IEEE Access
Multidisciplinary ⋮ Rapid Review ⋮ Open Access Journal

Received December 4, 2019, accepted January 27, 2020, date of publication February 3, 2020, date of current version February 14, 2020.

Digital Object Identifier 10.1109/ACCESS.2020.2971143

Multivariate Standard Addition Cobalt Electrochemistry Data Fusion for Determining Phosphate Concentration in Hydroponic Solution

VU NGOC TUAN[1,2,3], TRINH DINH DINH[4,5], ABDUL MATEEN KHATTAK[2,6], LIHUA ZHENG[1,2], XIAOFEI CHU[1,2], WANLIN GAO[1,2], AND MINJUAN WANG[1,2,7]

[1]Key Laboratory of Agricultural Informatization Standardization, Ministry of Agriculture and Rural Affairs, Beijing 100083, China
[2]College of Information and Electrical Engineering, China Agricultural University, Beijing 100083, China
[3]Faculty of Electrical and Electronic Engineering, Nam Dinh University of Technology Education, Nam Dinh 420000, Vietnam
[4]Quality Testing Lab, Center for Research and Development Science Technology Tien Nong, Thanh Hoa 442410, Vietnam
[5]College of Chemistry and Chemical Engineering, Beijing Institute of Technology, Beijing 102488, China
[6]Department of Horticulture, The University of Agriculture, Peshawar 25120, Pakistan
[7]Key Laboratory of Liquor Making Biological Technology and Application, Zigong 643000, China

Corresponding authors: Wanlin Gao (wanlin_cau@163.com) and Minjuan Wang (minjuan@cau.edu.cn)

This work was supported in part by the National Key Research and Development Program under Grant 2016YFD0200600-2016YFD0200602, and in part by the Liquor Making Biological Technology and Application of Key Laboratory of Sichuan Province under Grant NJ2019-02.

ABSTRACT Phosphate is one of the major elements affecting agricultural production. The accurate determination of phosphate concentration essential for plant growth, especially in a hydroponics system, allows regulating the balanced and suitable range set of nutrients to plants efficiently. This study proposed a data fusion model based on 70 samples for calibration and 30 samples for predicting concentrations of phosphate in an eggplant nutrient solution. Three multivariate analysis methods i.e. partial least squares model (PLS), Gaussian process regression (GPR), and artificial neural network (ANN) were studied and compared for their performance efficiencies. The results showed that combining the multivariate standard addition method (MSAM) in acquiring data from cobalt electrodes and ANN data fusion model came up with satisfactory outcomes. Both the method provided good performance with R^2 values of 0.98 and 0.96, and the root mean square error (RMSE) of 50 and 66 mg. L^{-1} respectively in calibration and evaluation tests. These values were much higher than those of conventional processing techniques. Moreover, the normal direct calibration method in acquisition signal from cobalt electrodes was also applied, which provided R^2 values of 0.7 to 0.8. These high values are sufficient for development to measure phosphate concentration in hydroponic solutions.

INDEX TERMS Phosphate sensing, multi-sensor data fusion, multivariate standard addition method (MSAM), partial least squares model (PLS), Gaussian process regression (GPR), neural network-ANN.

I. INTRODUCTION

Phosphate is an essential nutrient component needed for plants [1]. The optimization of phosphorus to improve crop productivity has been widely used in the traditional agricultural system. Moreover, the rapidly developing high production hydroponic farming has become an economically effective method of crop culture [2]. Optimized crop nutrient levels i.e. PO_4, NO_3, and K etc. (written as PO_4^{3-}, NO_3^-, and K^+ respectively in ionic form) are continuously used at all stages of hydroponic process in order to improve the quality and yield of plants [3], [4].

Appropriate phosphate supply is very important for the growth and development of plants, especially at the flowering and pollination stages. Moreover, phosphate ion occurs in three states of protonation whilst the change in pH level affects the phosphate sensor system. Thus determining phosphate concentration in a hydroponics solution is difficult and less accurate even with a cobalt electrode [5]. In addition, there is still a lack of an effective phosphate ion-selective electrode that allows to determine phosphate conveniently and precisely. Therefore, most approaches, created a reaction

The associate editor coordinating the review of this manuscript and approving it for publication was Yuan Zhang.

between cobalt electrode surface and phosphate to quantitatively detect phosphate ions from the treated soil or aqueous samples, for which a sufficient selective sensitivity could be confirmed [6]–[8]. Nevertheless, the cobalt electrode has not yet been used in practical applications because of some weaknesses such as the instability of the electrode signal band. To deal with this problem, some approaches have been used, such as the one proposed by Jung *et al.* [9] to adapt to the calibration equation they developed. However, the cobalt electrode's response potential to complex phosphate ion concentration in the solution and the high ionic strength made the simple calibration methods unsuitable [10].

The multi-sensor data fusion technique has been highlighted considerably in various fields, because it can enhance not only the accuracy of sensors but also the robustness of a system [11], [12]. These advantages have been investigated and applied in acquisition and analysis of sensor data in measurement systems. Therefore, the data fusion method was applied to effectively improve the performance of the measurement system in various fields [13]. Data fusion has been employed in precision agriculture, with soil property analysis and aggregating of databases for the development of nutrition distribution maps. La *et al.* [14] applied a soil nutrient data fusion to determine phosphate concentrations. Moreover, combining the modern multivariate calibration processing or sampling techniques are solutions that resolve drawbacks in measurement based on ISEs. For example, multivariate standard addition method (MSAM) and principal component analysis (PCA) or partial least squares regression (PLSR) model to overcome problems such as potential drift, ionic strength, and matrix effect etc. [15]–[20]. Currently, the multilayer perceptron neural network structure and learning performance have been substantially improved through the error back propagation algorithm. These exhibited particularly good performance being highly nonlinear and employed widely [21], [22]. Although the structure of the neural network is a black box and many mathematical studies have been conducted to ascertain and complete the approximation ability of various neural network types, possible omitting of step in designing a hypothesis for the model prompted to pursue ANNs in a variety of fields. The neural network was used to eliminate interference caused by changing temperature, especially, for the enhancement of sensors with ion-selective electrodes [23], interference of other ions with the primary ion [24], [25], fluctuating ionic strength [26], [27], and drifting potential [17], [28], [29] etc. Besides, the data fusion models have made a novel breakthrough in processing sensor signals to improve sensors' characteristics [30], fusing big data sensing [31], developing virtual sensors for non-availability of commercial electrodes ion types [32], [33]. Gutierrez *et al.* [34] proposed advantageous multiple ISEs and ANN model to monitor nutrients in hydroponic solutions.

Recently, fusion technique has been applied to detect phosphate in hydroponic solutions. For example, Jung *et al.* [35] used two types of sensors, i.e. NIR spectroscopy and cobalt electrode, to develop fusion for detecting PO_4 ion concentration in a paprika nutrient solution. They recorded R^2 of 0.90 in training and 0.89 in testing and RMSE of 96.70 mg.L^{-1} in training and 119.50 mg.L^{-1} in testing. La *et al.* [14] fused both ISE information in P, K, and NIR spectroscopy to estimate soil properties. The model fit resulted in $R^2 \geq 0.90$, RPD $\geq$ 3.19. Similarly, Chen *et al.* [32] used some ion-selective electrodes (NO_3, K, Ca), pH, and EC electrodes to develop phosphate and sulphate virtual sensing. Though many approaches based on unique or associative fusion techniques have been used, the cobalt electrode has not been deployed yet in practical applications. This is because of some weaknesses such as it needs frequently abrading and electrode surface treating constantly to ensure induce oxidation reaction, which affects changes in the electrode signal band [35], [36]. Moreover, MSAM is considered to be the most suitable method for compensating the shortcomings of ISEs sensing systems. However, ample information is not available about the use of the MSAM for phosphate concentration analysis.

This research proposes an approach combining MSAM in sampling and data fusion techniques. The combination will not only improve the characteristics of cobalt electrodes, but also the performance of the models. Fusing the raw data from cobalt electrochemistry aims to enhance the stabilization and accuracy of phosphate measurement in eggplant hydroponic solution, as depicted in figure 1. The electromotive force (EMF) data was acquired from four cobalt electrodes and supplied to pre-process and enhance the feature of input data and then an artificial intelligence algorithm was employed. The calibration model and comparative verification were performed using PLSR, GPR, and ANN to assess the efficiency of the model. The models were implemented and investigated through Matlab 2017b and toolboxes (Math work corporation, USA).

The present article mainly consists of five parts. Section I introduces the relevant literature and motivation of this study. Section II describes the materials and methods used for conducting the experiments. This section includes methodologies for processing sensor signals and development of models for fusion sensor data to determine phosphate concentration. In section III and IV the results and discussion of different tests are provided respectively, while section V offers the conclusions of the study.

II. MATERIALS AND METHODS

A. PREPARATION OF NUTRIENT SOLUTION SAMPLE

In this experiment, a circulating nutrient solution was prepared for growing eggplant hydroponically. The nutrient solution was prepared according to the Hoagland standard, where the stock solution was 200 times the normal concentration presented by Trejo-Téllez and Gómez-Merino [37], as shown in Table 1. The PO_4 concentration was set in the solution within the range of 6 –1350 mg.L^{-1}. The amount of KH_2PO_4 was changed to increase or decrease the PO_4 concentration.

