Jurnal Teknologi

Full Paper

THEORETICAL AND EXPERIMENTAL INVESTIGATION OF ESTIMATING CHANGE POINT IN MULTIVARIATE PROCESSES VIA SIMULTANEOUS COVARIANCE MATRIX AND MEAN VECTOR

Alireza Firouzi[a]*, Noordin Mohd Yusof[a], Muhammad Hisyam Lee[b], Robabeh Bashiri[c]

[a]School of Mechanical Engineering, Faculty of Engineering,
[b]Department of Mathematical Sciences, Faculty of Science, Universiti Teknologi Malaysia, 81310 UTM Johor Bahru, Johor, Malaysia
[c]Department of Environmental Technology, Faculty of Chemistry, University of Gdańskul, Wita Stwosza 63, 80-308 Gdańsk, Poland

Article history
Received
30 July 2021
Received in revised form
23 September 2021
Accepted
7 October 2021
Published Online
20 December 2021

*Corresponding author
falireza3@graduate.utm.my

Abstract

The identification of change points in statistical process control (SPC) data is the critical criterion for multivariate techniques when output is out-of-control condition. Therefore, monitoring all independent variables is essential and demands targeted attention to avoid errors at the systems control stage. However, estimating change-point in multivariate control charts is the main problem when these correlated quality characteristics monitor together. Therefore, we proposed a combination of an ensemble learning-based model of artificial neural networks with support vector machines to monitor process mean vector and covariance matrix shifts simultaneously to estimate the change point in a multivariable system. The performance of the final model indicated an estimated changing point with one sample over 6,000 simulated cases with a probability of 98 percent, which is a significantly high accuracy rating. Finding suggests the outcome of the project confirms that the proposed model can provide a precise estimating the change point by monitoring the mean vector and the covariance matrix simultaneously and, helps to identify those variable(s) responsible for an out-of-control condition. For further validation of the model, the performance of the proposed model has been compared with previous reported which confirms a better performance of the proposed model. Finally, the model was applied to monitor the performance of the solar hydrogen production system and the model identify the variables which have negative effects on the performance of the system.

Keywords: Multivariate normal process, simultaneous covariance matrix and mean vector, artificial neural networks (ANN), support vector machine (SVM), change point

Abstrak

Pengenalpastian titik perubahan dalam data kawalan proses statistik (SPC) adalah kriteria kritikal untuk teknik multivariate ketika output berada di luar kawalan. Oleh itu, memantau semua pemboleh ubah bebas adalah penting dan menuntut perhatian yang tepat untuk mengelakkan kesilapan pada tahap kawalan sistem. Walau bagaimanapun, mengira titik perubahan dalam carta kawalan multivariat adalah masalah utama apabila ciri-ciri kualiti berkorelasi ini dipantau bersama. Oleh itu, kami mencadangkan gabungan model berasaskan rangkaian saraf tiruan berasaskan pembelajaran ensemble dengan mesin vektor sokongan untuk memantau proses vektor dan perubahan matriks kovarians secara serentak untuk menganggarkan titik perubahan dalam sistem berbilang variabel. Prestasi model akhir menunjukkan titik perubahan yang dianggarkan dengan satu sampel lebih dari 6.000 kes simulasi dengan kebarangkalian 98 peratus, yang merupakan penilaian ketepatan yang sangat tinggi. Penemuan menunjukkan hasil projek mengesahkan bahawa model yang dicadangkan dapat memberikan anggaran tepat titik perubahan dengan memantau vektor min dan matriks kovarians

secara serentak dan, membantu mengenal pasti pemboleh ubah yang bertanggungjawab untuk keadaan di luar kawalan. Untuk pengesahan lebih lanjut model, prestasi model yang dicadangkan telah dibandingkan dengan laporan sebelumnya yang mengesahkan prestasi model yang dicadangkan lebih baik. Akhirnya, model ini digunakan untuk memantau prestasi sistem pengeluaran hidrogen suria dan model mengenal pasti pemboleh ubah yang mempunyai kesan negatif terhadap prestasi sistem tersebut.

Kata kunci: Proses normal multivariat, matriks kovarians serentak dan vektor min, rangkaian neural buatan (ANN), mesin vektor sokongan (SVM), titik perubahan

1.0 INTRODUCTION

In recent years competitive market and industry, the quality of a product or service is no longer measured by a single variable; however, several variables define the final product or service quality. It has known that these quality variables of products or services have correlated with each other, and it is essential to monitor them simultaneously. One of the main challenges in deploying multivariate control charts is identifying which elements are responsible for the control charts' out-of-control signal and identifying delay time as named change-point [1-4].

In the last decades, statistical process control (SPC) charts were the most popular tools to monitor the stability and variability of the industrial application [5-8]. The SPC charts have been utilized to identify either method is statistically under or out of control condition; however, the presence of autocorrelation and a specific pattern in the data cannot provide the possibility of correctly, quickly detecting, and classifying the existing fault [9-11]. It is crucial a quick detection of these shifts, and their causes for promoting required action at earlier stage of production [12, 13]. Many researchers proposed alternative monitoring processes like integrating SPC with ANNs to solve the SPC method's limitation. ANNs are enormously parallel computational systems that simulate a human brain. It has been reported that ANNs showed acceptable performance for a wide range of applications [14-16]. Furthermore, the highly reliable ANNs results provide a new platform for SPC during the last decade [17]. Applying ANNs in the detection of mean and/or variance shifts in the process assists in the interpretation of automating SPC plots [18]. Therefore,researchers have paid excellent attention to the ANNs to determine the change point by varying the mean vector because of its quite satisfying efficiency compared to other techniques. For instance, Ahmadzadeh *et al*. [19] add the sentence related their work Amiri *et al*. [20] write a new sentence about Amiri work. have utilized an exponentially weighted moving average (EWMA) control chart to show out-of-control condition which was integrated with the supervised learning method to estimate the step-change point in the mean vector of a multivariate normal process. In another work, Atashgar [21] displayed a supervised learning ANN to detect the change point with a linear trend in the mean vector of the bivariate neural network. The primary outcome of these results showed that the modified ANNs could identify the step point change in the mean vector, recognize out-of-control situations, and the variable or variables that contributed in the changes. This strategy is also able to measure change point and deviation variables. However, all reported works only identify the change point with the mean vector without considering covariance matrix changes [14, 18, 22].

Meanwhile, other researchers looked at changing points in the multivariate Covariance Matrix (MCM) using the ANN algorithm method [19]. Control charts based on the sample covariance matrix, such as can presents only shifts that change the determinant's value. Thus, considerable work was dedicated to reviewing change points for the multivariate mean; however, detecting structural in the covariance matrix has not been studied in the literature.

Aue *et al*. [20] developed nonparametric point of change estimates based on the well-known CUSUM method for a fixed dimension. Dette, H. and D. Wied [21] had proposed a general approach to identify essential change points in a time series parameter. Furthermore, Kao *et al*. [22] show that big dimensional stability tests of the significant covariance matrix, extreme size distortions result. On the other hand, Firouzi *et al*. [23] performed the first-ever inspection of control conditions using the MEWMS (multivariate exponentially weighted mean Square) construct, followed by the ANN algorithm, for appraising change points. Our team reported integrating the covariance matrix with the ANN algorithm without considering the mean vector. However, all stated models have not affected a practical approach. It is essential to consider changing the covariance matrix on the mean-vector in the multivariate process. A successful monitoring program requires monitoring both mean vector and covariance matrix shifts, the importance of simultaneously monitoring process mean and variability has been increased [24].

