XLI CILAMCE
XLI Ibero-Latin American Congress on Computational Methods in Engineering
16-19 November, 2020 FOZ DO IGUAÇU PARANÁ - BRAZIL

Analysing a statistic index in identify change in data behaviour : Apply case in Itaipu Dam

Guilherme L. Campos[1], Sheila R. Oro[1], Cláudio I. Osako[2]

[1]*Federal Technology University - Paraná, Francisco Beltrão, Brazil*
Line Santa Bárbara, s/n, PR, 85601-970, Francisco Beltrão, Paraná
guilherme_lopescampos@hotmail.com, sheilaro@utfpr.edu.br
[2]*Hydroeletric power plant Itaipu Dam*
Foz do Iguaçu - Av. Tancredo Neves, 6.731 CEP 85856-970 Foz do Iguaçu, Paraná, Brasil
osako@itaipu.gov.br

Abstract. The safety of dam is a set of techniques whom to obtain maximum information and data to forecast possible structures damage. The statistical indexes are multivariable approach that to account information from: deformation, uplift, shift. The present paper, apply the factorial analysis to modelling a statistical index, in the purposes of describe jointly the structural status and assess sensibility of index, on the way to complement and describe the better control chart and quantify the relation between new data and extrapolations of control limits. The control chart in this paper, it was used as feedback of type : exponentially weighted with moving average. The analysis through plot, relate amount of occurs and standard deviation manipulated in the exponential regression between variables. The sensibility brings better interpretation for occurs and in decision- making.

Keywords: Dam monitoring, Factor analysis, IMCRB, Safety of Dam, Multivariate analysis.

1 Introduction

The monitoring of dam, to concern application of a set of techniques and operations that allows structure situation and safety parameters. [1] Instruments used in this action permitted measure : inherent shifts , uplift , farthest effects, environmental temperature and of concrete, reservoir level (headwater level, tailwater level), seepage, leakage, structure damage and others.[2] Through phenomenon monitored it is possible established two variable group: environmental and effects. The effect variables to concern of condition of structure, whereas, measure propriets that surfacing at the place of dam. [3]

The mainly instruments that are used in the monitoring of a concrete dam are: piezometer, strain gauge, deformeter and pendulum. Piezometer, are installed in jointed foundation on different level of depth, to measure uplift of water that infiltrated in the interior structure, could be of type standpipe or geonor, it aspects that distinguish is the way to obtain a data, manually or electronically, respectively. Strain gauge, measure a deformation of rocky massif in function a reference point. Deformeter, is install between joints to ascertain three – dimensional displacement; Pendulum, segregated on two categories : direct and invert, measure the displacement at top of dam and foundations using a reference point, respectively. [4]

Identifying the variability of instruments readings, it is possible to apply multivariate methods to analyse evolution from the data together .[5]. Clustering methods or order reduction, it was applicated in the context of dam monitoring, such as, Factor analysis. This method be based on group a set of data in small group of factors, that represents the joint variability of data, losing minimal information in the modeling. The previously results get notable income in cases applicated at Itaipu Dam, located in Foz do Iguaçu, Brazil. [5]

To perform the modeling of index that represents the global trend of key blocks, several methods can be used , such as: statistical models, machine learning methods, small probability, outliers attached data mining, among others. [3]

Recently, it was developed Joint Monitoring Index of Dam Blocks Response (IMCRB), through Factor analysis, constitute a dimensionless statistic parameter, that summarises all information of a set instruments in a small factor number [5]. At [6], it was applied this methodologic for two key blocks of Itaipu Dam, type gravity dam, obtained 90.1% variability of data by pendulum, piezometer, deformeter and strain gauge.

The IMCRB is characterise as latent variable it is not possible directly measure. Modeling by scores and weighted factors, using the control chart to analysis monitoring process. Permit combine a set of instruments in a just variable and analysing through limits control, it was a interesting about the index's sensibility, in the other words, quantify the difference between each point at relation mean that could unstable of monitoring process, at direct relation, extrapolate the limits in control chart. The knowledge of sensibility would substantiate decision making as the proposition with original information's (instruments) and (environmental phenomena's).

The purpose of paper is modelling IMCRB to buttress dam I10 of Itaipu dam, herewith estipulate control limits for exponentially weighted moving average control chart.