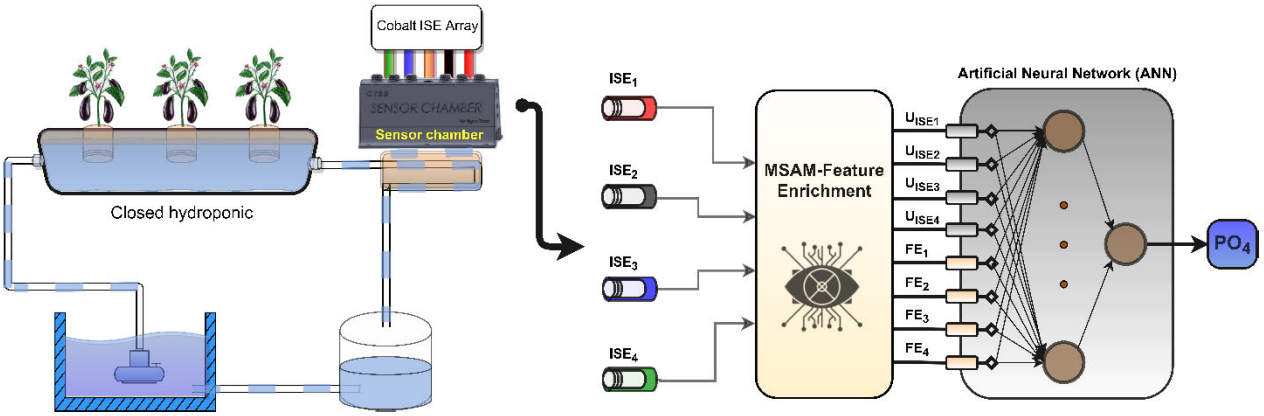

FIGURE 1. The novel combination of MSAM-feature enrichment cobalt ISEs and Artificial neural network model for detecting phosphate concentration in eggplant hydroponic solutions.

TABLE 1. Compositions of preparing 200 times concentrated solutions for eggplant hydroponic culture.

Component	Amount per litter
KNO_3	202 g
$Ca(NO_3)_2.4H_2O$	472 g
$MgSO_4.7H_2O$	493 g
NH_4NO_3	80 g
KH_2PO_4	136 g
H_3BO_3	2.86 g
$MnCl_2.4H_2O$	1.81 g
$ZnSO_4.7H_2O$	0.22 g
$CuSO_4.5H_2O$	0.051 g
$Na_2MoO_4.2H_2O$	0.12 g
Iron (Sprint 138 iron chelate)	15 g

To ensure efficient performance of models and experiments, the Kennard–Stone algorithm [38] was applied for selecting samples. One hundred experimental samples were created randomly. Out of those, 70 samples were used for training and developing the calibration and verification models and the rest (30) for testing. The actual concentration values of the created samples were analyzed at the Laboratory of Agricultural Informatization Standardization, Ministry of Agriculture and Rural Affairs, China Agricultural University, Beijing, China.

B. FABRICATION OF COBALT ELECTRODES AND SAMPLING PREPARATION

The cobalt metallic material is a selective material reactive to phosphate in the dihydrogen phosphate ion form [39], [6]. In the present study, cobalt electrodes were immersed into the sensor chamber (made of ABS-acrylonitrin butadien styren plastic using 3D-printer), which combined with the electric pump formed a simple flow injection analysis module. Accordingly, the phosphate selective electrode was created from a 99.99% pure, 5 mm diameter cobalt rod (Sigma-Aldrich, Chemical Company, USA). A length of 5 mm cobalt rod was soldered to a 1 mm copper wire, as shown in Figure 2a. The cobalt connected copper wire structure was enclosed in an ABS plastic body of 100 mm length, 12 mm external diameter (similar in diameter with other commercial ISEs for convenience of installation in the sensor chamber), and 6 mm internal diameter (created with a 3D-printer). Epoxy was injected between the cobalt metal and the plastic to prevent from making contact with other materials. The copper wire was then connected to a Bayonet Neill-Concelman (BNC) cable. The active electrode surface was prepared by polishing it on 400 and 1200 grit emery paper and a soft cloth to obtain a stable reaction. Subsequently, the electrode was pre-treated with deionized water for about 30 minutes until the EMF became uniform. This was followed by conditioning in 0.04 $mol.L^{-1}$ potassium hydrogenphthalate buffer (KHP) for further 20 minutes [40]. A double-junction reference electrode (Orion 900200, Thermo Fisher, MA, USA) in the center of the chamber was combined with the pretreated cobalt electrodes to obtain the EMF values of each sample. In this manner, the electrodes were connected to a conditioning amplifier based on the INA116 precision instrumentation amplifiers (Texas Instruments, USA) and acquired by a multichannel data acquisition device NI USB DAQ 6218 (National Instrument Corporation, USA). A sampling procedure was carried out for all the samples of the data set. First, 40 ml KHP buffer solution with 0.025M concentration was injected into the sensor chamber. An electric pump (KLP05-6, Kamoer Company, China) retained the solution flow to the electrode surface until the EMF values stabilization (roughly for 2 minutes). Then the standard concentration of hydroponic solution sample was injected into the chamber at a 1:3 ratio and cycled for 2 to 5 minutes (depending upon the concentration of the sample) until the potentials of the electrodes were stable.

A lab program based on LabVIEW (LabVIEW 2017, National Instrument Corporation, USA) was developed to monitor the specific stabilized potentials and saved to excel file as data set for developing the models (Figure 2b). On the other hand, the direct calibration method (in which the standard solution was mixed with KHP 0.025M concentration buffer solution at a 1:3 ratio and the electrodes were immersed

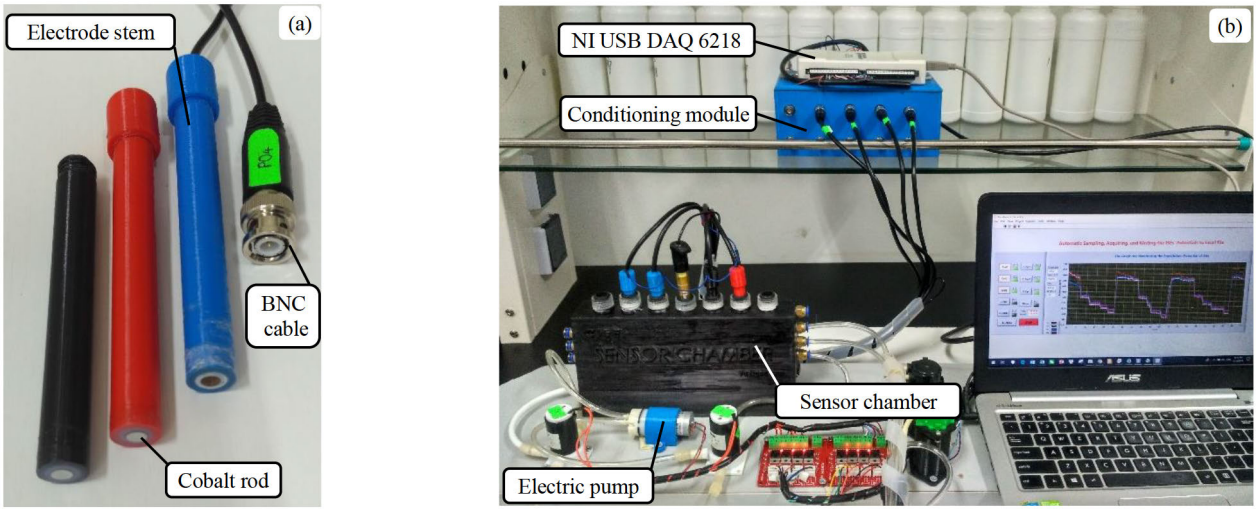

FIGURE 2. Fabricated cobalt electrodes (a) and the measurement system used in this study (b).

directly into mixture) was also carried out for comparison purposes [35].

C. PRE-PROCESSING DATA

1) NORMALIZATION OF DATA

Normalized data inputs allow the model to learn the optimal parameters for each input node rapidly. Additionally, it is useful to ensure that inputs are within the range of −1 to 1 to avoid weird mathematical artifacts associated with floating-point number precision. The inputs and outputs are normalized such that the values on each channel occupy the same range before training. Thus, the input vector C_i for the entire suite of training samples can be expressed as:

$$C_{nor} = \frac{(y_{max} - y_{\min})}{(C_{imax} - C_{\min})}.(C_i - C_{i\,min}) + y_{min} \quad (1)$$

where y_{min} and y_{max} are a standard range (normally −1 to 1), which can be set independently for each signal, C_{nor} is normalized concentration, C_{imin} is the minimum value of the C_i vector, and C_{imax} is the corresponding maximum value. The C_i can be inverted to yield the original concentration value as

$$C_i = (C_{nor} - y_{min}).\frac{(C_{imax} - C_{i\,min})}{(y_{max} - y_{min})} + C_{imin} \quad (2)$$

2) MSAM AND DATA FEATURE ENRICHMENT

Standard addition is employed to be the most effective method for compensating matrix effects [41], potential drift, and ionic strength [42]. However, extrapolation is an inherent disadvantage of standard addition method (SAM). Therefore, in this study, multivariate standard addition method (MSAM) calibration procedure with one spike was used to substitute the traditional multiple spikes univariate calibration. In this approach, a signal received from the blank sample was used as offset, which could be combined with the spiked signal (the signal from standard addition) to enrich the feature of the sensor, as presented in the following equation:

$$U_{FEij} = U_{xij} - U_{0ij} \quad (3)$$

where i is the number of samples from 1 to 100. j is the number of ion-selective electrodes from 1 to 4. U_{0ij}, U_{xij} are the potential of corresponding ISE at the concentration C_0 and C_x respectively. U_{FEij} represents the difference value between U_{0ij} and U_{xij}, which were used to enrich the feature of the dataset to improve the performance of the model, called enriched data values. Therefore, using the data set fed into the models is divided into two scenarios, i) the input of data set is four columns including U_{ISE1} to U_{ISE4} , and ii) the input of the data set is eight columns, i.e. U_{ISE1} to U_{ISE4}, and U_{FE1} to U_{FE4}. Three models (PLSR, GPR, and ANN) were used in the experiment with all the data sets to find out the most appropriate model and evaluate the efficiency of the proposed method in determining phosphate concentration.