This paper proposes a novel approach to estimate the change point's precise location by considering the

covariance matrix and mean vector simultaneity. Firstly, the MEWMS and MEWMA(multivariate exponentially weighted moving average control charts are utilized to identify the control or occ situation and compare it together. when each control chart has been shown a faster signal occ condition, then we can start estimating the change point and also investigate the assignable cause(s). Secondly, the ANN(fitting)-SVM(classifier) method is applied to determine the status of the change point. and finally with one illustrated example, can be the accuracy of this method.

2.0 METHODOLOGY

We present the machine learning algorithm procedure as a preliminary study on process simulation, a change-point estimation model, a support vector machine, classification for the cause of out-of-control conditions, an estimator algorithm, and the performance appraisal of a merged ANN method with SVM.

2.1. Algorithm Procedure

In general terms, a system is identified by its parameters, including $Mean\ (\mathrm{M}) = [\mu_i]$ and Covariance $(\Sigma) = [\sigma_{ij}]$. The correlation matrix is defined by equation (1):

$$r_{jk} = \frac{\Sigma_{1=1}^{n}(X_{ij}-\bar{X}_j)\times(X_{ik}-\bar{X}_k)}{\sqrt{\Sigma_{1=1}^{n}(X_{ij}-\bar{X}_j)^2}\times\sqrt{\Sigma_{1=1}^{n}(X_{ik}-\bar{X}_k)^2}} \quad (1)$$

Here, it is assumed that the correlation matrix is constant during the process control. When the correlation matrix is constant, then there are (2*n*) independent variables, n mean *(M)* and *n* sigma *(∑)* values in the system.

If in control, statistics are expressed by zero indices, then a change in μ and Σ causes a change in all statistics elements since the parameters are correlated. Equations (2) and (3) help us make a new distribution, utilizing simulation where σ is a vector of change in standard deviations and δ is a standard deviation shift.

$$M = M_0 + \Delta M \times P \quad (2)$$

$$\Sigma = \begin{bmatrix} \delta_1^2\sigma_1^2 & \rho\delta_1\delta_2\sigma_1\sigma_2 \\ \rho\delta_1\delta_2\sigma_1\sigma_2 & \delta_2^2\sigma_2^2 \end{bmatrix} \quad (3)$$

The assumptions of the simulation are: (i) out of control does not alter the probability distribution of the variables, and (ii) variables follow a multivariate normal distribution. The new statistics were calculated using the formula mentioned earlier to simulate new random numbers when a change in statistics happens. The multivariate standard distribution as seen with Equation (4) is as follows:

$$N(M,\Sigma) = (2\pi)^{-\frac{n}{2}}|\Sigma|^{-\frac{1}{2}}\, e^{(X-M)^T\Sigma^{-1}(x-M)} \quad (4)$$

In mathematical form: $X_0 \sim N(M_0, \Sigma_0)$ and after the change: $X \sim N(M,\Sigma)$.These are (2*n)* independent variables in which their variations change the distribution parameters mean and *n* standard deviation changes at constant correlation assumption to determine the changes' scenarios. Therefore, there are 2^{2n}-1 scenarios of changes, excluding the no-change procedure.

2.2. Preliminary Study on Process Simulation

The MATLAB multivariate normal random generator was used to simulate the process. This function helped yield random numbers with a multivariate normal distribution, with a given M and σ. A set of random numbers was developed, and the mean and the covariance matrix are evaluated and recorded for each sample size. There were two sections for the simulation algorithm: (1) simulating an under-control system and (2) simulating the system after changing standard deviation or means value. The first part had a run the length of 50 samples, and the only simulation run for the process remained under control, and the data was kept final calculations. We explored these sample sizes, either process is in-control or out of control condition, by two statistics of MEWMA and MEWMS [23]. MEWMS has monitored variability in the multivariate process. Let $X_1, X_2, \ldots, X_\tau, X_{\tau+1}, \ldots, X_T$ are independent vectors from observations in which X is a normal distribution with p variables. the control chart of MEWMS discovered the occurred shift in the mean and Variability the control statistic in equations (5) and (6) of the chart as follow see Firouzi *et al*. (2020):

$$Y_{ij} = \Sigma_0^{-\frac{1}{2}}(X_{ij} - \mu_0) \quad (5)$$

$$S_t = (1-\lambda)S_{t-1} + \frac{\lambda}{n}\Sigma_{j=1}^{n} Y_{ij}Y_{ij}' \quad (6)$$

After the change is triggered in the second step, a new random of 100 samples are generated, and control chart statistics are evaluated.

2.3. Change Point Estimation Model

For the main algorithm, 2^{2n}-1 scenarios for each 400 sample are generated and followed by determining control chart statistics, mean vector, and covariance matrix. For each sample runs, two sets of machine learning inputs (I_1 and I_2) were chosen. I_1 inputs entailed the vector of mean and covariance matrix. Meanwhile, I_2 inputs comprised entirely of the vector

of control chart statistics, as shown in Equations (7) and (8) below:

$$I_1 = \left[M_{1:100}^{T} \ \left(\Sigma_{ij}\right)_{(1:100)}\right] \quad (7)$$

$$I_2 = \left[T_{MEWMA_{1:100}} \ T_{MEWMS_{1:100}}\right] \quad (8)$$

2.4. Support Vector Machine (SVM)

Supervised Learning is the most common paradigm for performing the machine learning (ML) process. It has widely used for data where there is an accurate mapping between input-output data. The Supervised Learning algorithm identifies the features explicitly and carries out predictions or classification accordingly [25]. As the training period progresses, the algorithm can identify the relationships between the two variables to predict a new outcome. It shows that supervised learning algorithms are task oriented. SVM is a supervised learning algorithm that analyses data for classification and regression analysis [26]. A supervised learning algorithm that consists of the Gaussian process (GP) can predict the value of an unseen point from training data by employing 'lazy learning' and measuring the similarity between those points in question (the kernel function) [27, 28]. Therefore, Equation (9) leans on the provisions of the Gaussian construct for representation regarding the probability density function of a normally distributed random variable with the expected value μ and variance σ^2 set as follow:

$$g(x) = \frac{1}{\sigma\sqrt{2\pi}} \exp\left(-\frac{(x-\mu)^2}{2\times\sigma^2}\right) \quad (9)$$

2.5. Classification for the Cause of Out of Control

A classifier model is required to clarify whether the cause is a change in mean vector only, standard deviation vector, or both mean and standard deviation. Because a control chart sometimes fails to prompt an out-of-control (OOC) signal. Therefore, there is not a record for length of change, which a program should chart instead of landing at a finite length of value. The cause of out-of-control must be evaluated at the initial stage to diminish the condition's effects on the predictor's accuracy.
.. , n... Here, we applied different sets of classifiers to SVM data to achieve the highest level of accuracy with all cases. Two inputs (I1 and I2) were designed to classify the cause. To create an accurate model, we set the classifier with several approaches, including whether a set of input relates to (1) a mean (M) changes only or (2) standard deviation (σ) changes only or (3) change in both parameters. Table 1 shows the overall strategy, including the input of symbols and targets for each approach.

