

Improving Ion-Sensitive Field-Effect Transistor Selectivity with Backpropagation Neural Network

WAN FAZLIDA HANIM ABDULLAH^{1,3}, MASURI OTHMAN², MOHD ALAUDIN MOHD ALI³,
MD SHABIUL ISLAM³

¹Fakulti Kejuruteraan Elektrik, Universiti Teknologi MARA, Shah Alam,
40450 Selangor, MALAYSIA

²MIMOS Berhad, Taman Teknologi Malaysia,
57000 Kuala Lumpur, MALAYSIA

³Institute of Microengineering and Nanoelectronics, Universiti Kebangsaan Malaysia,
43600 Bangi, Selangor, MALAYSIA

wanfaz@salam.uitm.edu.my <http://www.uitm.edu.my>

Abstract: - The ion-sensitive field-effect transistor (ISFET) that is designed to detect a specific ionic activity is susceptible to interfering ions in mixed-ion environments causing the sensor to produce deceptive signals. The objective of this work is to improve the interpretation of ISFET signals in mixed-ion environments. The focus of the research is relating sensor signal to the targeted ion concentration by applying supervised neural network as post-processing stage as a method to overcome low selectivity issues. In this paper, we acquire ISFET voltage response data in potassium and ammonium mixed-ion solutions for the training of a multilayer perceptron with backpropagation algorithm. A constant-voltage constant-current readout interface circuit is applied to maintain constant bias of the sensor throughout the data collection process. Primary data from measured observations was fed to a feed-forward multilayer perceptron trained to classify levels of ionic concentrations in various levels of mixed-ion solutions. Accuracy of sensor response interpretation of ionic activity estimation is compared between with and without neural network post-processing stage. Neural network performance was also compared for voltage values with and without pre-processing voltage signals by referencing sensor response in deionized water. Further improvement of the network was approached by using an ensemble of similar structures of networks trained with backpropagation constructed using the bagging algorithm. Results show that neural network fed with dc voltage response from 4-sensor array is able to improve concentration estimation by 15% improvement compared to direct estimation based on a look-up table. Pre-processing the sensor response significantly improves the sensor signal repeatability correlation factor by 15.5% and reduces mean-square error by 98.3%, with a typical 20% improvement in output-target regression factor network performance. Averaging from ensemble system is shown to give a further 5% improvement on the output-target regression factor with consistently stable ion concentration estimations.

Key-Words: - microsensors, electrochemical devices, MOSFET, sensor array, supervised learning, selectivity

1 Introduction

The ion-sensitive field-effect transistor (ISFET) is an electrochemical sensor that produces electrical response in accordance to ionic concentration due to the ionic activity at the exposed gate window. ISFETs are potentiometric sensors that produce response by virtue of potential reaction at the electrolyte/membrane interface similar to the more frequently used ion-selective electrode [1]. The membrane at the gate window acts as the receptor and the basic structure of a metal-oxide field-effect transistor functions as a transducer [2, 3]. The construction of ISFET requires the fabrication of metal-oxide field-effect (MOSFET) transistor thus giving ISFET the advantage of being solid-state silicon based and compatible with standard

MOSFET fabrication technology thus opening the door of mass-production techniques and miniaturization benefits to chemical sensing [4-6]. Sharing the same silicon platform is an added convenience for integrating sensing and computational modules in smart sensor systems that are portable. This makes ISFET an appealing technology for environmental, agricultural and clinical applications that require traditional laboratory analysis to be available on site.

In view of positioning the ISFET as the candidate of choice for chemical sensing, the ISFET will have to be proven reliable and robust under specified physical conditions and different chemical environments within the area of application. One of the challenges for the ISFET sensor is the need to

demonstrate high selectivity [7]. Selectivity is the ability to respond to primarily only one ion species in the presence of other species. In the presence of mixed ions of equal charge number and similar ionic radii to the main ion of interest, ISFETs exhibit response towards the interfering ionic activity. The sensor signal in mixed-ion solution would then represent combined activity rather than information from a single ion type [8-10]. The common approach for sensor selectivity is to find the selectivity coefficient. The standard procedure for potentiometric sensors are provided by the International Union of Pure and Applied Chemistry (IUPAC) for ion-selective electrodes that is applicable to ISFETs [11]. The approach involves parameter estimation and interpolation of chemical data.