2 Methods

The study realized focus on a quantitative and describing research, so looking for establishment link among data provided of different instruments that are used in dam monitoring at Itaipu dam [2, 7, 8]. Subsequently, are presenting of data and steps realized to obtain IMCRB e controls limits of plot exponentially weighted,

2.1 Data

The hydroeletric power plant Itaipu, located in Foz do Iguaçu, Paraná, Brazil, it has dimension of 7.919 meters length, 196 meters height, build in multiples types of dam, such as: Earth, Rockfill, Gravity, Buttress. The blocks that are more instrumented, are referred to as key blocks.

The paper, look upon data of the instruments installed in key block I10 (Table 1), kind buttress dam, located in left of Itaipu Dam, derive from manual readings, undertaken from January 1990 to December 2019, accessible through center for advanced studies on dam safety (CEASB) incorporated at Itaipu technology park.

Table 1. Amount of instruments - Key block I10

Instruments	Amount	Physics Variable	Measure unit
Strain gauge	10	Deformation of foundations	mm
Deformeter	6	Three dimensional displacement	mm
Piezometer	5	Uplift	msnm
Pendulum	4	Displacement at top of dam and foundations	mm
Total	25	-	-

The data was described as: non – structure for contains different class of variables like double, string, categorical and datatime; [9]. It was organized in electronic spreadsheet and adjusted to a monthly periodicity. The process of structure in this way, it was necessary to twice reasons: 1) By periodicity chosen, there was instruments that contains more than a measure per month or missing data; 2) To identify standards in each class of instruments used.

2.2 Method

The step used in the scientific research it was presented in figure 1. The control loop was rated as: multivariable, closed, feedback, there is no accounting of delay without time, due to the rapid response of the effect variables. [10, 11].

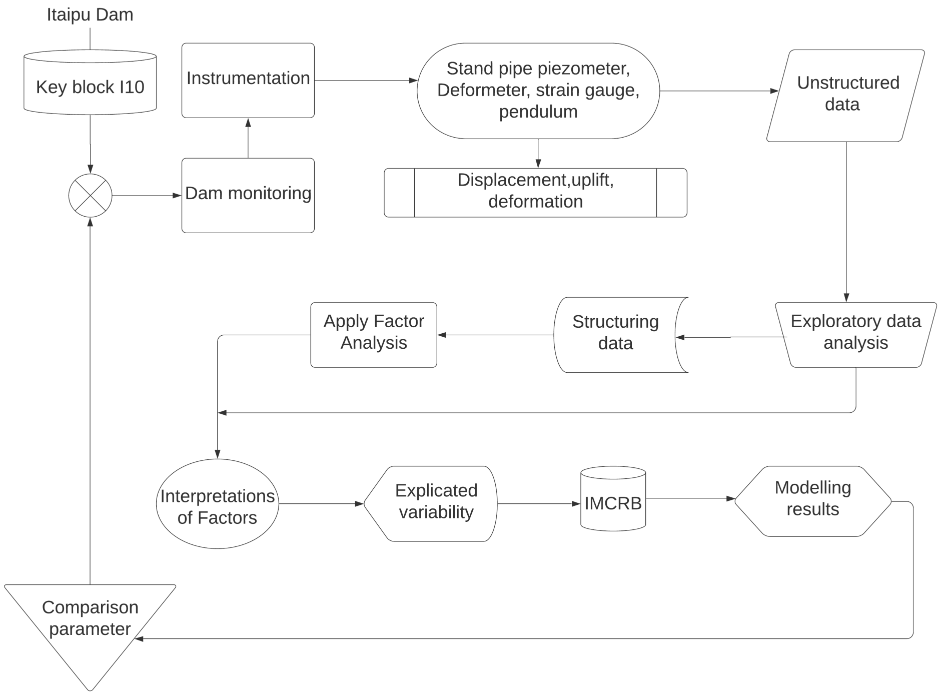

Figure 1. Flow chart research (Authors,2020)

Through factors identified in factor analysis, it was modelling IMCRB, in accordance with equation 1.

$$IMCRB = a_1F_1 + a_2F_2 + ... + a_kF_k \tag{1}$$

3 Factor analysis variables

The multivariable method of factor analysis has the intrinsic character the reduction of space dimension, through orthogonal model, that observable variable are represented by linear combination of latent variable (common factors), specific factors (random noisy) and a constant term (equation 2). This modelling it can be applied to generate indexes, to get insights, jointly general description of data variability.

$$X_j - \mu_j = \sum_{l=1}^{k} L_{jl}F_l + \Psi_l, j = 1, ..., p \tag{2}$$

L: Loadings factors; Ψ : Commom factors; F: Specific factors; k: Amount of factors; j: Amount of variables.