Table 1 The details of classifiers for each cause of OOC

Symbol	Input variable	Target
A	I_2	Whether the cause only changes in M
B	I_2	Whether the cause only changes in σ vector
C	I_2	Whether the cause changes in both
D	I_1	Whether the cause only changes in M
E	I_1	Whether the cause only changes in σ vector
F	I_1	Whether the cause changes in both

I: input, M: mean, and σ: sigma

2.6. Design Estimator Algorithem

The MATLAB classification toolbox is applied to examine all existing classifiers to indicate the best classification model. Upon designing the models for all three classifiers, finding fitters for each designated classifier is necessary. Therefore, ANN is utilized per the rules Levenberg and Bayesian. Furthermore, the number of hidden layers is varied to identify the effect on the accuracy of output. The number of one and three layers are for the Bayesian rule.Meanwhile, three and five hidden layers are for the Levenberg rule. The R square (R^2) for test and train is measured to evaluate the accuracy. Training (classifiers and fitters) and testing are applied for 85% and 15% of all models' data, respectively.

After developing the models, the distance or length (L) between the first out-of-control point on the chart and the change trigger point were recorded for each simulated sample run. Length of change on MEWMS and MEWMA control charts were labelled L_S and L_M, respectively. The following algorithm is used for estimating the size as follow:

- Construct the input I_2 based on the statistics of the control charts.
- Using the M only classifier (SVM), monitor if the cause of change in M only.
- If the cause of change is M only, calculate the L using the corresponding fitter.
- If the cause is not M only, detect with σ only classifier (SVM) if the reason is σ only.
- If the cause is σ only, calculate L using the corresponding fitter.
- Using both case classifiers (SVM), detect if the Figure 1: Algorithm of the estimator
- change is in both parameters.
- If the change is in both parameters, using the corresponding fitter, estimate L.

If none of the cases above is detected, then calculate the length, as shown in Equation (10):

$$L = \frac{L_{M\ only} + L_{\sigma\ only} + L_{both}}{3} \quad (10)$$

The above procedures are visualized in Figure 1. The simulated data is applied for the testing procedure; however, actual data can be utilized from a process

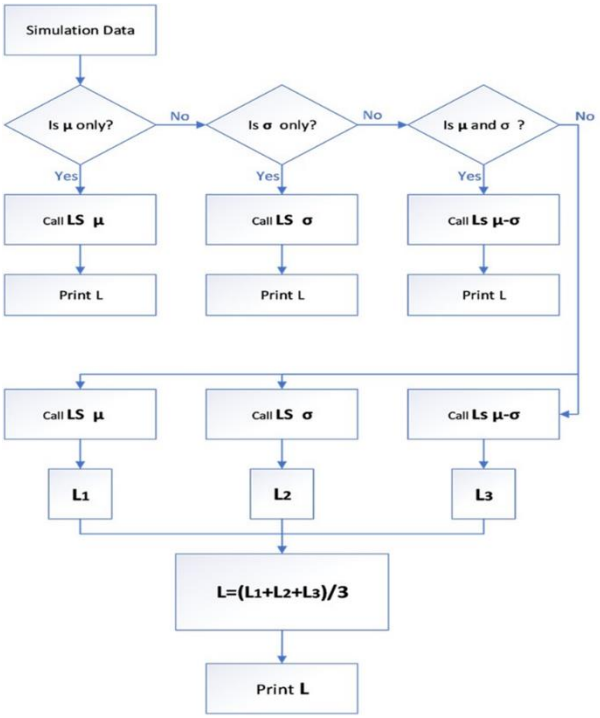

Figure 1 Algorithm of the estimator

2.7 Evaluating the Estimation Performance

Upon completing the estimator model, its performance should be appraised, as illustrated in Figure 2. Hence, a large sample set is simulated, which is different from the training sample set, and the length of change is estimated using the developed model. Then, the error is assessed, and from the estimated error, the probability density function (PDF) of the error is calculated. After calculating the error PDF, the confidence interval for the error is estimated.

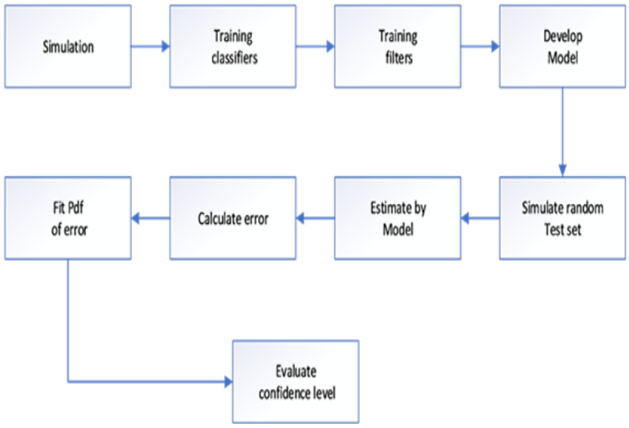

Figure 2 Algorithm of the estimator

3.0 RESULTS AND DISCUSSION

3.1 Analysis of the Outcomes of SVM Gaussian

Multiple models were used to locate the most suitable classifier with two sample sizes of 100 and 400 for 15 scenarios, as illustrated in Figures 3 and 4. Notably, increasing the sample size affected the performance; however, the SVM Gaussian with medium size revealed robust performance in both cases.

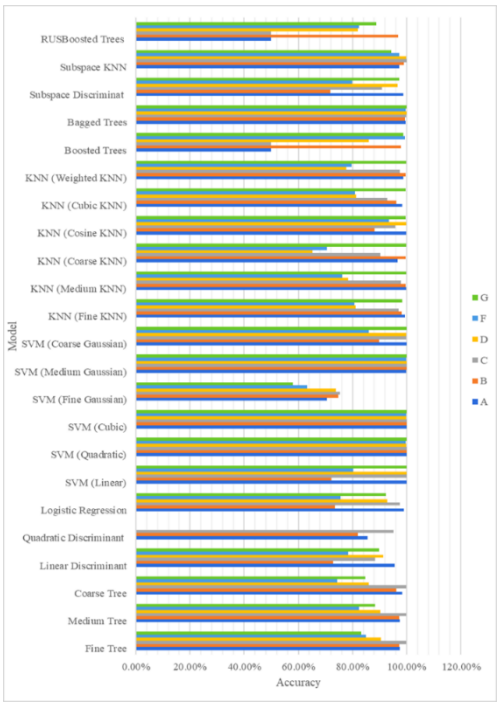

Figure 3 Accuracy of different classifiers (refer to Table 1) with a sample size of 100 for each scenario