TABLE 2. Forming data set by combining the original cobalt electrochemistry signal and MSAM-Feature enrichment values.

S.No	C	U_{ISE1}	U_{ISE2}	U_{ISE3}	U_{ISE4}	U_{FE1}	U_{FE2}	U_{FE3}	U_{FE4}
1	C_{01}	U_{011}	U_{012}	U_{013}	U_{014}	NaN	NaN	NaN	NaN
	C_{x1}	U_{x11}	U_{x12}	U_{x13}	U_{x14}	U_{FE11}	U_{FE12}	U_{FE13}	U_{FE14}
2	C_{02}	U_{021}	U_{022}	U_{023}	U_{024}	NaN	NaN	NaN	NaN
	C_{x2}	U_{x21}	U_{x22}	U_{x23}	U_{x24}	U_{FE21}	U_{FE22}	U_{FE23}	U_{FE24}
...	...	...	...	...	...	...	...	...	...
100	C_{0100}	U_{01001}	U_{01002}	U_{01003}	U_{01004}	NaN	NaN	NaN	NaN
	C_{x100}	U_{x1001}	U_{x1002}	U_{x1003}	U_{x1004}	U_{FE1001}	U_{FE1002}	U_{FE1003}	U_{FE1004}

S.No: Sample No; C-Concentration; U_{ISE1}, U_{ISE2}, U_{ISE3}, U_{ISE4}, U_{FE1}, U_{FE2}, U_{FE3}, U_{FE4} *– Potential of electrodes* ISE_1, ISE_2, ISE_3, ISE_4 *and feature enrichment values of them, respectively; NaN- Not a number.*

D. DEVELOPMENT OF DATA FUSION MODEL FOR PHOSPHATE SENSING

1) PARTIAL LEAST SQUARE REGRESSION

Partial least square regression (PLSR), a stepwise linear model that extracts latent variables and seems to fit very well with a lot of spectroscopical problems. In this research, PLSR was used to predict phosphate concentration via the sensor signals (original calibration data matrix **X**) form cobalt electrodes. Unlike the principal component analysis regression (PCR) model, the PLSR scores are estimated by including both sensor signals and concentration information. Specifically, the PLS calibration phase consists of the following inverse model [43]:

$$\mathbf{y}_{\mathrm{n}} = \mathbf{T}_{\mathrm{A}}\mathbf{v}_{\mathrm{n}} + \mathbf{e} \tag{4}$$

$$\mathbf{T}_{\mathrm{A}} = \mathbf{X}^{\mathrm{T}}\mathbf{U}_{\mathrm{A}} \tag{5}$$

where $\mathbf{v}_{\mathrm{n}}$ (of size Ax1) is the vector of PLS regression coefficients defined in the PLS latent space. $\mathbf{T}_{\mathrm{A}}$ is the truncated score PLS matrix, the projection of the original data matrix **X** onto the space defined by the loadings contained in $\mathbf{U}_{\mathrm{A}}$, calculated by equation (5). $\mathbf{U}_{\mathrm{A}}$ is loading matrix which is estimated by the eigenvectors of the square matrix $\mathbf{X}\mathbf{X}^{\mathrm{T}}$ [44]. **e** is a vector collecting the concentration modeling errors and the $\mathbf{v}_{\mathrm{n}}$ is derived by equation (6)

$$\mathbf{v}_{\mathrm{n}} = \mathbf{T}_{\mathrm{A}}\mathbf{y}_{\mathrm{n}} \tag{6}$$

In PLS, both the loadings (**P**) and the weight loadings (**W**) participate in estimating the calibration score matrix. As in PCR, they are also truncated to the first (**A**) column, which should retain the main portion of both the sensor signals and concentration variance. The specific expression for the PLS score matrix is:

$$\mathbf{T}_{\mathrm{A}} = \mathbf{X}^{\mathrm{T}}\mathbf{W}_{\mathrm{A}}(\mathbf{P}_{\mathrm{A}}^{\mathrm{T}}\mathbf{W}_{\mathrm{A}})^{-1} \tag{7}$$

In the prediction phase, the regression coefficients are employed to estimate the analyte concentration in a future sample. A previous step is required to find the test sample scores, which proceeds by means of the truncated loading matrices $\mathbf{W}_{\mathrm{A}}$ and $\mathbf{P}_{\mathrm{A}}$. The specific PLS expression for the test sample score vector is:

$$\mathbf{t}_{\mathrm{A}}(\mathbf{P}_{\mathrm{A}}^{\mathrm{T}}\mathbf{W}_{\mathrm{A}})^{-1}\mathbf{W}_{\mathrm{A}}^{\mathrm{T}}\mathbf{X} \tag{8}$$

and the prediction equation is:

$$\mathrm{y} = \mathbf{v}_{\mathrm{n}}^{\mathrm{T}}\mathbf{t}_{\mathrm{A}} \tag{9}$$

where y is the predicted analyte concentration in the test sample, whose sensor signal is the vector **x** of equation (7).

2) GAUSSIAN PROCESS REGRESSION

The Gaussian Process is an appropriate method for defining the preferred distribution of flexible regression and classification models where the functions of regression or class probability are not limited to simple parametric forms. One advantage of the Gaussian Process is the wide variety of its covariance functions, leading to functions with varying degrees of smoothness, or different types of continuous structures, allowing to choose from among them appropriately. These models can specify one or more input variable distributions among the functions. If the mean answer is calculated in a regression model with Gaussian errors, inferences can be made by the matrix calculations. Here, data systems with samples of more than one thousand each can be used.

In the regression task, the data set D consisted of N input vectors $x_1, x_2, \ldots x_N$(signals of dimension D from sensors) with corresponding continuous outputs $y_1, y_2, \ldots y_N$ (phosphate concentration of samples). The outputs are assumed to be noisily observed from an underlying functional mapping $f\,(x)$. The object of the regression task is to estimate $f\,(x)$ from the data D. The data sets can be corresponded with a Gaussian process. The median for the Gaussian cycle is usually assumed to be zero everywhere. In such situations, the covariance function relates one result to the other. With the Gaussian noise model, each observation can be associated with a kenel function, $k\left(x, \acute{x}\right)$. Each observation y can be connected to the main function through the Gaussian noise model [45].

$$y = f\,(x) + N\left(0, \sigma_f^2\right) \tag{10}$$

where $N\left(0, \sigma_f^2\right)$ is the noise of normal distribution function with mean 0 and variance σ_f^2. Regression means looking for $f\,(x)$. The $k\left(x, \acute{x}\right)$ is calculated as follows:

$$k\left(x, \acute{x}\right) = \sigma_f^2 exp\left[\frac{-(x - x')^2}{2l^2}\right] + \sigma_n^2\delta\left(x, \acute{x}\right) \tag{11}$$

where $\delta\left(x, \acute{x}\right)$ is the Kronecker delta function, which is one if $x = x'$ and zero otherwise, l is length-scale. The predicted values of observations are the same in accordance with equation (10), but variances vary due to observational noise process. To prepare the GPR for covariance function, equation (11) is incorporated in all possible combinations of these points and findings are summarized in three covariance matrices:

$$\mathrm{K} = \begin{bmatrix} \mathrm{k}\,(\mathrm{x}_1, \mathrm{x}_1) & \mathrm{k}\,(\mathrm{x}_1, \mathrm{x}_2) & \ldots & k\,(\mathrm{x}_1, \mathrm{x}_\mathrm{n}) \\ \mathrm{k}\,(\mathrm{x}_2, \mathrm{x}_1) & \mathrm{k}\,(\mathrm{x}_2, \mathrm{x}_2) & \ldots & k\,(\mathrm{x}_2, \mathrm{x}_\mathrm{n}) \\ \vdots & \vdots & \vdots & \vdots \\ \mathrm{k}\,(\mathrm{x}_\mathrm{n}, \mathrm{x}_1) & \mathrm{k}\,(\mathrm{x}_\mathrm{n}, \mathrm{x}_2) & \ldots & k\,(\mathrm{x}_\mathrm{n}, \mathrm{x}_\mathrm{n}) \end{bmatrix} \tag{12}$$

$$\mathrm{K}_* = [\mathrm{k}\,(\mathrm{X}_*, \mathrm{X}_1)\,\mathrm{k}\,(\mathrm{X}_*, \mathrm{X}_2) \ldots \mathrm{k}\,(\mathrm{X}_*, \mathrm{X}_\mathrm{n})]\;\mathrm{K}_{**} = \mathrm{k}\,(\mathrm{X}_*, \mathrm{X}_*) \tag{13}$$

The diagonal elements of matrix K are in the form of $\sigma_f^2 + \sigma_n^2$, and non-diagonal elements approach zero when the x extends a large domain [46].