Factor analysis to perform a spectral decomposition of original observations in three components : Scores factors, communality, and loadings factors. The scores factors (equation 3) are normalize data, obtainted by each observation in correspond a factor in a weighting of observations values. [12].

$$\hat{f} = (\tilde{L}_z'\tilde{L}_z)^{-1}\tilde{L}_z'z_j \tag{3}$$

Communality (equation 4) is a statistical parameter that quantify a quality of factorial model in represents each observed variable, with values present in the interval zero and one, closer than one, better representations. . h_i^2 : Communality, Ψ_i : Specific variance. [13]

$$h_i^2 = 1 - \Psi_i \tag{4}$$

Loadings factors ($l_{jk} = cov(Z_jF_k)$), (Z : normalize data) explain the correlation between each variable in function to factor .[12]

The analysis of sensibility of statistical index, it was quantified by upset in the final observation of IMCRB times series, by means of simulation in variability instruments classified in communality more than 0.9. Eight instruments satisfy this condition : six strain gauge and two piezometers. It was generated 12 normal random number, at the level 1,2,3,4,5 standard deviation parameter in relation on mean, for each instrument. The remaining variables was the method of time series forecasting Seasonal Naïve [14], which repeat the last twelve observation (month). The results it was analysing by an exponentially weighted moving average control chart, using damping parameter optimized ($\lambda = 0.4$) at three standard deviation parameters.

3.1 Exponentially weighted moving average control chart

Exponential weighted moving average (EWMA) control chart it was utilized to detect small changes in mean, for this, to embrace present and past observation to predict if an observation is an occurrence in the control limits, in addition, because the purposes of studs is measure sensibility, control chart that detect small changes in mean is the best way to using. Considering the different type of methodologic for detect, the EWMA chart outstanding results to interpret IMCRB, jointly a description of floating control limits that represents piezometer data, particularly because has the trend component more intensified. To generate a cart, it is necessary two mainly parameter : (λ), account for sensibility of limits control, present in the range zero and one, closer than zero, better critical. Usual values are 0.2 at 0.4, but the chosen have studied uniqueness of the process to be controlled and unspoken knowledge. Parameter L, measure distance of interval at function of standard deviation, typical value 2.7; To control a process of dam monitoring, these parameters used are: de λ = 0.4 e L=3, com μ zero check data are normalized $\sigma = 0.57$ in the statistical programming environment RStudio [15].

$$Limits = \mu_0 \pm L\sigma\sqrt{\frac{\lambda}{2-\lambda}.(1-(1-\lambda)^{2i})} \tag{5}$$

4 Results

Before applying factor analysis, it was certificated the adequation criteria (table 2). KMO test, resulted in a value close to 1, indicated sample suitability. Bartlett test highlight to a greater than critical value ($\chi^2_{228} = 264.22$), appointing that instrument shows significant correlations. The multivariable normality was verification by a QQ plot (figure 2), condition necessary to evaluate loadings factors through maximum likelihood estimate. The value of Maximized log-likelihood value, results close to zero, shows excellent fit of factorial model. Horn's parallel analysis betoken three factors are appropriate to represents a set of twenty-five variables.

Table 2. Hallmark - Factor analysis

Parameter	Value	Method
KMO	0.93	Kaiser Meyer Olkin
χ^2_{calc}	3.19×10^3	Bartlett Sphericity Test
Maximized log-likelihood value	-9.16	Maximized log-likelihood
Amount of factors	3	Horn's Parallel Analysis