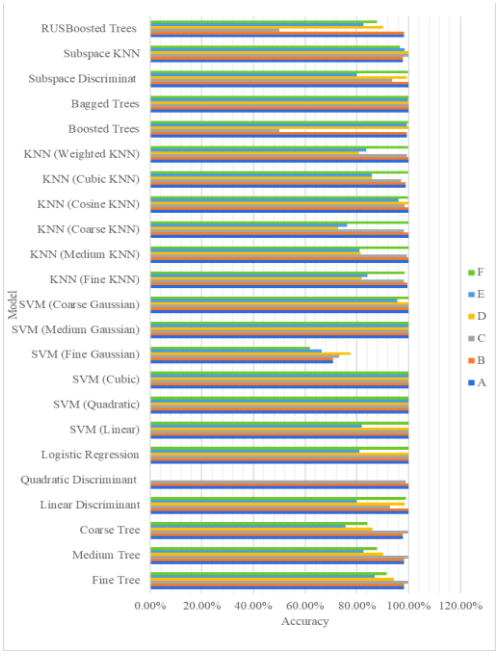

Figure 4 Accuracy of different classifiers (refer to Table 1) for the sample size of 400 for each scenario

The outcome of Figures 3 confirm that SVM medium Gaussian shows the best accuracy, and thus it was applied in the models for what. Afterward, the cause of the change was identified by modelling and adopting three classifiers. These charts revealed that the I_2 input has higher accuracy and is computationally less expensive for classification than I_1. Therefore, it stands that the MEWMS control chart charts were accurately modelled as a part of the relationship shared dimension reduction [18].

For the next step, we applied the best classifier to determine the best regression model and settings. The multiple models were developed using an ANN fitter with different settings, with a similar procedure used for the classifier. We compared Levenberg-Marquardt , Bayesian rules and the number of hidden layerstogether R^2 test, R^2 train, and R^2 validation are shown in Tables 2 and 3 for Levenberg and Bayesian rules, respectively. The results confirm that a lower number of hidden layers show better performance on the R^2 test, signaling possible overfitting in a high number of hidden layers [23].

Table 2 R^2 results for ANN model with Levenberg rules

Model No.	Input	Target	Cause of out of control	R^2_{train}	$R^2_{validation}$	R^2_{test}
The number of hidden layers: 3						
1	I_1	L_S	Change in M only	90.04	87.04	88.70
2	I_1	L_M	Change in M only	94.33	81.16	80.90
3	I_1	L_S	Change in σ only	90.25	54.30	47.60
4	I_1	L_S	Change in both	86.28	63.63	64.69
5	I_1	L_M	Change in both	86.05	80.00	80.01
6	I_2	L_M	Change in M only	96.93	93.42	93.38
7	I_2	L_S	Change in M only	94.96	90.06	90.5
8	I_2	L_S	Change in σ only	92.48	87.49	86.88
9	I_2	L_M	Change in both	95.45	92.87	93.03
10	I_2	L_S	Change in both	96.85	92.60	94.10
The number of hidden layers: 5						
1	I_1	L_S	Change in M only	86.01	52.65	53.27
2	I_1	L_M	Change in M only	90.40	74.23	72.95
3	I_1	L_S	Change in σ only	90.90	42.91	34.51
4	I_1	L_S	Change in both	81.40	58.19	54.17
5	I_1	L_M	Change in both	87.00	78.48	78.71
6	I_2	L_M	Change in M only	96.87	94.18	94.14
7	I_2	L_S	Change in M only	96.03	85.80	84.78
8	I_2	L_S	Change in σ only	92.60	85.95	84.56
9	I_2	L_M	Change in both	95.03	93.93	92.73
10	I_2	L_S	Change in both	97.39	94.25	94.03

I: input, M: mean, and σ: sigma

Table 3 R^2 results for ANN model with Bayesian rules

Model No.	Input	Target	Cause of out of control	R^2_{train}	$R^2_{validation}$	R^2_{test}
The number of hidden layers: 1						
1	I_1	L_S	Change in M only	83.11	0	66.52
2	I_1	L_M	Change in M only	92.45	0	86.82
3	I_1	L_S	Change in σ only	80.58	0	56.00
4	I_1	L_S	Change in both	91.70	0	61.8
5	I_1	L_M	Change in both	85.61	0	81.77
6	I_2	L_M	Change in M only	88.29	0	87.93
7	I_2	L_S	Change in M only	95.45	0	93.13
8	I_2	L_S	Change in σ only	90.14	0	92.44
9	I_2	L_M	Change in both	92.25	0	90.98
10	I_2	L_S	Change in both	97.72	0	92.36
The number of hidden layers: 3						
1	I_1	L_S	Change in M only	96.85	0	57.05
2	I_1	L_M	Change in M only	93.75	0	80.63
3	I_1	L_S	Change in σ only	96.84	0	38.76
4	I_1	L_S	Change in both	93.95	0	36.23
5	I_1	L_M	Change in both	95.02	0	87.00
6	I_2	L_M	Change in M only	94.78	0	90.32
7	I_2	L_S	Change in M only	97.81	0	93.91
8	I_2	L_S	Change in σ only	100.0	0	42.98
9	I_2	L_M	Change in both	100.0	0	60.01
10	I_2	L_S	Change in both	99.95	0	82.76

Figures 3 and 4 illustrate R^2 test for the different number of hidden layers based on Tables 2 and 3, respectively. For both Figures, models 1 to 5, corresponding to I_1, demonstrate lower accuracy than models 6 to 10, relating to I_2. Moreover, it can be observed that I_1 has a lower accuracy than I_2. Therefore, choosing I_2 input is more promising than I_1. Figure 3 and Table 2 represent three hidden layers with a higher R^2 test than five hidden layers in the Levenberg rules. Furthermore, five hidden layers modes also show R^2 train results, which are a sign of overfitting. Thus, three hidden layers are chosen because both the R^2 train and R^2 test are high; typically, it means the model is well trained with little chance of overfitting [29]. Figure 4 and Table 3 show the analysis of the data for the Bayesian model. In this model, there is no R^2 validation, and one hidden layer model shows much better values in terms of the R^2 test than the three hidden layers model. Figure 5 displays the compare R^2 test to 3 and 5 hidden layers for Levenberg rules. As we can see, 3 hidden layers are better than 5 hidden layers.

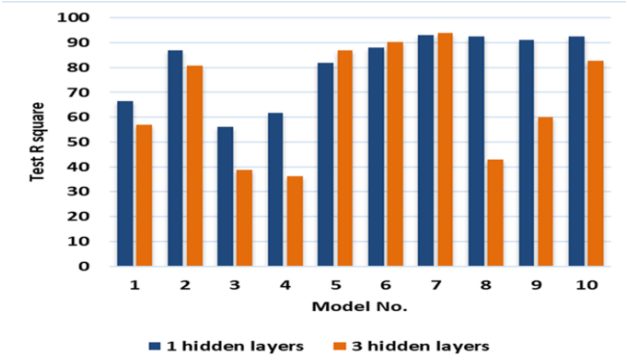

Figure 5 R^2 test for different modelsfrom Levenberg rule

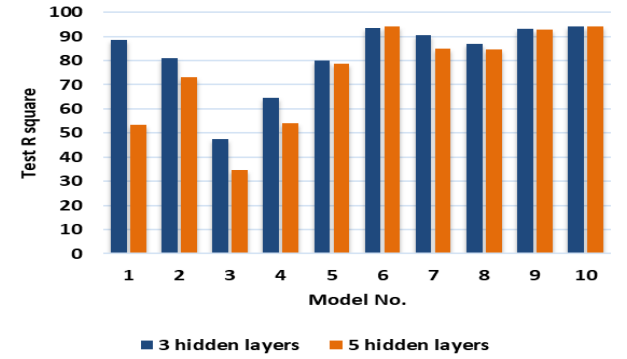

Figure 6 R^2 test for different models from Bayesian rules

Figure 6 reveals the compare R^2 test to 3 and 5 hidden layers for Bayesian rules as shown that one hidden layer is better than 3 layers.