For chemical sensors, neural networks can enhance sensor performance and allow control of area of applications [12]. For ionic sensors, post-processing stage involving machine learning has been proposed to estimate ionic concentration change and to extract the main ion activity from the mixed response such as blind source separation techniques [13-15] that requires time-dependent voltage response. Pattern recognition methods are also used with an array of sensors providing a series of input features for classification [16].

In this work, ISFET voltage response is obtained for the purpose of recognizing the potassium ion (K^+) logarithmic value of concentration in the presence of ammonium ion (NH_4^+) in ranges of concentration level typical to agricultural surroundings. In contrast to the other methods involving machine learning, this work handles dc output response of the sensors as constant averaged values independent of time. Sensor voltage response was acquired to act as input data while prepared sample concentrations from standard calculations was used as target for training data. Sensor response are captured from a readout interface circuit that satisfies the need of a MOSFET biasing, in an array of 4 sensors, of K^+ and NH_4^+ types. Measurement and data acquisition is setup to provide the training data for feedforward neural network back-propagation training algorithm.

2 Problem Formulation

The post-processing stage experimental method includes primary data acquisition and the formation of the artificial neural network. Training data collection is planned to cover the required range of concentrations for ionic sensors. The work is multidisciplinary from semiconductor theory for

field-effect transistor device and electronics instrumentation, to training data collection involving electrochemistry for chemical sensing, as well as neural network architecture and learning algorithm.

2.1 Device Measurements

ISFETs are commonly modelled with the electrical characteristics based on the following standard MOSFET drain-source current, I_{DS} , equation [17]:

$$I_{DS} = \mu C_{ox} \frac{W}{L} \left((V_{GS} - V_{TH}) V_{DS} - \frac{V_{DS}^2}{2} \right) \quad (1)$$

where W/L is channel width/channel length of gate area, μ is surface mobility, C_{ox} is oxide capacitance per unit area, V_{TH} is device threshold voltage, V_{GS} and V_{DS} are gate and drain applied biases respectively.

Fig.1 depicts ISFET circuit connections resembling a typical MOSFET biasing setup except that the gate with ion selective membrane is not directly biased. Instead, the membrane is exposed to the ionic solution with a reference electrode in the setup for signal return and influencing V_{GS} . With the source connection grounded, the voltage applied to the reference electrode, V_{ref} , represents V_{GS} . Electrochemical effects at the membrane/electrolyte interface alter the flatband voltage thus causing the ISFET threshold voltage, $V_{thsensor}$, to be modified [3, 17]:

$$V_{thsensor} = V_{th} - E_{chem} \quad (2)$$

where V_{th} is the device threshold voltage without membrane and E_{chem} is the electrochemically induced voltage.

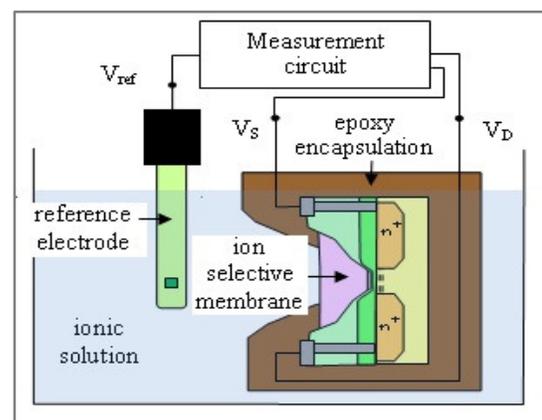


Fig. 1: Schematic diagram of ISFET under operational conditions. [18]

Based on the concepts of electrochemistry, E_{chem} is related to ionic species as governed by Nernst and

to the interfering ion by Nikolsky equation which is summarized into the following equation:

$$E_{chem} = E_o \pm 2.303 \frac{RT}{z_i F} \log_{10} \left(a_i + \sum_j K_{ij} a_j^{z_i/z_j} \right) \quad (3)$$

where subscripts ‘i’ refer to the main ion of interest, subscripts ‘j’ to the interfering ion, E_o is potential at 1 mol/dm³ ionic activity, R is gas constant, F is Faraday constant, T is temperature, a is ionic activity, K_j is potentiometric selectivity coefficient and z are charge numbers [19]. Incorporating the electrochemical effects threshold voltage modifications to (1) results in the following expression [20]:

$$I_{ds} = \mu C_{ox} \frac{w}{L} \left((V_{ref} - (V_{th} - E_{chem})) V_{ds} - \frac{V_{ds}^2}{2} \right) \quad (4)$$