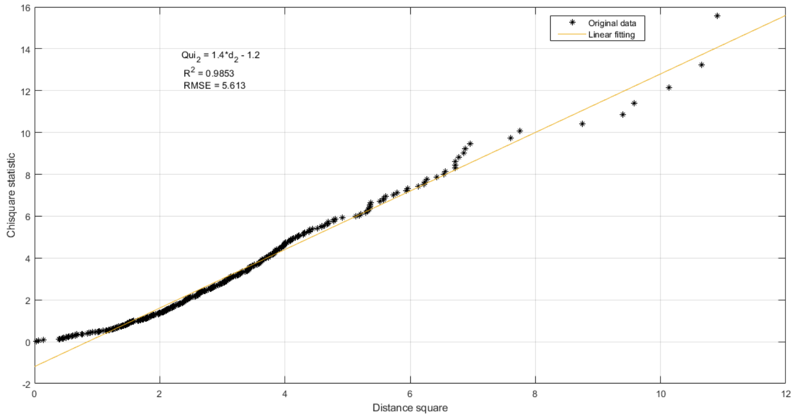

Figure 2. Multivariable normal

Ranking instruments in each factor, it was realized according to correlation value, reflected loadings factors. The first factor was dominated overwhelmingly for strain gauge in less contribution of piezometers. Second factor, through pendulums and deformeter in the meantime third factor, strain gauge and piezometer.

Analysing loadings factors, jointly scores factors plot (figure 3), the first factor is to present a ascend trend sequencially an inversion point stemming from extensometer. The characteristic of seasonal oscillation , at the second factor, reflecting pendulums measures. Third factor, implicit an ascendant trend, represents information coming from strain gauge and piezometer.

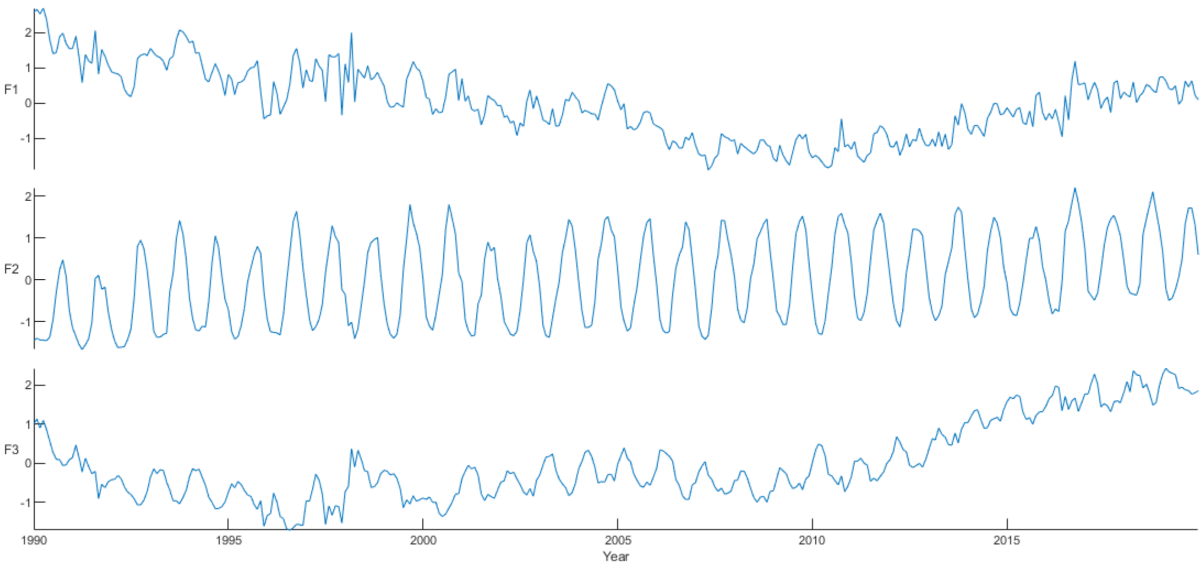

Figure 3. Factor scores

To modelling the IMCRB for key block I10 (equation 6), weighted scores factors in function to factors influence, whom, it was obtained of representation of each factors using the ration of eigenvalue greater than one and total amount of variables, sort in descend way. Obtained in this method, for each factor, respectively. The level of variability explication it is obtain by sum of each results, thus, factor model was successful compute thereabouts seventy-four percent of total variability of measure instruments, from 1990 to 2019.

$$IMCRB = 0.58F_1 + 0.10F_2 + 0.06F_3 \tag{6}$$

4.1 Sensibility - IMCRB

The instrumentation ranking it was realized using the communality value. Six strain gauge and two piezometers satisfy criteria of communality greater than 0.9, to be these instruments the biggest instruments represented by factor model. These instruments were selected to ascertain a sensibility of IMCRB, halfway through generate random number altering of variability to one to five level of changes of standard deviation in function mean. The results were plotted in control chart type exponentially weighted moving average (EWMA, figure 4), with parameter $\lambda = 0.4$ and control limits calculated at ninety-five of confidence. These plots are equivalent, has the difference only of final observation in times series. Each perturbation is identified by (x_n), in n : interference level.