3) NEURAL NETWORK MODEL

There are four layers of neurons (nodes) in the architecture of ANN and MSAM-FE-ANN systems, described in this paper. These include an input layer, two hidden layers, and an output layer. The predictor signals are fed into the

TABLE 3. Types of activation functions used in the neural network.

Name	Logistic (sigmoid)	Tanh	Radbas	Relu
Equation	(15) $f(x)=\frac{1}{1+e^{-x}}$	(16) $f(x)=\frac{2}{1+e^{-2x}}-1$	(17) $f(x)=e^{-x^2}$	(18) $f(x)=\begin{cases}0, & x<0\\ x, & x\geq 0\end{cases}$

TABLE 4. The structure of the neural network model for predicting phosphate concentration.

Parameter	Values
Number of hidden layers	1, 2
Hidden layer size	5 to 60
Hidden layer transfer function f(x)	tansig, logsig, radbas
Output layer transfer function	purelin
Learning method	Levenberg-Marquardt (LM)
Learning rate	0.001; 0.002
Max number of epochs	1000
Training goal	10^{-6}

input layer and transmitted to the hidden layers and output layer through the weighted connections and the transfer functions [47]. Figure 3 illustrates the simple multilayer structure of an ANN and the neuron [48]. The input variable x_i connects to each single input node and each input node transmits a weight value W_{ij} to the hidden layer. Subsequently, these input-weight products are added and the sum is passed through an activation function a_{ij}, to determine whether and to what extent that signal should progress further through the network to affect the ultimate outcome. Figure 3b depicts the principle of a neuron in the neural network, and equation (14) is the mathematical expression of the i^{th} neuron:

$$y_i = f\left(\sum_i w_{ij}{}^* x_i + b_i\right) \tag{14}$$

where x_i is the input information of neuron i, w_{ij} is the network connection weight, f is the activation function, b_i is the bias, and y_i is the output value. The regular learning rule is to use a gradient descent algorithm to constantly adjust the weight and threshold of the network through backpropagation, in order to minimize the error square sum of the network.

To avoid local minima and overfitting, the Levenberg–Marquardt algorithm was used for training the network. We adopted some functions as the transfer (activation) function (table 3) for neural network [49]. The main network parameters (number of layers and nodes in the hidden layer, transfer function, learning algorithm, and learning rates) were subjected to various trials to find out the best fitting condition. The network structure parameters were set as in table 4. The number of neurons in the hidden layer was set in the range of 5 to 60 to consider the effect of number of neurons in the hidden layer on the training results. Considering that the neural network prediction results have a certain volatility, each neural network with different neurons in the hidden layers was trained several times. The best observed values of two performance indices were used to

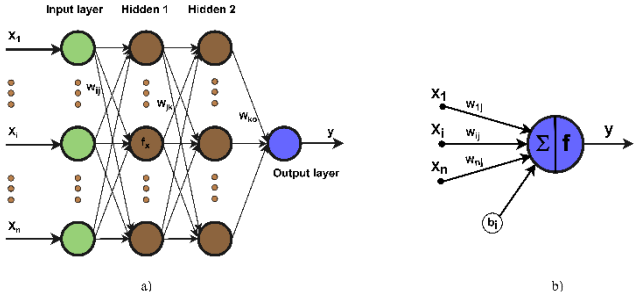

FIGURE 3. Neural network structure diagram (a) and the neuron (b).

define the suitable number of neurons in the hidden layers, determined to be 25, 35 nodes for hidden layer 1 and hidden layer 2 respectively.

4) PREDICTION PERFORMANCE INDICES

To compare the effectiveness of data fusion techniques in predicting phosphate concentration, two prediction performance indices were estimated, i.e. the root mean square error (RMSE) and determination coefficient (R^2). The smaller the values of RMSE, the closer the predicted value is to the true value, which means better prediction accuracy. Besides that, the closer R^2 value is to unity, the better the machine learning prediction is. RMSE is calculated using the following formula:

$$RMSE = \sqrt{\frac{1}{n}\sum_{i=1}^{n}\left(y_i - \hat{y}_i\right)^2} \tag{15}$$

where n is the total number of data in the training set or test set, y_i is the actual phosphate concentration value, and $\hat{y}_i$ is the predicted phosphate concentration value.

R^2 is an index that measures the degree of agreement between the test data and the fitting function and is calculated as per equation (16).

$$R^2 = 1 - \frac{\sum_i \left(y_i - \hat{y}_i\right)^2}{\sum_i \left(\bar{y}_i - \hat{y}_i\right)^2} \tag{16}$$

where $\bar{y}_i$ is the average value of the test set.

III. RESULTS

A. COBALT ELECTRODE REACTION

As shown in Figure 4, four cobalt electrodes were used to detect phosphate concentrations in the range of 6–1350 $mg.L^{-1}$ using the direct calibration method [50]. The Nikolsky-Eisenman equation calibration plot showed the logarithmic relationship between the phosphate concentration and electrode EMF. In this manner, the determination coefficient (R^2) ranged from 0.68 to 0.71, which cannot be employed as that of actual sensors. The signs on the graph also show that a considerable variation in the reactions occured with the polished surface of electrodes. The cobaltous oxide layer on the cobalt electrodes formed and deviated during exposure to the sample solution. This caused changes in the state of electrode conditioning and required

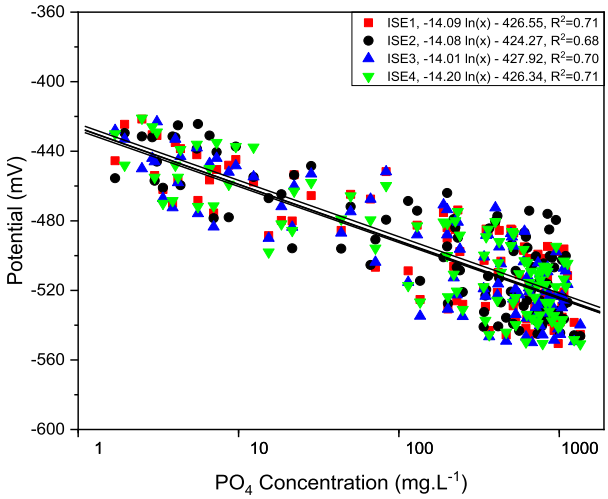

FIGURE 4. Direct calibration plots of cobalt-electrode response to PO_4 concentrations.

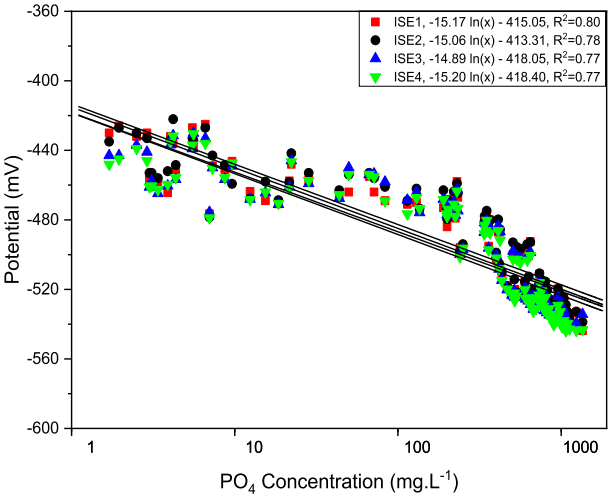

FIGURE 5. Standard addition calibration plots of cobalt-electrode response to PO_4 concentrations.

a normalization process. This problem could be reduced by using automated measuring and conditioning equipment having a buffering solution to stabilize the oxidation surface for a long time [51].

Figure 5 illustrates the calibration plots of four electrodes based on the standard addition method. In this approach, the R^2 was 0.77 to 0.80, which was higher than that of the direct calibration technique (Fig 4). The improvement of detecting coefficient of electrodes in this manner could be due to the reason that the ISEs' surfaces were being re-conditioned in each sample by the buffer solution (KHP) during the first step of sampling phase [52]. Nevertheless, these achievements still do not fulfill the requirements of an actual measurement system. Thus, to enhance the efficiency of electrodes, three models including PLSR, GPR, and ANN were deployed.

Figure 6 shows the results of the test sample obtained through the prediction of the two hidden layers with 12 and

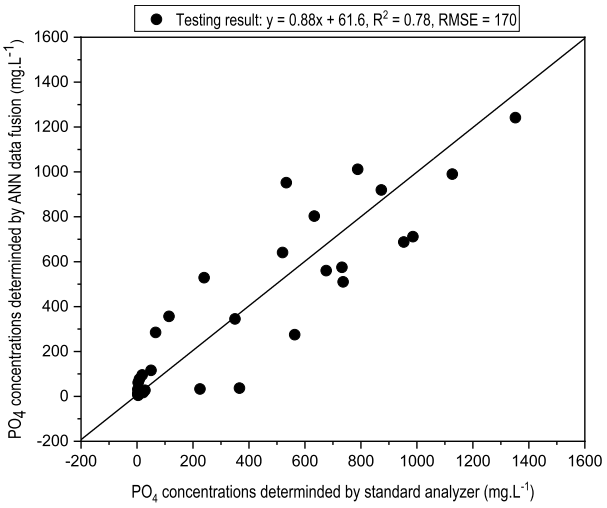

FIGURE 6. Prediction results using ANN analysis with the EMF response of cobalt electrodes based direction sampling method.

