0	0,00%
IN-AMENAS	738	3	0,41%

Table 1. La proportion d'erreur pour chaque station

n	31	
P	0,10%	
LCS	1,80%	
LSS	0,67%	
LSI	-0,47%
LCI	-1,03%

Table 2. Paramètres de la carte de contrôle.

des erreurs commises dans les METAR, provient du codage du groupe vent. Le diagramme indique 11 erreurs dans le groupe vent, ce qui représente 38 % de toutes les erreurs. 72 % des erreurs globales proviennent des deux premiers types : vent et phénomène.

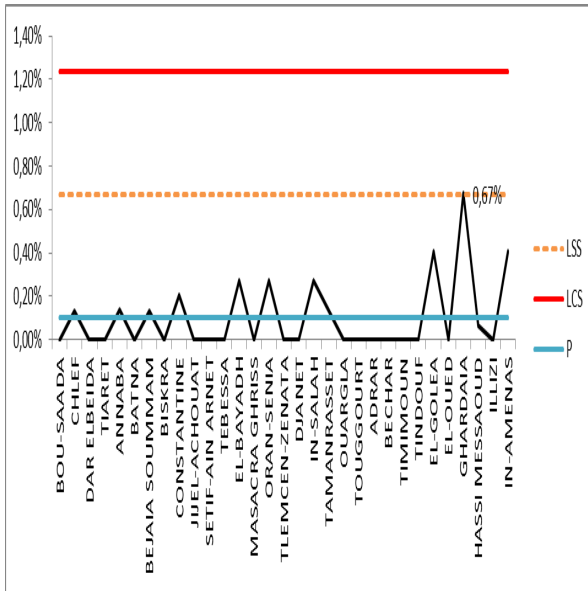


Figure 2. La carte P des METAR pour le mois de Juillet 2019.

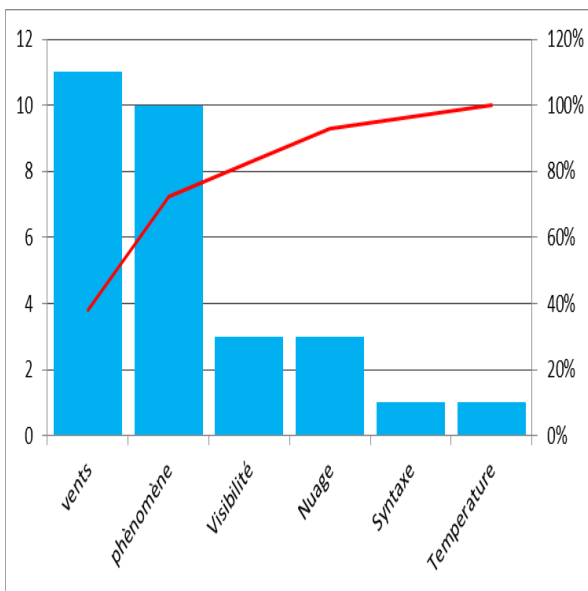


Figure 3. Carte de Pareo des erreurs.

3. Conclusion

Une approche de contrôle qualité des METAR par le SPC a été réalisée grâce à l'utilisation de la carte de contrôle, et le principe de Pareto. Ces deux derniers représentent les principaux outils de la maîtrise statistique des procédés. L'étude facilite l'évaluation de la qualité des METAR, par la détection des anomalies, et la détermination des causes de la plupart des erreurs. Dans plus de 90 % des METAR reçus, les erreurs commises ont porté sur les phénomènes suivants : vents, temps présent, visibilité et nuage).

References

- [1] Magnus Pettersson. Spc with applications to churn management. *Quality and Reliability Engineering International*, 20(5):397–406, 2004.
- [2] F Monfort. Application de la carte de contrôle à la surveillance du taux en filature peignée. efficacité du test. centrage initial. *Revue de Statistique Appliquée*, 3(4):27–43, 1955.
- [3] Benjamin Fournier, Nicolas Rupin, Maxence Bigerelle, Denis Najjar, and Alain Iost. Application of the generalized lambda distributions in a statistical process control methodology. *Journal of process control*, 16(10):1087–1098, 2006.
- [4] Shuen-Lin Jeng, Jye-Chyi Lu, and Kaibo Wang. A review of reliability research on nanotechnology. *IEEE Transactions on reliability*, 56(3):401–410, 2007.
- [5] Nick Bunkley. Joseph juran, 103, pioneer in quality control, dies. *New York Times*, 3:50–55, 2008.
- [6] A Narula. 80/20 rule of communicating your ideas effectively, 2005.
- [7] Ankunda R Kiremire. The application of the pareto principle in software engineering. *Consulted January*, 13:2016, 2011.
- [8] Walter Andrew Shewhart. *Economic control of quality of manufactured product*. Macmillan And Co Ltd, London, 1931.
- [9] Donald M Berwick. Controlling variation in health care: a consultation from walter shewhart. *Medical care*, pages 1212–1225, 1991.
- [10] Douglas C Montgomery. *Statistical quality control*. Wiley Global Education, 2012.