The error histogram graph, extracting from Tables 2 and 3, are presented in supplementary information Figures 1S, 2S, and 3S. Error distribution for Levenberg models are flatter than Bayesian rules; however, both cases become sharper when hidden layers increase. Additionally, the mean change for each model only shows more extended distributions. The MEWMA chart is less sensitive to the changes, and the length of change is significantly larger, increasing or extending the possible error values over a larger domain[30]. Moreover, the training error is more sharply distributed over a small amount around the origin, and the test errors are more widely distributed with a lower R^2 test. The sign of overfitting is observed on higher hidden layers more than a lower number of hidden layers. The Levenberg rule models are more accurate than Bayesians and more stable when the number of layers changes. Therefore, selecting the Levenberg rule with three hidden layers for models' numbers 6 to 10 with better input is possible. R^2 test data, one may choose the best predictor for input I_2 for models' number 6 to 10 except for model number 8, which corresponds to change in standard deviation only. Therefore, a Bayesian model with one hidden layer is selected for model number 8. Table 4 and Figure 7 (a) to (c) show the models chosen for each estimator area.

Table 4 Selected fitting models for each cause of change

Model No.	Cause of out of control	Target	Number of layers	Rule
7	Change point if M changes only	L_s	3	Levenberg
8	Change point if σ changes only	L_s	1	Bayesian
10	Change point of both are changing	L_s	3	Levenberg

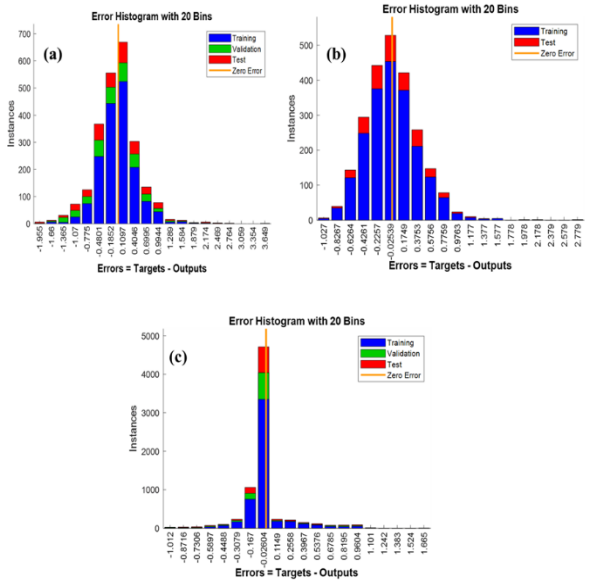

Figure 7 The selected error histogram for each estimator areas in where (a) run number 7, (b) run number 8, and (c) run number 10

3.2 Fitting PDF Error and Evaluate CI

It is worth mentioning that model number 6 with target L_M with higher R^2test was not chosen since L_S target shows better performance than L_M even for changes in M only. Two sets of error data are calculated to compare these two models. In the first set, the error is calculated when the fitter with L_m target is used for change in mean only, which is model number 6 in Table 2. The corresponding error and pdf for this case are shown with subscript 1 (set_1). The second set is the error data when model 7 is used, denoted with subscript 2 (set_2). Figure 8 shows the probability distribution for error for each case. One exciting fact is that, despite the higher R square, the error distribution for model 6 is flatter than 7, and the confidence interval for 7 is better than 6. The fact is that the R square shows the correlation, while the plot here is the error itself. MEWMA is less sensitive to the change than MEWMS, and the absolute value of error for MEWMA is higher than MEWMS. For this reason, all selected models include Ls unless the control chart MEWMS does not show any out of control.

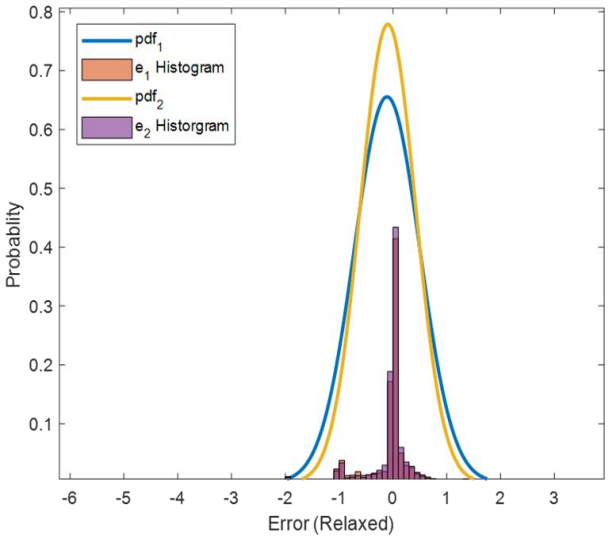

Figure 8 PDF and error histogram for the two error sets

The confidence intervals are plotted in Figure 9 and Table 5 for two sets based on the obtained probability distribution for error. Set_2 shows higher values for a reason discussed earlier. The Z is the value that in Equation (10) a follow:

$$P(|e| \leq Z) = \gamma \quad (10)$$

The graph shows that with the confidence of 98%, the estimated length of change is only one sample apart from the real change point, which is a significant accuracy.

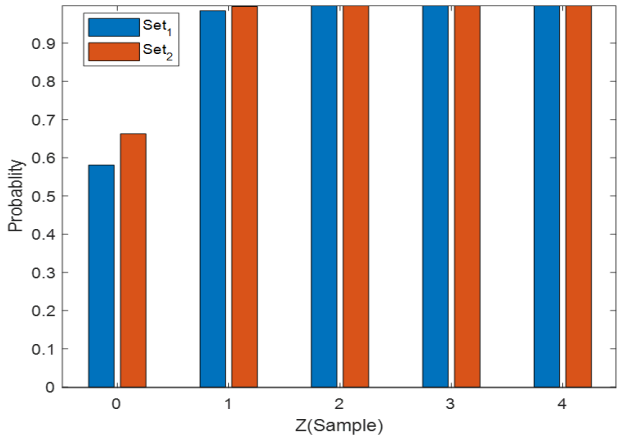

Figure 9 Confidence intervals (CI) corresponding to the discrete error

Table 5 Discrete values for confidence intervals

	$e = 0$	$\|e\| \leq 1$	$\|e\| \leq 2$	$\|e\| \leq 3$	$\|e\| \leq 4$
Set_1	0.5810	0.9847	0.9999	0.9999	0.9999
Set_2	0.6630	0.9960	0.9999	0.9999	0.9999

3.3 Validation of Proposed Model

The change point estimators were verified for a designed continuous solar hydrogen production system (Figure 4S) to control the electrolyte concentration and temperature on system performance.[31] Two critical quality characteristics are considered: temperature (X1) with a specification of $75^{oC} \pm 1$ and glycerol concentration (X2) with a specification of 5M ±1. To apply the model, we organized all hydrogen production data in 100 sample sets in which each sample contains ten points of data simulated for two parameters. The total accumulative data size is 2000 data points; thus, presenting all data is not possible, and only control charts statistics are shown for each sample in Table 6. Figures 10 and 11 show the comparison between proposed change point estimators for MEWMS and MEWMA control charts, respectively.