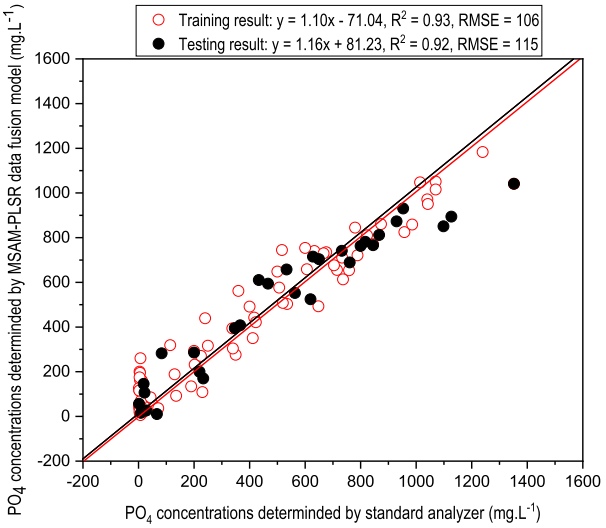

FIGURE 7. Calibration and test sample prediction results with MSAM - PLSR analysis performed using the enriched data.

18 nodes and Tanh and the Logistic function ANN model with the data set from the direction calibration method. The RMSE obtained for the test sample was 170 $mg.L^{-1}$ with an R^2 of 0.78, although the result was slightly better than that obtained using the linear calibration model (Figure 4). Thus, the procedure was still not satisfactory for determining phosphate component in the hydroponic system.

B. PERFORMANCE OF THE PROPOSED MODELS

As mentioned earlier, we attempted to develop prediction models through data fusion. First, the raw cobalt EMF data acquired through a multivariate standard addition method were pre-processed to produce the feature enrichment data. Then the PLSR, GPR, and ANN models were employed for the fusion methods.

Figures 7, 8, and 9 depict the performances of the developed models for predicting samples. As shown in Figure 7,

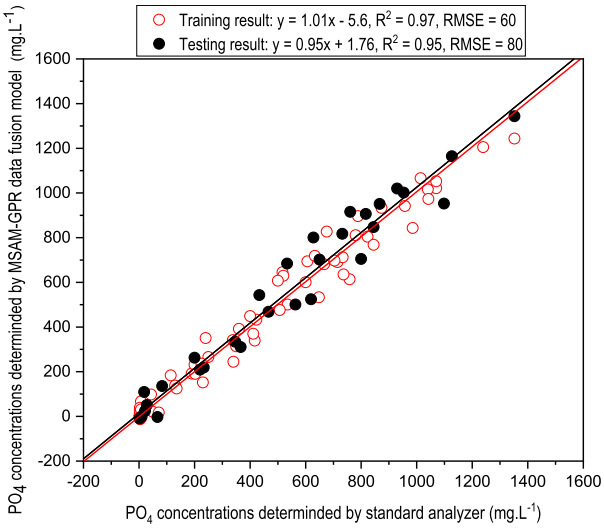

FIGURE 8. Calibration and test sample prediction results with MSAM - GPR analysis performed using the enriched data.

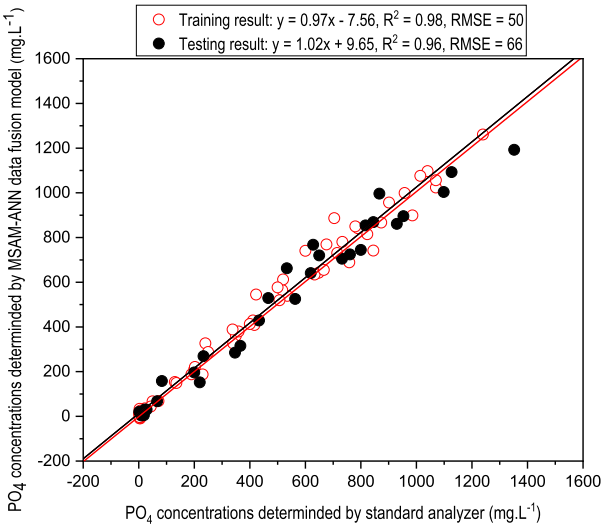

FIGURE 9. Calibration and test sample prediction results with MSAM – ANN analysis performed using the enriched data.

the calibration model of fusing data model developed by PLSR technique had a slope of 1.1 and an R^2 of 0.93. The predicted results of the test samples had a slope of 1.16 and an intercept of 81.23, with an R^2 of 0.92 and an RMSE of 115 $mg.L^{-1}$. Whereas, the fused database GPR model demonstrated slightly better results (as shown in Figure 8), where the slope of the calibration phase was 1.01 and the R^2 value was 0.97. The testing prediction results of the GPR calibration model achieved R^2 of 0.95 and RMSE of 80 $mg.L^{-1}$. The results of GPR model using the difference of kernel (covariance) functions are presented in table 5. The squared exponential kernel gave better results than the other four. The graphs revealed that performances of both the PLSR and GPR improved compared to those of cobalt electrode EMF data.

Figure 9 shows the best fit of developed ANN model with a slope of 0.97, an offset of 7.56, an R^2 of 0.98, and an RMSE

TABLE 5. Types of kernel functions used in GPR model.

Prameters \ Kernel functions		Squared exponential	Exponential	Matern 3/2	Matern 5/2	Rational quadratic
RMSE results	Training	60	38	72	72	69
	Testing	80	91	100	99	96
R^2 results	Training	0.974	0.98	0.96	0.96	0.965
	Testing	0.95	0.94	0.93	0.93	0.936

TABLE 6. The structures of the neural network model for predicting phosphate concentration.

Hidden 1		Hidden 2		Learning rate	Training results		Testing results	
Number of Nodes	Activation function	Number of Nodes	Activation function		RMSE	R^2	RMSE	R^2
20	Logsig	30	Tansig	0.001	68	0.95	88	0.94
60	Radbas	30	Tansig	0.001	45	0.98	95	0.93
50	Radbas	40	Logsig	0.002	43	0.98	90	0.94
25	Logsig	30	Radbas	0.001	70	0.95	100	0.92
25	**Tansig**	**35**	**Logsig**	**0.001**	**50**	**0.98**	**66**	**0.96**
30	Tansig	45	Radbas	0.002	65	0.95	92	0.92

of 50 $mg.L^{-1}$. The slope of predicted results of the developed model with the test samples was 1.02, a residual of 9.65, with an R^2 of 0.96 and an RMSE of 66 $mg.L^{-1}$. Table 6 exhibits the results of the ANN model, in both the training and testing phases, under each proposed training condition. The Tanh and Logistic were two functions through which the model acquired the best results, where the parameters were 25 and 35 neurons for the first and second hidden layers and the learning rate was 0.001.

The results of three calibration models (PLSR, GPR, and ANN) based on enriched data sets fed from MSAM and expanding feature pre-process are summarized in table 7. In the case of PLS, when the number of electrodes was adjusted (from 1 to 4) with or without the feature enrichment (FE) signals, the best results were obtained, especially when four electrodes' signals and four enriched features were used. They produced an R^2 of 0.93, an RMSE of 106 $mg.L^{-1}$, and a validation R^2 of 0.92, and corresponding RMSE of 115 $mg.L^{-1}$. The best results in GPR were those of preprocessing using four electrodes and four FE signals. The GPR results had an R^2 of 0.97 and an RMSE of 60 $mg.L^{-1}$ in the calibration phase, and in the validation phase, the R^2 and RMSE were 0.95 and 80 $mg.L^{-1}$ respectively. These were slightly higher than the R^2 and RMSE of the PLSR model. The results of ANN model, when the number of electrodes changed from one to four, had significant variations too. The prediction of the MSAM-ANN model achieved good results while using three electrodes with FEs. Especially, with number of data feature was eight (four electrodes and four FEs), the best results was obtained with an R^2 of 0.96 and RMSE of 66 $mg.L^{-1}$. It was quite satisfying as with this approach, the proposed model could be employed for the

TABLE 7. A comparison of the results of a combination number of electrode inputs (ISEs) with feature enriched data inputs (FE) and the used models; chosen item (☑), unchosen item (–).

Number of Electrodes	Inputs		Models			Training results		Testing results	
	ISEs	ISEs+FE	PLSR	GPR	ANN	RMSE	R2	RMSE	R2
1	☑	☑	☑	-	-	151	0.87	132	0.85
	☑	☑	-	☑	-	111	0.91	135	0.86
	☑	☑	-	-	☑	90	0.93	145	0.86
2	☑	-	☑	-	-	149	0.879	129	0.86
	☑	-	-	☑	-	108	0.92	124	0.88
	☑	-	-	-	☑	106	0.92	140	0.89
	☑	☑	☑	-	-	144	0.88	127	0.87
	☑	☑	-	☑	-	93	0.94	133	0.88
	☑	☑	-	-	☑	71	0.97	137	0.896
3	☑	-	☑	-	-	143	0.88	114	0.89
	☑	-	-	☑	-	102	0.927	108	0.907
	☑	-	-	-	☑	110	0.92	121	0.90
	☑	☑	☑	-	-	134	0.89	108	0.90
	☑	☑	-	☑	-	95	0.93	100	0.92
	☑	☑	-	-	☑	60	0.97	87	0.94
4	☑	-	☑	-	-	123	0.90	132	0.89
	☑	-	-	☑	-	84	0.95	93	0.93
	☑	-	-	-	☑	86	0.95	90	0.94
	☑	☑	☑	-	-	106	0.93	115	0.92
	☑	☑	-	☑	-	60	0.97	80	0.95
	☑	☑	-	-	☑	**50**	**0.98**	**66**	**0.96**

actual sensors. These results suggest that the MSAM and the FE pre-processing stage would play an essential role in improving the performance of the models in developing the cobalt electrochemistry fusion data for detecting phosphate in a hydroponic solution.