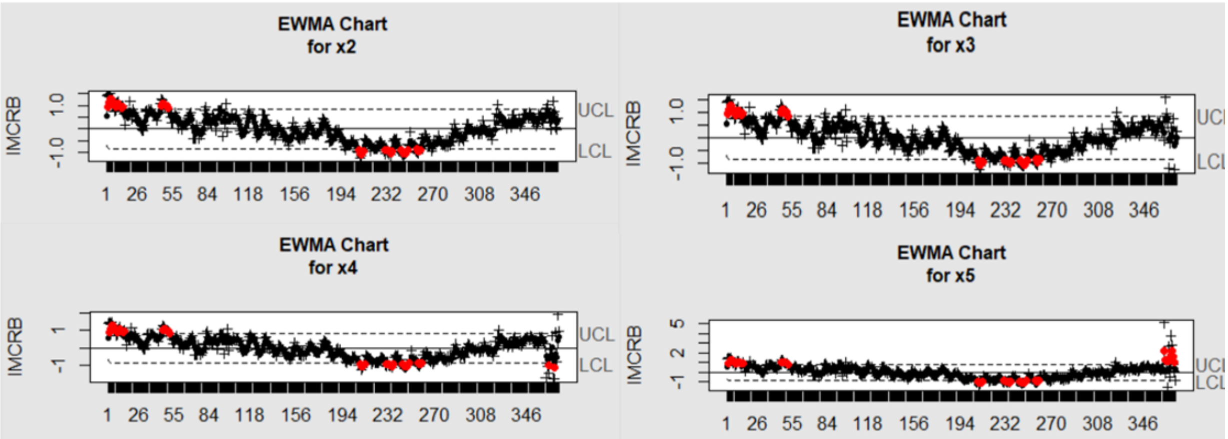

Figure 4. Control chart - EWMA

The figure 5, relate the total amount of occurs in function of level standard deviation (sd) in the generation of random number, in the others words, the first point, number of total extrapolation in one sd, sequentially, made until five sd, create four region: Stable region is composed through first three points, whom the perturbation in a one, two and three sd, were not enough to extrapolated limits control. Onset region, in three sd, the amount of occurs rise softly, indicated a growth exponential. Elevation region, Demonstrate the hypotheses of growth, increasing in two occurs. Last part, acceleration region, suppose that majority of future points, stay out of control limits.

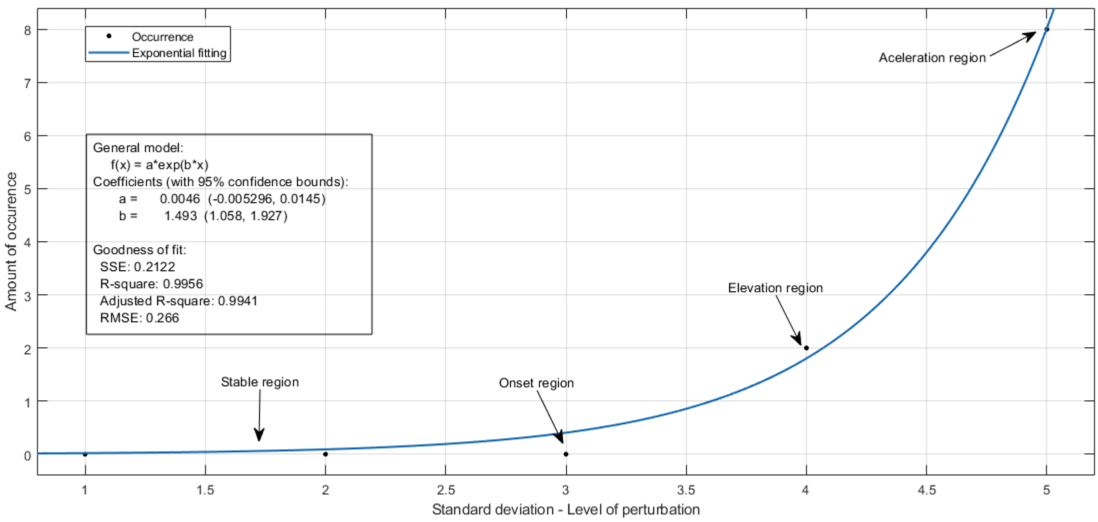

Figure 5. Sensibility

5 Conclusion

Joint Monitoring Index of Dam Blocks Response , building through factor analysis, resulting excellent performance in represents of instruments measure installed at the key block I10 of Itaipu dam. The three factors identified segregate strain gauge, pendulums, piezometers and deformeter.