Table 6 Control Chart Statistics for all data

Sample	MEWMA	MEWMS	Sample	MEWMA	MEWMS	Sample	MEWMA	MEWMS
1.000	0.220	2.158	35.000	0.064	1.981	69.000	39.132	12.027
2.000	0.170	2.083	36.000	0.176	1.974	70.000	47.427	12.882
3.000	0.007	2.167	37.000	0.078	2.081	71.000	47.475	12.507
4.000	0.031	2.109	38.000	0.148	2.067	72.000	54.597	13.111
5.000	0.068	2.171	39.000	0.045	2.117	73.000	53.138	13.708
6.000	0.031	2.124	40.000	0.009	2.071	74.000	54.821	13.202
7.000	0.164	2.093	41.000	0.118	2.070	75.000	50.872	13.093
8.000	0.060	2.138	42.000	0.060	2.130	76.000	54.539	13.741
9.000	0.122	2.022	43.000	0.020	2.173	77.000	55.176	13.828
10.000	0.210	2.135	44.000	0.030	2.239	78.000	55.410	13.664
11.000	0.167	2.151	45.000	0.067	2.221	79.000	54.488	13.137
12.000	0.070	2.224	46.000	0.034	2.126	80.000	55.939	13.401
13.000	0.062	2.215	47.000	0.066	2.005	81.000	63.141	13.657
14.000	0.008	2.236	48.000	0.082	1.980	82.000	65.082	13.546
15.000	0.196	2.164	49.000	0.229	2.000	83.000	69.195	13.698
16.000	0.065	2.125	50.000	0.137	2.005	84.000	70.978	13.278
17.000	0.001	2.143	51.000	0.194	2.975	85.000	59.042	12.612
18.000	0.039	2.045	52.000	0.752	3.860	86.000	53.627	13.040
19.000	0.124	2.060	53.000	2.871	5.325	87.000	60.694	13.290
20.000	0.183	2.084	54.000	7.300	6.487	88.000	57.566	13.684
21.000	0.096	2.103	55.000	9.055	7.175	89.000	64.195	14.734
22.000	0.047	2.034	56.000	12.777	7.622	90.000	65.873	14.778
23.000	0.048	1.999	57.000	15.411	8.248	91.000	77.515	15.490
24.000	0.026	2.012	58.000	15.872	8.483	92.000	76.902	16.103
25.000	0.163	2.182	59.000	18.247	8.965	93.000	72.515	15.272
26.000	0.140	2.219	60.000	22.643	9.217	94.000	68.343	15.110
27.000	0.367	2.358	61.000	24.484	9.948	95.000	66.492	15.246
28.000	0.138	2.361	62.000	25.560	9.687	96.000	67.842	15.274
29.000	0.109	2.263	63.000	31.523	10.660	97.000	67.247	15.633
30.000	0.036	2.180	64.000	31.923	10.841	98.000	68.954	15.511
31.000	0.013	2.068	65.000	34.527	10.635	99.000	64.949	15.795
32.000	0.013	1.974	66.000	38.468	10.656	100.000	63.087	15.758
33.000	0.042	1.885	67.000	43.173	11.025			
34.000	0.021	1.951	68.000	42.789	11.708			

Both charts show out-of-control (OOC) conditions, upper control chart limit (UCL) control chart statistics, and OOC detected. The CPA is an actual change point at set sample 50, and CPE is an estimated change point by the model developed here. The error is less than one, so considering the integer value, the accuracy of ±1 sample here. In this case, the calculated value is precisely 50 if the bracket function applies to the result. In both charts, the distance between OOC and CPA is the length of change (actual). The results show that hydrogen production performance gradually dropped after increasing the temperature by more than 75^{oC,} while glycerol concentration negatively affected hydrogen production in Figure 10 shown changepoint in the MEWMS control chart in out-of-control conditions.

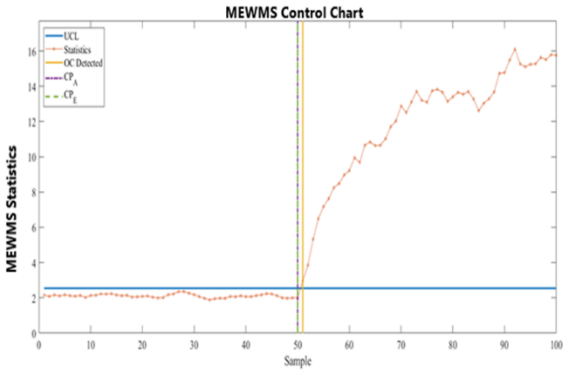

Figure 10 MEWMS Control Chart shows the estimate change point after out-of-control condition

Figure 11 presents change point estimation to consider using the MEWMA control charts. To compare about two control charts, MEWMS is better than MEWMA to control conditions.

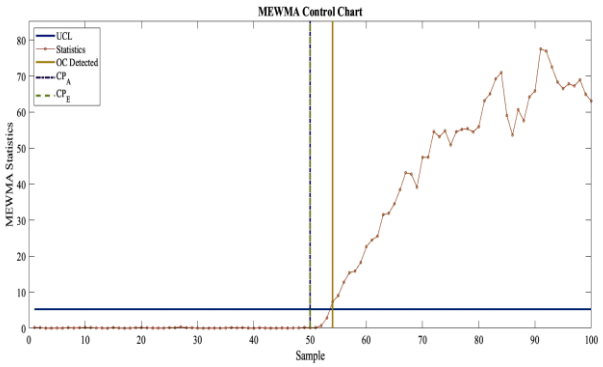

Figure 11 MEWMA control chart shows the estimated change point after out-of-control condition

3.4 Comparison with Another Estimator

Comparison performance of the method current study with proposed ANN-SVM compare with Ahmadzadeh *et al.* [32] is estimating mean change with using ANN method and used MEWMA control chart is provided in Table 7.

Table 7 comparison current study with Ahmadzadeh *et al.* [20] for p=2, $\Delta\mu_1$

		$e = 0$	$\lvert e\rvert \leq 1$	$\lvert e\rvert \leq 2$	$\lvert e\rvert \leq 3$	$\lvert e\rvert \leq 4$
My study	ANN-SVM	0.663	0.9960	0.9999	0.9999	0.9999
Ahmadzadeh *et al* [20]	ANN	6E-10	0.9875	0.9999	0.9999	0.9999

As it can be shown in Table 7, As we can be seen, the error of zero, to the estimator of this study 0.663 and for another estimator near 0 and error of one is 0.9960 better than 0.9875, because of their case estimator has a shift, while the estimator of this study is near zero, the accuracy of the method in this study is more precise. On the other hand, given that the other of the errors have equal probability, then, we conclude that our method is more accurate, because mean of error for our study has been zero, but the other study near 1.23.