IV. DISCUSSIONS

In this study, we employed the two measurement techniques and three fusion methods to develop a more accurate prediction model through fusing cobalt electrochemistry data. In the direct calibration technique, the results of the cobalt electrode exhibited low repeatability. The highest determination coefficient (R^2) of the linear regression model was 0.71, which is similar the results obtained in previous studies [35], which reported a high possibility of alteration in oxidation layer (CoO) formation on the cobalt electrodes. To combat this issue, a buffering solution is used to maintain the oxidation layer over a long period [53] or the normalization is used based on two, three, or even multiple points. However, this needs to keep prepared the standard solution at all times for fitting the nonlinear data [54], which is a disadvantage of this method. Thus, in this study, we used merely cobalt electrochemistry raw data.

In the MSAM technique, the multivariate standard addition procedure for sampling and pre-processing data of cobalt electrochemistry were carried out. At first in the sampling procedure, the electrodes were kept in contact with the KHP buffer solution flow. This was significant as it not only made electrodes work stably but also reduced the fluctuation of electrodes potential by changing the states of the electrode conditioning [36]. Using MSAM reduces the conjunction potential and drift potential of electrodes [42], and the ionic strength [20], [55]. Figure 5 shows that the distributions of electrodes potentials were more correlative with the phosphate concentration than those of direction method sampling (figure 4), especially in a high range of concentration in which the electrode reaction with phosphate was more stable [36]. The combination of MSAM and electrodes potential in KHP solution is required for enriching the data information from physical sensor data (depicted in table 2). That is called feature enrichment (FE). In previous research, FE normally was used for image processing and text classification etc. that used complicated algorithm and took more time in processing [55], [56].

PLSR and GPR, reported in previous studies [20], [57]–[59], were used to process enriched data and contributed towards the improvement in phosphate sensing performance. Although the enriched data improved the performance of the models significantly (see table 7), they also increased the nonlinearity. Therefore, the performance of PLSR has been improved slightly (compare the efficiency of PLSR between four ISEs signals only and four ISEs signals plus four FEs signals in combination). Table 7 also demonstrates that when the number of electrodes was less, the performance of GPR model was better than that of ANN, and ANN was slightly over fitted (at two electrodes). However, when the number of electrodes is sufficient, the performance of ANN is better [47]. This can be explained by the generalization ability of ANN that needs the data set big enough, whereas the GPR can work well with small and nonlinear data. To evaluate the performance of ANN the main parameters were adjusted in several cases. The results of proposed ANN were significantly better than those of the PLSR and GPR. Thus, we can employ only four electrodes and combine MSAM, FE, and ANN to determine the phosphate concentration closer to the requirements of the real hydroponic system.

The proposed models were calibrated to predict phosphate in range 6 to 1350 $mg.\ L^{-1}$. This allowed observing the abnormal phosphate ions changes in the nutrient solution at a regular concentration within 100 to 500 $mg.\ L^{-1}$ range, as well as in the applicable hydroponic nutrient solutions. Moreover, the wide-range measurement of the models would fulfill the requirements of an automated control nutrient system even when the phosphate concentration changes to a large extent (tens to thousands of $mg.\ L^{-1}$) due to possible technical issues e.g., pump, valve errors etc. [35].

In this study, we used four cobalt electrodes, MSAM and FE techniques, and the data fusion method. Therefore, the proposed system is not only easy to deploy in the field but also can predict the phosphate concentrations rapidly (in a few minutes) except for long conditioning and preparation time. Furthermore, the materials used in this study are common, easy to fabricate with low cost. These contribute towards reducing the system cost that detect nutrient concentration in

the hydroponic solution. The proposed method is also useful in measuring equipment for examining soil and wastewater pollution where phosphate contamination is in large quantity [60], [61].

V. CONCLUSION

In this paper, we have presented the combination of MSAM sampling and data fusion method to develop PLSR, GPR and ANN models for determining phosphate ion concentrations in an eggplant nutrient solution. The following are the main findings in our research.

The MSAM was used to collect the samples of phosphate concentration solution for preparing the data set to develop the data fusion models. The MSAM based sampling showed that it improved the detecting efficiency of four cobalt electrodes. This reduced the cost of the measurement system and improved quality as well. The expansion of the sensor data features based on MSAM was proposed, where the models were fed more information about the data to predict more accurately. Furthermore, feature enrichment was quite simple due to the zero-order of the electrodes signal and the results showed that it was relatively significant in this scenario.

Three models including PLSR, GPR and ANN were used to fuse the enriched feature data to improve the models performance. The calibration sample results showed that the PLSR and GPR models provided better performances (R^2 of 0.925 and 0.97, respectively) than the models without data fusion. The results of measuring the test samples for verification showed R^2 values of 0.92 and 0.95. Nevertheless, the ANN prediction model achieved better results than the other two models. The calibration phase of the model showed an R^2 of 0.98 and RMSE of 50 mg. L^{-1} and the test phase showed an R^2 value of 0.96 and RMSE of 66 mg. L^{-1}.

The results obtained in this work revealed that combining MSAM-FE and fusion method facilitated the effective detection of phosphate. Furthermore, the proposed approach can be used for monitoring phosphate in a closed hydroponic nutrient solution rapidly and effectively. This method is expected to be used in future for developing low-cost and more accurate sensors for a variety of fields.

ACKNOWLEDGMENT

The authors would like to thank for all the help of the teachers and students of the related universities.