The sensibility of IMCRB, measured by disturbance of greater variables represents by multivariable model, indicated latency between three initial perturbation, in exponential increase after three sd. Therefore, IMCRB display lower sensibility in minimal change in patter of readings, but on considerable alteration, detect each extrapolation. The low sensibility it is potently desired because little change, can be informed transitory and self-correcting situation that could normalize on the next seasonal period, or yet as response a correlated instrument. As of four standard deviation, suggest that start an investigation in the structural conditional of key block, measures, and instruments present in the first factor.

The low sensibility identified, shows that through punctual analysis of IMCRB at the exponential weighted moving average (EWMA) it is possible compare profile oscillations of measures, in addition, if there is trends, it is possible forecast the extrapolations limits and using the results as parameters to anticipate necessary action to return of process to stable situation.

Acknowledgements

Guilherme Lopes de Campos, Federal University technology – Paraná, scholarship holder of Itaipu technology Park – (FPTI-BR)

References

[1] Konakoglu, B., Cakir, L., & Yilmaz, V., 2020. Monitoring the deformation of a concrete dam: a case study on the deriner dam, artvin, turkey. *Geomatics, Natural Hazards and Risk*, vol. 11, n. 1, pp. 160–177.
[2] Hellgren, R., Malm, R., & Ansell, A., 2020. Performance of data-based models for early detection of damage in concrete dams. *Structure and Infrastructure Engineering*, vol. 0, n. 0, pp. 1–15.
[3] Li, B., Yang, J., & Hu, D., 2020. Dam monitoring data analysis methods: A literature review. *Structural Control and Health Monitoring*, vol. 27, n. 3, pp. e2501. e2501 STC-19-0265.R1.
[4] Silveira, J. F. A., 2003. *Instrumentação e Comportamento de Fundações de Barragens de Concreto*. Editora Oficina de Textos.
[5] Oro, R. S., 2016. *Índice de monitoramento do comportamento estrutural dos blocos de concreto de barragens - uma abordagem multivariada*. PhD thesis, Universidade Federal do Paraná, Curitiba, Brazil.
[6] , 2018. *Dam World - Third International Dam World Conference*, volume 1 of *Dam Monitoring and Instrumentation - Safety Assessment*, Brazil, Foz do iguaçu. Ibracon.
[7] Wang, S., Xu, Y., Gu, C., Bao, T., Xia, Q., & Hu, K., 2019. Hysteretic effect considered monitoring model for interpreting abnormal deformation behavior of arch dams: A case study. *Structural Control and Health Monitoring*, vol. 26, n. 10, pp. e2417. e2417 STC-18-0338.R1.
[8] Collado, C., Sampieri, R., & Lucio, P., 2013. *Metodologia de pesquisa*. Penso.
[9] Martinez, W., Martinez, A., & Solka, J., 2017. *Exploratory Data Analysis with MATLAB*. Chapman & Hall/CRC Computer Science & Data Analysis. CRC Press.
[10] Kang, F., Li, J., Zhao, S., & Wang, Y., 2019. Structural health monitoring of concrete dams using long-term air temperature for thermal effect simulation. *Engineering Structures*, vol. 180, pp. 642 – 653.
[11] Mata, J., Tavares de Castro, A., & Sá da Costa, J., 2014. Constructing statistical models for arch dam deformation. *Structural Control and Health Monitoring*, vol. 21, n. 3, pp. 423–437.
[12] Mingoti, S., 2005. *Análise de dados através de métodos de estatística multivariada: uma abordagem aplicada*. Editora UFMG.
[13] Härdle, W. & Simar, L., 2019. *Applied Multivariate Statistical Analysis*. Springer International Publishing.
[14] Kotu, V. & Deshpande, B., 2018. *Data Science: Concepts and Practice*. Elsevier Science.
[15] Montgomery, D., 2019. *Introduction to Statistical Quality Control*. EMEA edition. John Wiley & Sons Incorporated.