4.0 CONCLUSION

In this study, the integration of Artificial Neural Networks (ANN) with Support Vector Machine (SVM) as one of the new methods of ML showed significant accuracy to estimate the change point. The consequences of this study have been shown the MEWMS is much more sensitive to change in statistics than the MEWMA chart. Also, the SVM medium gaussian classifier illustrates significant performance in classifying the cause of change. On the other hand, ANN with the Levenberg rule with l_2 provides better accuracy when mean changes are involved, but, ANN with the Bayesian rule with l_2 shows higher accuracy for changes in standard deviation only. Hence, The Ls statistics fitting model with a lower absolute error is more suitable for estimating the length. The model provided here can estimate the change point with one sample difference over 6000 tested cases (simulated) with a probability of 98%, which is an accurate and reliable model for a practical approach. The model has the potential for further study on other machine learning algorithms, which results in lower time and resource consumption is a considerable contribution the outcome of this project shows the combination of a theoretical and experimental method for solar hydrogen production is a step-stone toward practical application by monitoring parameters during the experiment and enhance the performance of the system. To comparison this study with Ahmadzadeh *et al.* method, the estimator for this study performs, better than the that estimator in terms of accuracy and validity.

Acknowledgment

The authors wish to acknowledge the school of engineering department of the Universiti Teknologi Malaysia (UTM) for Technical and financial support.

References

[1] J. H. Sullivan, W. H. Woodall. 2000. Change-point Detection of Mean Vector or Covariance Matrix Shifts Using Multivariate Individual Observations. *IIE Transactions*. 32: 537-549.

[2] S. T. A. Niaki, M. Khedmati. 2014. Monotonic Change-point Estimation of Multivariate Poisson Processes using a Multi-

Attribute Control Chart and MLE. *International Journal of Production Research.* 52: 2954-2982.
[3] K.-P. Lu, S.-T. Chang, M.-S. Yang. 2016. Change-point Detection for Shifts in Control Charts using Fuzzy Shift Change-point Algorithms. *Computers & Industrial Engineering*. 93: 12-27.
[4] B. Liu, C. Zhou, X. Zhang. 2019. A Tail Adaptive Approach for Change Point Detection. *Journal of Multivariate Analysis.* 169: 33-48.
[5] D. C. Montgomery. 2013. *Introduction to Statistical Quality Control.* New York, John Wiley & Sons.
[6] E. Schechtman, G. Bandner, S. Meginy. 2007. Detecting a Change in a Scale Parameter-A Combination of SPC and Change Point Procedures. *International Journal of Production Research*. 45: 5535-5545.
[7] J. Park, C.-H. Jun. 2015. A New Multivariate EWMA Control Chart Via Multiple Testing. *Journal of Process Control*. 26: 51-55.
[8] Y. Kwon, J.-H. Won, B. J. Kim, M. C. Paik. 2020. Uncertainty Quantification Using Bayesian Neural Networks in Classification: Application to Biomedical Image Segmentation. *Computational Statistics & Data Analysis.* 142: 106816.
[9] A. McCracken, S. Chakraborti, A. Mukherjee. 2013. Control charts for Simultaneous Monitoring of Unknown Mean and Variance of Normally Distributed Processes. *Journal of Quality Technology*. 45: 360-376.
[10] S. Hu, L. Zhao, Y. Yao, R. Dou. 2016. A Variance Change Point Estimation Method Based on Intelligent Ensemble Model for Quality Fluctuation Analysis. *International Journal of Production Research*. 54: 5783-5797.
[11] A. Batsidis, L. Horváth, N. Martín, L. Pardo, K. Zografos. 2013. Change-point Detection in Multinomial Data using Phi-Divergence Test Statistics. *Journal of Multivariate Analysis.* 118: 53-66.
[12] W. H. Woodall, D. C. Montgomery. 2014. Some Current Directions in the Theory and Application of Statistical Process Monitoring. *Journal of Quality Technology*. 46: 78-94.
[13] J. Huh. 2010. Detection of a Change Point based on Local-likelihood. *Journal of Multivariate Analysis.* 101: 1681-1700.
[14] F. A. P. Peres, F. S. Fogliatto. 2018. Variable Selection Methods in Multivariate Statistical Process Control: A Systematic Literature Review. *Computers & Industrial Engineering.* 115: 603-619.
[15] I. Sabuncuoglu. 1998. Scheduling with Neural Networks: A Review of the Literature and New Research Directtions. *Production Planning & Control.* 9: 2-12.
[16] S. Cheon, J. Kim. 2010. Multiple Change-point Detection of Multivariate Mean Vectors with the Bayesian Approach. *Computational Statistics & Data Analysis*. 54: 406-415.
[17] D. C. Schmidt, J. Haddock, S. Marchandon, G. C. Runger, W. A. Wallace, R. N. Wright. 1998. A Methodology for Formulating, Formalizing, Validating, and Evaluating a Real-time Process Control Advisor. *IIE transactions*. 30: 235-245.
[18] A. O. Memar, S. T. A. Niaki. 2011. Multivariate Variability Monitoring using EWMA Control Charts based on Squared Deviation of Observations from Target. *Quality and Reliability Engineering International*. 27: 1069-1086.
[19] S. Formentin, M. Mazzoleni, M. Scandella, F. Previdi. 2019. Nonlinear System Identification via Data Augmentation. *Systems & Control Letters*. 128: 56-63.
[20] A. Aue, S. Hörmann, L. Horváth, M. Reimherr. 2009. Break Detection in the Covariance Structure of Multivariate Time Series Models. *The Annals of Statistics.* 37: 4046-4087.
[21] H. Dette, D. Wied. 2014. Detecting Relevant Changes in Time Series Models. arXiv preprint arXiv:1403.8120.
[22] C. Kao, L. Trapani, G. Urga. 2018. Testing for Instability in Covariance Structures. *Bernoulli.* 24: 740-771.
[23] A. Firouzi, N. B. M. Yusof, M. H. Lee. 2020. Multivariate Change Point Estimation in Covariance Matrix Using ANN. *IOP Conference Series: Materials Science and Engineering.* IOP Publishing. 012101.
[24] E. Doğu, İ. D. Kocakoç. 2013. A Multivariate Change Point Detection Procedure for Monitoring Mean and Covariance Simultaneously. *Communications in Statistics-Simulation and Computation.* 42: 1235-1255.
[25] A. Tharwat. 2020. Behavioral Analysis of Support Vector Machine Classifier with Gaussian Kernel and Imbalanced Data. arXiv preprint arXiv:2007.05042.
[26] F. Lolli, E. Balugani, A. Ishizaka, R. Gamberini, B. Rimini, A. Regattieri. 2019. Machine Learning for Multi-criteria Inventory Classification Applied to Intermittent Demand. *Production Planning & Control*. 30: 76-89.
[27] O. Okwuashi, C. E. Ndehedehe. 2020. Deep Support Vector Machine for Hyperspectral Image Classification. *Pattern Recognition.* 107298.
[28] M. Ghosh, Y. Li, L. Zeng, Z. Zhang, Q. Zhou. 2020. Modeling Multivariate Profiles using Gaussian Process-controlled B-splines. *IISE Transactions*. 1-12.
[29] S. B. Rahimi, A. Amiri, R. Ghashghaei. 2019. Simultaneous Monitoring of Mean Vector and Covariance Matrix of Multivariate Simple Linear Profiles in the Presence of within Profile Autocorrelation. *Communications in Statistics-Simulation and Computation*. 1-18.
[30] S. Knoth. 2017. ARL Numerics for MEWMA Charts. *Journal of Quality Technology*. 49: 78-89.
[31] N. M. Mohamed, R. Bashiri, C. F. Kait, S. Sufian. 2018. Photocatalytic Water Splitting over Titania Supported Copper and Nickel Oxide in Photoelectrochemical Cell; Optimization of Photoconversion Efficiency. *IOP Conference Series: Materials Science and Engineering*. IOP Publishing. 012007.
[32] F. Ahmadzadeh. 2018. Change Point Detection with Multivariate Control Charts by Artificial Neural Network. *The International Journal of Advanced Manufacturing Technology*. 97: 3179-3190.