REFERENCES

[1] M. Naeem, A. A. Ansari, and S. S. Gill, *Essential Plant Nutrients: Uptake, Use Efficiency, and Management*. Cham, Switzerland: Springer, 2017.
[2] D. Savvas, "Hydroponics: A modern technology supporting the application of integrated crop management in greenhouse," *J. Food Agricult. Environ.*, vol. 1, no. 1, pp. 80–86, Jan. 2003.
[3] M. M. Maboko, C. P. Du Plooy, and S. Chiloane, "Yield and mineral content of hydroponically grown mini-cucumber (Cucumis sativus L.) as affected by reduced nutrient concentration and foliar fertilizer application," *HortScience*, vol. 52, no. 12, pp. 1728–1733, Dec. 2017, doi: 10.21273/hortsci12496-17.
[4] H. Zekki, L. Gauthier, and A. Gosselin, "Growth, productivity, and mineral composition of hydroponically cultivated greenhouse tomatoes, with or without nutrient solution recycling," *J. Amer. Soc. Horticultural Sci.*, vol. 121, no. 6, pp. 1082–1088, Nov. 1996.
[5] J. Y. Lee, A. Rahman, H. Azam, H. S. Kim, and M. J. Kwon, "Characterizing nutrient uptake kinetics for efficient crop production during Solanum lycopersicum var. Cerasiforme Alef. Growth in a closed indoor hydroponic system," *PLoS ONE*, vol. 12, no. 5, May 2017, Art. no. e0177041, doi: 10.1371/journal.pone.0177041.
[6] H. J. Kim, J. W. Hummel, K. A. Sudduth, and S. J. Birrell, "Evaluation of phosphate ion-selective membranes and cobalt-based electrodes for soil nutrient sensing," *Trans. ASABE*, vol. 50, no. 2, pp. 415–425, Oct. 2013.
[7] D. Xiao, H.-Y. Yuan, J. Li, and R.-Q. Yu, "Surface-modified cobalt-based sensor as a phosphate-sensitive electrode," *Anal. Chem.*, vol. 67, no. 2, pp. 288–291, Jan. 1995, doi: 10.1021/ac00098a009.
[8] D.-H. Jung, H.-J. Kim, W.-J. Cho, S. H. Park, and S.-H. Yang, "Validation testing of an ion-specific sensing and control system for precision hydroponic macronutrient management," *Comput. Electron. Agricult.*, vol. 156, pp. 660–668, Jan. 2019, doi: 10.1016/j.compag.2018.12.025.
[9] D. H. Jung, H. J. Kim, G. L. Choi, T. I. Ahn, J. E. Son, and K. A. Sudduth, "Automated lettuce nutrient solution management using an array of ion-selective electrodes," *Trans. ASABE*, vol. 58, no. 5, pp. 1309–1319, Sep./Oct. 2015.
[10] F. Sales, M. P. Callao, and F. X. Rius, "Multivariate standardization for correcting the ionic strength variation on potentiometric sensor arrays," *Analyst*, vol. 125, no. 5, pp. 883–888, 2000.
[11] A. P. Kene and S. K. Choudhury, "Analytical modeling of tool health monitoring system using multiple sensor data fusion approach in hard machining," *Measurement*, vol. 145, pp. 118–129, Oct. 2019, doi: 10.1016/j.measurement.2019.05.062.
[12] Y.-H. Liao and J.-C. Chou, "Weighted data fusion use for ruthenium dioxide thin film pH array electrodes," *IEEE Sensors J.*, vol. 9, no. 7, pp. 842–848, Jul. 2009, doi: 10.1109/jsen.2009.2024045.
[13] D. L. Hall and J. Llinas, "An introduction to multisensor data fusion," *Proc. IEEE*, vol. 85, no. 1, pp. 6–23, Jan. 1997, doi: 10.1109/5.554205.
[14] W. J. La, K. A. Sudduth, H. J. Kim, and S. O. Chung, "Fusion of spectral and electrochemical sensor data for estimating soil macronutrients," *Trans. ASABE*, vol. 59, no. 4, pp. 787–794, 2016, doi: 10.13031/trans.59.11562.
[15] J. Fonollosa, L. Fernández, A. Gutiérrez-Gálvez, R. Huerta, and S. Marco, "Calibration transfer and drift counteraction in chemical sensor arrays using direct standardization," *Sens. Actuators B, Chem.*, vol. 236, pp. 1044–1053, Nov. 2016, doi: 10.1016/j.snb.2016.05.089.
[16] M. Khaydukova, V. Panchuk, D. Kirsanov, and A. Legin, "Multivariate calibration transfer between two potentiometric multisensor systems," *Electroanalysis*, vol. 29, no. 9, pp. 2161–2166, Sep. 2017, doi: 10.1002/elan.201700190.
[17] V. Panchuk, L. Lvova, D. Kirsanov, C. G. Gonçalves, C. Di Natale, R. Paolesse, and A. Legin, "Extending electronic tongue calibration lifetime through mathematical drift correction: Case study of microcystin toxicity analysis in waters," *Sens. Actuators B, Chem.*, vol. 237, pp. 962–968, Dec. 2016, doi: 10.1016/j.snb.2016.07.045.
[18] D. Melucci and C. Locatelli, "Multivariate calibration in differential pulse stripping voltammetry using a home-made carbon-nanotubes paste electrode," *J. Electroanal. Chem.*, vol. 675, pp. 25–31, Jun. 2012, doi: 10.1016/j.jelechem.2012.04.020.
[19] L. Kortazar, J. Sáez, J. Agirre, J. K. Izaguirre, and L. A. Fernández, "Application of multivariate analysis to the turbidimetric determination of sulphate in seawater," *Anal. Methods*, vol. 6, no. 10, pp. 3510–3514, Mar. 2014, doi: 10.1039/c4ay00335g.
[20] C. Pérez-Ráfols, J. Puy-Llovera, N. Serrano, C. Ariño, M. Esteban, and J. M. Díaz-Cruz, "A new multivariate standard addition strategy for stripping voltammetric electronic tongues: Application to the determination of Tl(I) and In(III) in samples with complex matrices," *Talanta*, vol. 192, pp. 147–153, Jan. 2019, doi: 10.1016/j.talanta.2018.09.035.
[21] B. Idjeri, M. Laghrouche, and J. Boussey, "Wind measurement based on MEMS micro-anemometer with high accuracy using ANN technique," *IEEE Sensors J.*, vol. 17, no. 13, pp. 4181–4188, Jul. 2017, doi: 10.1109/jsen.2017.2701502.
[22] L. T. Duarte, C. Jutten, and S. Moussaoui, "A Bayesian nonlinear source separation method for smart ion-selective electrode arrays," *IEEE Sensors J.*, vol. 9, no. 12, pp. 1763–1771, Dec. 2009, doi: 10.1109/jsen.2009.2030707.
[23] R. Bhardwaj, "Temperature compensation of ISFET based pH sensor using artificial neural networks," in *Proc. IEEE Regional Symp. Micro Nanoelectron.*, 2017, pp. 155–158.

[24] L. Wang, Y. Cheng, D. Lamb, Z. Chen, P. J. Lesniewski, M. Megharaj, and R. Naidu, "Simultaneously determining multi-metal ions using an ion selective electrode array system," *Environ. Technol. Innov.*, vol. 6, pp. 165–176, Nov. 2016, doi: 10.1016/j.eti.2016.10.001.

[25] S. Bermejo, G. Bedoya, and J. Cabestany, "Hybrid neural networks for ISFET source separation," *Proc. SPIE*, vol. 5116, pp. 109–119, Apr. 2003.

[26] S. M. Cormier, G. W. Suter, and L. Zheng, "Derivation of a benchmark for freshwater ionic strength," *Environ. Toxicol. Chem.*, vol. 32, no. 2, pp. 263–271, Feb. 2013, doi: 10.1002/etc.2064.

[27] D. E. Culham, M. Meinecke, and J. M. Wood, "Impacts of the osmolality and the lumenal ionic strength on osmosensory transporter ProP in proteoliposomes," *J. Biol. Chem.*, vol. 287, no. 33, pp. 27813–27822, Aug. 2012, doi: 10.1074/jbc.m112.387936.

[28] A. Rudnitskaya, "Calibration update and drift correction for electronic noses and tongues," *Frontiers Chem.*, vol. 6, Sep. 2018, Art no. 433, doi: 10.3389/fchem.2018.00433.

[29] K. Yan and D. Zhang, "Calibration transfer and drift compensation of e-noses via coupled task learning," *Sens. Actuators B, Chem.*, vol. 225, pp. 288–297, Mar. 2016, doi: 10.1016/j.snb.2015.11.058.

[30] Z. Ahmad and J. Zhang, "Combination of multiple neural networks using data fusion techniques for enhanced nonlinear process modelling," *Comput. Chem. Eng.*, vol. 30, no. 2, pp. 295–308, Dec. 2005, doi: 10.1016/j.compchemeng.2005.09.010.

[31] W. Ji, V. I. Adamchuk, S. Chen, A. S. M. Su, A. Ismail, Q. Gan, Z. Shi, and A. Biswas, "Simultaneous measurement of multiple soil properties through proximal sensor data fusion: A case study," *Geoderma*, vol. 341, pp. 111–128, May 2019.

[32] F. Chen, D. Wei, and Y. Tang, "Virtual ion selective electrode for online measurement of nutrient solution components," *IEEE Sensors J.*, vol. 11, no. 2, pp. 462–468, Feb. 2011, doi: 10.1109/jsen.2010.2060479.

[33] A. Mueller, "Development of a combined multi-sensor/signal processing architecture for improved *in-situ* quantification of the charge balance of natural waters," Massachusetts Inst. Technol., Cambridge, MA, USA, Tech. Rep. 813047462-MIT, 2012.

[34] M. Gutiérrez, S. Alegret, R. Cáceres, J. Casadesús, O. Marfá, and M. D. Valle, "Nutrient solution monitoring in greenhouse cultivation employing a potentiometric electronic tongue," *J. Agricult. Food Chem.*, vol. 56, no. 6, pp. 1810–1817, Mar. 2008.

[35] D.-H. Jung, H.-J. Kim, H. Kim, J. Choi, J. Kim, and S. Park, "Fusion of spectroscopy and cobalt electrochemistry data for estimating phosphate concentration in hydroponic solution," *Sensors*, vol. 19, no. 11, p. 2596, Jun. 2019, doi: 10.3390/s19112596.

[36] W. H. Lee, Y. Seo, and P. L. Bishop, "Characteristics of a cobalt-based phosphate microelectrode for *in situ* monitoring of phosphate and its biological application," *Sens. Actuators B, Chem.*, vol. 137, no. 1, pp. 121–128, Mar. 2009, doi: 10.1016/j.snb.2008.10.032.

[37] L. I. Trejo-Téllez and F. C. Gómez-Merino, "Nutrient solutions for hydroponic systems," in *Hydroponics—A Standard Methodology for Plant Biological Researches*. Rijeka, Croatia: InTech, 2012.

[38] R. W. Kennard and L. A. Stone, "Computer aided design of experiments," *Technometrics*, vol. 11, no. 1, p. 137, 1969, doi: 10.2307/1266770.

[39] S. Berchmans, T. B. Issa, and P. Singh, "Determination of inorganic phosphate by electroanalytical methods: A review," *Anal. Chim. Acta*, vol. 729, pp. 7–20, Jun. 2012, doi: 10.1016/j.aca.2012.03.060.

[40] R. D. Marco, B. Pejcic, and Z. Chen, "Flow injection potentiometric determination of phosphate in waste waters and fertilisers using a cobalt wire ion-selective electrode," *Analyst*, vol. 123, no. 7, pp. 1635–1640, 1998.

[41] M. Stüber and T. Reemtsma, "Evaluation of three calibration methods to compensate matrix effects in environmental analysis with LC-ESI-MS," *Anal. Bioanal. Chem.*, vol. 378, no. 4, pp. 910–916, Feb. 2004.

[42] C. C. Rundle. (2000). *A Beginners Guide to Ion-Selective Electrode Measurements*. [Online]. Available: http://www.nico2000.net/Book/Guide1.html

[43] A. C. Olivieri, *Introduction to Multivariate Calibration: A Practical Approach*. Cham, Switzerland: Springer, 2018.

[44] S. Wold, H. Martens, and H. Wold, "The multivariate calibration problem in chemistry solved by the PLS method," in *Matrix Pencils*. Berlin, Germany: Springer-Verlag, 1983, pp. 286–293.