The readout circuit employed to track the threshold voltage modifications due to the ionic activity at the membrane/electrolyte interface is given in Fig. 2 [21, 22]. Based on isothermal point of operation, I_{ds} , V_{ref} and V_{ds} are set at 100 μ A, 0 V and 0.5 V respectively. E_{chem} is monitored based on the output of the readout circuit that is connected to the source of the transistor. To ensure only the electrochemically induced voltage is monitored, the circuit maintains constant biasing of the sensor. The current sources for source and sink ensure constant drain current whereas followers and the bias resistor ensure constant voltage across the drain-source for isothermal point operation. E_{chem} due to ionic activity changes is monitored by the source voltage fluctuations of the transistor sensor. The voltage adjustment makes use of the operational amplifiers characteristics where inputs track each other’s potential as well as infinite input impedance that draws negligible current.

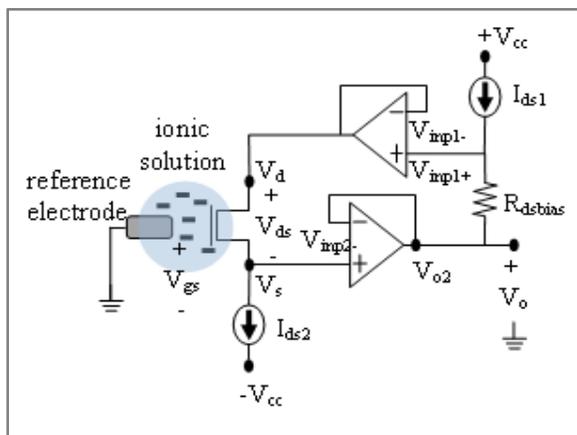


Fig. 2: Readout interface circuit employed in measuring sensor response.

2.2 Training Data Collection

Training data is collected from K^+ and NH_4^+ sample solutions between -6 to -1 log of ion concentration by mixing potassium and ammonium chloride with deionized water (DIW). The logarithmic values of potassium ion concentration (pK) and of ammonium ion concentration (pNH₄) are used; the intended target data format is {pK, pNH₄}. The electrical response for the particular target data is taken from an array of sensors through interfacing circuits that would function as time-independent input data { V_1, V_2, \dots, V_m } to a neural network post-processing stage.

The goal of sample preparation is to represent the input space within the detection limit of available K^+ and NH_4^+ ISFET. The chosen ions are known to be most interfering with each other having the same charge type and charge number with the closest ionic radii size. The reagents used as ionophores, an organic compound that facilitates the transport of ions across the cell membrane, are valinomycin (for K^+ sensors) and nonactin (for NH_4^+ sensors). The reference electrode used is a silver/silver chloride (Ag/AgCl) double junction

Sample preparation is based on IUPAC recommendations for fixed interference method that varies one type of ion activity by adding titrants to a total solution while keeping the other ion unchanged [11, 23]. Required volumes in the titration process is calculated from dilution equation and the log of activity of cation, γ , based on the Debye-Huckel approximation:

$$I_{total} = \frac{V_o \cdot I_o + \sum_i V_{titrant1,i} \cdot I_{titrant1,i} + \sum_i V_{titrant2,i} \cdot I_{titrant2,i}}{V_{total}} \quad (5)$$

$$\log (\gamma) = - \left(0.509 * \frac{\sqrt{I_{total}}}{1 + (2.5 * 0.328 * \sqrt{I_{total}})} \right) \quad (6)$$

where ‘I’ is ionic strength, ‘V’ is volume of solution, subscripts ‘o’ for initial sample, ‘titrant’ and ‘total’ indicates added titrant and total solution respectively [24]. Equations (3) and (4) are applied in achieving the planned training data consisting of main ion of interest and interfering ion.

Fig. 3 demonstrates the sample preparation strategy for training data collection. For the chosen type of ion (either K^+ or NH_4^+) for the initial solution, seven different levels of ionic strength initial solutions are prepared with {pK|pK \in Z, -6 \le pK \le -1} and one control solution represented by DIW.