Appendix

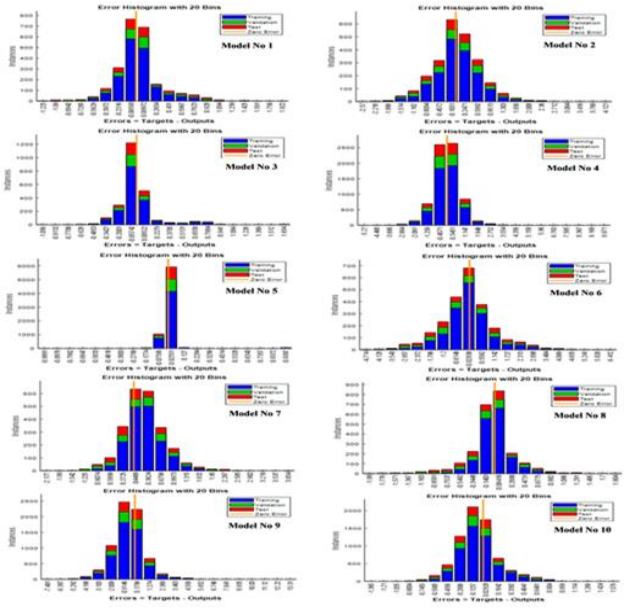

Figure 1S Error histogram for training models in Table 2 for Levenberg rule with 5 hidden layers

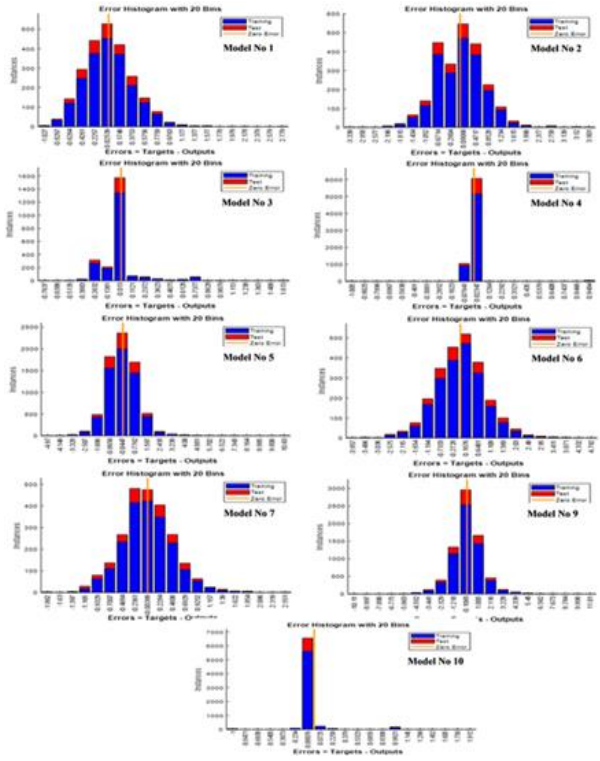

Figure 2S Error histogram for training models in Table 3 for Bayesian rule with 1 hidden layer (Model 8 is presented in the manuscript)

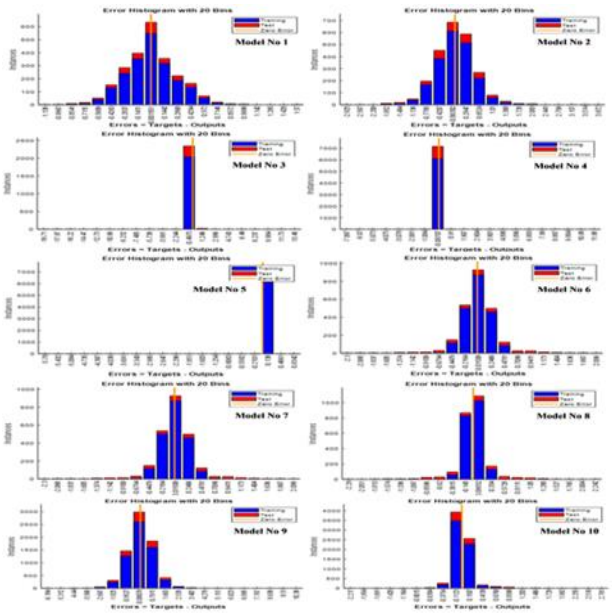

Figure 3S Error histogram for training models in Table 3 for Bayesian rule with 3 hidden layers

Overview of Continuous Solar Hydrogen Production Setup The Photoelectrochemical Layer-Integrated Cell with Nafion Separator (PeLICANS) is integrated with Dye-Sensitized Solar Cell (DSC). Each layer of the cell is purposefully designed to carry their respective functions within the cell. This 'layered' cell configuration is chosen as an innovative approach to conventional PEC cells, which are largely built around physical electrolyte reservoirs and have restrictions on future modifications. The green and seamless solution to this is integrating an innovative and standalone solar panel consisting of photoelectrochemical (PEC) cell with highly efficient bimetallic photocatalyst and dye solar cell (DSC) which operates well in diffused light into the façade of the building as walls. Applying light source on the surface of photoanode produces electron-hole pairs in the PEC cell. The excited electrons transfer The excited electrons can migrate to DSC and feed them back to the PEC cell's counter-electrode (Pt) without applying extra bias.

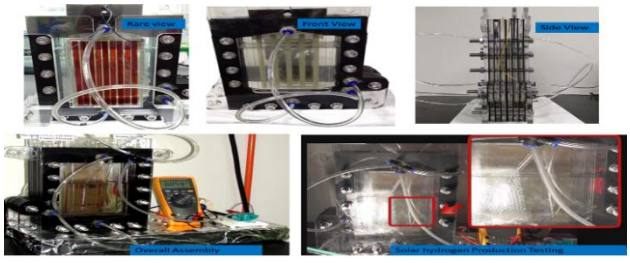

Figure 4S Different sections of continuous solar hydrogen production setup