[45] M. Ebden, "Gaussian processes: A quick introduction," 2015, *arXiv:1505.02965*. [Online]. Available: https://arxiv.org/abs/1505.02965

[46] C. Rasmussen and C. Williams, *Gaussian Processes for Machine Learning*. Cambridge, MA, USA: MIT Press, 2006.

[47] M. Bos, A. Bos, and W. Van Der Linden, "Processing of signals from an ion-elective electrode array by a neural network," *Anal. Chim. Acta*, vol. 233, pp. 31–39, Jan. 1990.

[48] I. Basheer and M. Hajmeer, "Artificial neural networks: Fundamentals, computing, design, and application," *J. Microbiol. Methods*, vol. 43, no. 1, pp. 3–31, Dec. 2000, doi: 10.1016/s0167-7012(00)00201-3.

[49] H. Tang, K. C. Tan, and Z. Yi, *Neural Networks: Computational Models and Applications*. Berlin, Germany: Springer-Verlag, 2007.

[50] G. Horvai, K. Tóth, and E. Pungor, "A simple continuous method for calibration and measurement with ion-selective electrodes," *Anal. Chim. Acta*, vol. 82, no. 1, pp. 45–54, Mar. 1976, doi: 10.1016/s0003-2670(01)82202-9.

[51] R. K. Meruva and M. E. Meyerhoff, "Mixed potential response mechanism of cobalt electrodes toward inorganic phosphate," *Anal. Chem.*, vol. 68, no. 13, pp. 2022–2026, Jan. 1996, doi: 10.1021/ac951086v.

[52] S. O. Engblom, "Determination of inorganic phosphate in a soil extract using a cobalt electrode," *Plant Soil*, vol. 206, no. 2, pp. 173–179, 1999.

[53] V. O. Ebuele, D. G. Congrave, C. D. Gwenin, and V. Fitzsimmons-Thoss, "Development of a cobalt electrode for the determination of phosphate in soil extracts and comparison with standard methods," *Anal. Lett.*, vol. 51, no. 6, pp. 834–848, Apr. 2018.

[54] H.-J. Kim, W.-K. Kim, M.-Y. Roh, C.-I. Kang, J.-M. Park, and K. A. Sudduth, "Automated sensing of hydroponic macronutrients using a computer-controlled system with an array of ion-selective electrodes," *Comput. Electron. Agricult.*, vol. 93, pp. 46–54, Apr. 2013, doi: 10.1016/j.compag.2013.01.011.

[55] P. Zhang and Z. He, "Using data-driven feature enrichment of text representation and ensemble technique for sentence-level polarity classification," *J. Inf. Sci.*, vol. 41, no. 4, pp. 531–549, Aug. 2015, doi: 10.1177/0165551515585264.

[56] S. K. Srivastava, S. K. Singh, and J. S. Suri, "Effect of incremental feature enrichment on healthcare text classification system: A machine learning paradigm," *Comput. Methods Programs Biomed.*, vol. 172, pp. 35–51, Apr. 2019, doi: 10.1016/j.cmpb.2019.01.011.

[57] M. Baret, P. Fabry, D. L. Massart, C. Menardo, and M. Fraysse, "Calibration procedures for a halide ion-selective electrode array," *Analusis*, vol. 26, no. 8, pp. 267–276, Oct. 1998.

[58] N. García-Villar, J. Saurina, and H. Hernández-Cassou, "Potentiometric sensor array for the determination of lysine in feed samples using multivariate calibration methods," *Fresenius J. Anal. Chem.*, vol. 371, no. 7, pp. 1001–1008, Dec. 2001, doi: 10.1007/s002160101042.

[59] Z. Hu, Y. Jin, Q. Hu, S. Sen, T. Zhou, and M. T. Osman, "Prediction of fuel consumption for enroute ship based on machine learning," *IEEE Access*, vol. 7, pp. 119497–119505, 2019, doi: 10.1109/access.2019.2933630.

[60] V. Gagnon, G. Maltais-Landry, J. Puigagut, F. Chazarenc, and J. Brisson, "Treatment of hydroponics Wastewater using constructed wetlands in winter conditions," *Water, Air, Soil Pollut.*, vol. 212, nos. 1–4, pp. 483–490, Oct. 2010, doi: 10.1007/s11270-010-0362-8.

[61] M. I. Lone, Z.-L. He, P. J. Stoffella, and X.-E. Yang, "Phytoremediation of heavy metal polluted soils and water: Progresses and perspectives," *J. Zhejiang Univ. Sci. B*, vol. 9, no. 3, pp. 210–220, Mar. 2008, doi: 10.1631/jzus.b0710633.

VU NGOC TUAN received the B.S. degree in electrical and electronic engineering from the Ho Chi Minh City University of Technology and Education, Ho Chi Minh City, Vietnam, and the M.S. degree in automation from Le Quy Don Technical University, Hanoi, Vietnam. He is currently the Doctoral Fellow of agricultural electrification and automation with the College of Information and Electrical Engineering, China Agricultural University, Beijing, China. He is a member of the Key Laboratory of Agricultural Informatization Standardization, Ministry of Agriculture and Rural Affairs. His research interests include measurement and control systems, ion-selective electrode sensors, multisensor techniques, virtual sensors, and sensor data fusion.

TRINH DINH DINH received the B.Sc. degree in chemical technology from Hanoi University of Industry, in 2011, and the M.Sc. degree in chemical engineering from Military Technical Academy, Hanoi, Vietnam, in 2015. He is currently pursuing the Ph.D. degree with the College of Chemistry and Chemical Engineering, Beijing Institute of Technology, China. His research interests include development of ion-selective materials, and sensor membrane, nano applications in fertilizer, and nanomaterial for sensor application.

ABDUL MATEEN KHATTAK received the Ph.D. degree in horticulture and landscape from the University of Reading, U.K., in 1999. He was a Research Scientist in different agricultural research organizations before joining Agricultural University Peshawar, Pakistan, where he is currently a Professor with considerable experience in teaching and research at the Department of Horticulture. He has conducted academic and applied research on different aspects of tropical fruits, vegetables, and ornamental plants. He was also at Alberta Agriculture and Forestry, Canada, as a Research Associate, and with the Organic Agriculture Centre of Canada, as a Research and Extension Coordinator (for Alberta province), where he helped in developing organic standards for greenhouse production and energy saving technologies for Alberta greenhouses. He is also a Visiting Professor with the College of Information and Electrical Engineering, China Agricultural University, Beijing. He has published 55 research articles in scientific journals of international repute. His research interests include greenhouse production, medicinal, aromatic and ornamental plants, light quality, supplemental lighting and temperature effects on greenhouse crops, aquaponics, and organic production.

LIHUA ZHENG graduated from Beijing Agricultural Engineering University. She also received the Ph.D. degree in agricultural engineering from China Agricultural University, in 2007. She started her career as a Teacher at China Agricultural University, in July 1992. From March 1998 to May 1999, she worked at Yamanasi Software Co., Ltd., Japan, and developed two commercial systems for the company. From March 2008 to June 2008, she studied computer relevant project teaching at Portsmouth University, U.K. From February 2011 to February 2012, as a Visiting Scholar, she visited the University of Florida, USA. She has been dedicating herself to teaching and research. She has lectured 14 computer-related courses over the years. She is currently a Professor with the College of Information and Electrical Engineering, China Agricultural University. She is a Key Member of the Key Laboratory of Modern Precision Agriculture System Integration Research, Ministry of Education, China. She presided over five national scientific research projects in recent years. She concentrates on using computing technologies to resolve agricultural issues, specifically agricultural informatization, near-infrared spectral analysis for agricultural products, and application of the IoT in agriculture.

XIAOFEI CHU received the bachelor's degree from China Agricultural University, Beijing, China, in 2018, where he is currently pursuing the master's degree with the College of Information and Electrical Engineering. His research focuses on the method of soil treatment by non-thermal plasma.

WANLIN GAO received the B.S., M.S., and Ph.D. degrees from China Agricultural University, in 1990, 2000, and 2010, respectively. He is currently the Dean of the College of Information and Electrical Engineering, China Agricultural University. He is a Principal Investigator of over 20 national plans and projects, a member of the Science and Technology Committee of the Ministry of Agriculture, a member of the Agriculture and Forestry Committee of Computer Basic Education in Colleges and Universities, and a Senior Member of the Society of Chinese Agricultural Engineering. He has published 90 academic articles in domestic and foreign journals among them and over 40 are cited by SCI/EI/ISTP. He has written two teaching materials, which are supported by the National Key Technology Research and Development Program of China during the 11th Five-Year Plan Period, and has written five monographs. Moreover, he holds 101 software copyrights, 11 patents for inventions, and eight patents for new practical inventions. His research interests include the informationization of new rural areas, intelligence agriculture, and the service for rural comprehensive information.

MINJUAN WANG received the Ph.D. degree from the School of Biological Science and Medical Engineering, Beihang University, under the supervision of Prof. H. Liu, in 2017. She was a Visiting Scholar with the School of Environmental Science, Ontario Agriculture College, University of Guelph, from 2015 to 2017. She is currently an Associate Professor with the School of Information and Electrical Engineering, China Agricultural University. Her research interests include bioinformatics and the Internet of Things key technologies.