The second type of ion is added to each of the initial solutions 11 times by pre-calculated volume

of titrants based on (5) and (6) successively of which some are control values outside sensor detection limits. This results in 12 possible values of varying pK in fixed pNH₄ and vice versa. This approach results in a total of 168 samples including 48 control samples. The setup is repeated on different days using different sensors for investigations on repeatability and reproducibility. For continuous reading throughout the titration of one set on solution, the output of the interface circuit is fed to a data logger NI DAQ USB6251 as shown in Fig. 4. The system communicates with the PC and record readings at 10 Hz sampling rate.

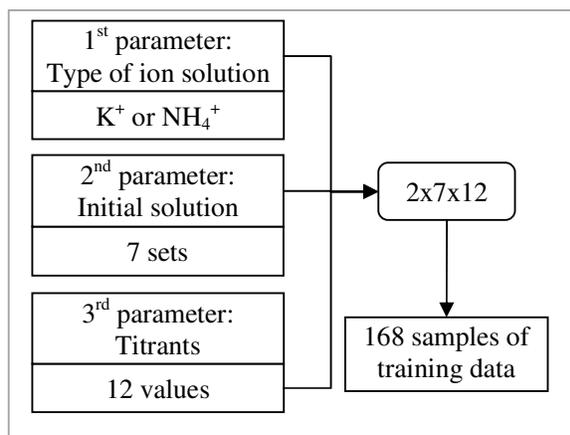


Fig. 3: Design parameters for sample preparation strategy to collect training data

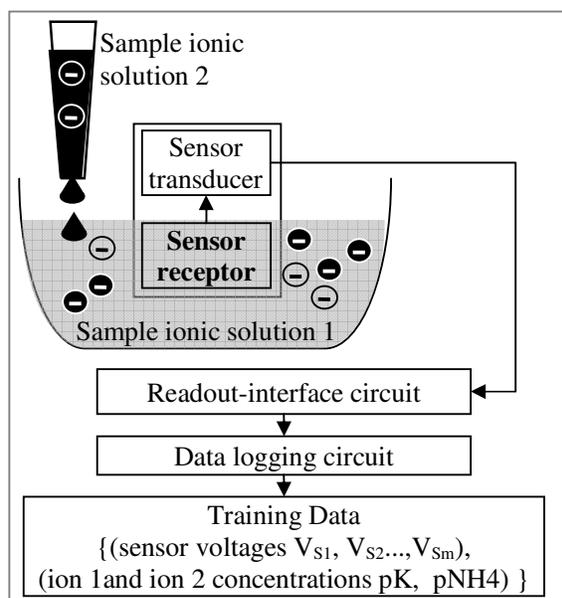


Fig. 4: Schematics of training data collection measurement setup

2.3 Neural Network Design

The standard method in overcoming the selectivity issue is by chemical approach involving calibration and graphical extrapolation involving mathematical model equations for chemical phenomenon requiring pre-determined device characteristic values. The role of the neural network to be designed is to provide a method to relate electrical response to chemical information that are free of graphical approach and device modelling equations and to be able to tolerate parameter variations. It is hypothesized that a supervised neural network as a post-processing stage can improve the estimation of the main ion concentration in the presence of interferants.

A feedforward multilayer perceptron (MLP) architecture with hidden layers is constructed with 4-input dimension for the sensor array of potassium and ammonium sensor with 2-output values of K⁺ and NH₄⁺ molarity. The network synaptic weights and bias free parameters are randomly initialized prior to training within -1 and 1. The trained neural network is intended to function as post-processing stage to the sensor signal for the estimation of ionic concentration from an ionic solution as demonstrated in Fig 5.

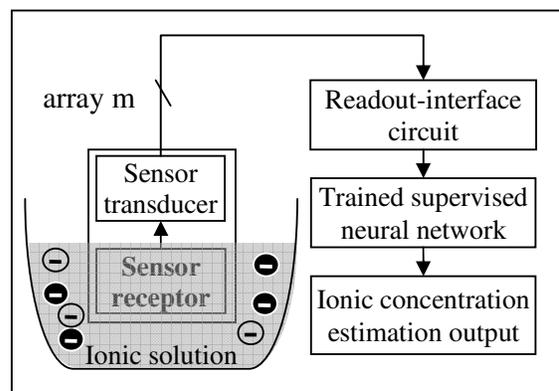


Fig. 5: Schematics of training data collection measurement setup

2.3.1 Backpropagation Neural Network

The network is trained with supervised learning back-propagation algorithm. Training, testing and validation data allocation size are set to 60%, 20% and 20% of the total data from 398 measurements respectively. The hidden layer uses a hyperbolic tangent activation function. The activation function of the output neuron is linear [25]. The general back-propagation algorithm applied includes a forward computation where the outputs of each neuron from the hidden layer to the output layer are

found based on the synaptic weights and activation function. Then backward computation that propagates the error signal from the output layer to the hidden layer takes place. In this process the local gradients for each neuron are calculated and the weights are adjusted based on the steepest descent method as in the following [26]:

$$\Delta w(n) = w(n+1) - w(n) = -\eta \nabla \xi(w) \quad (7)$$

where ξ is cost function and η is the learning parameter. The instantaneous value of the cost function takes the mean square error value.

$$\xi(w) = \frac{1}{2} e^2(n) \quad (8)$$

where e is error and n refers to the n -th sample.

Comparison with more back-propagation variations using resilient backpropagation [27], scaled conjugate gradient algorithms, Quasi-Newton [26] and Levenberg-Marquadt algorithms [28] were done using standard algorithms provided by the Matlab Neural Network toolbox [29].

2.3.2 Multiple Feedforward System

Multiple classifiers of the same MLP architecture and learning algorithm were constructed with variations achieved through bootstrap aggregating resampling methods and randomness of free parameter initial values. The objective is to create a multiple classifier system based on feedforward multilayer perceptron to reduce the risk of relying on a poorly trained classifier. Multiple outputs were combined through averaging [30].

Fig. 6 is the algorithm flowchart implemented in Matlab to generate individual classifiers and creating diversity between classifiers by using bootstrapped replicas of the training data. 'F' represents percentage of samples randomly drawn from the pool data obtained and replaced for the next training. 'T' is the specified total number of classifiers in the system. Each classifier would have different subsets of training data by bootstrapping as well as initial free parameter weights and bias values for training [31, 32].

With all individual 'T' MLPs formed and trained, the multiple classifier system is created. Each of the individually trained classifier values are called in parallel for individual output classification. During decision making, all T classifier decides on an output individually. The outputs would then be either averaged for regression continuous values or vote for classification problems. For classification problems, each classifier votes for an output

category and the votes are counted using simple majority voting. The output category with the maximum number of votes is chosen as the output of the system.

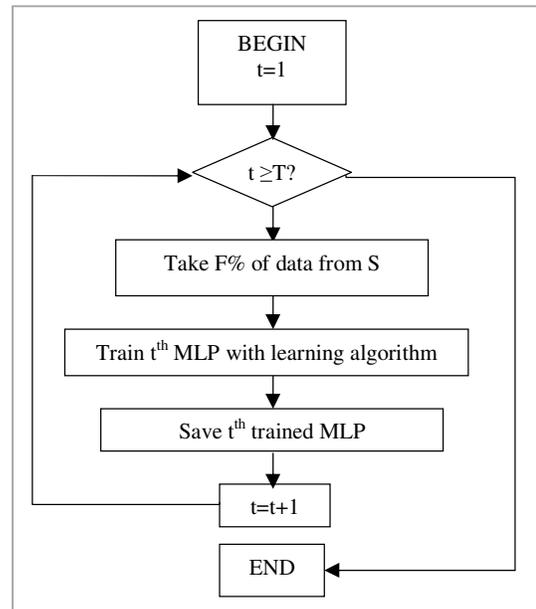


Fig. 6: Flowchart for generating individual classifiers for T multiple classifiers by bagging.

3 Problem Solution

Results are presented first by looking at the sensor performance from the training data collected followed by neural network performance having been trained with the measured data. Comparison is made between applying neural network processing stage and lookup table along with the effectiveness of pre-processing the sensor responses prior to the network that has been trained by backpropagation.

3.1 Sensor Performance

Fig.8 shows the measured I_{ds} - V_{gs} transfer characteristic of K^+ ISFET in K^+ solution with the transconductance curve superimposed with the second y-axis. For the sake of comparing effects of different K^+ concentrations on the characteristics, two different solutions of 10^{-3} and 10^{-5} M of K^+ solutions are considered. It can be observed that stronger ionic concentration causes the transfer characteristics to be shifted to the left, indicating lower threshold voltage. V_{th} is calculated based on the following equation [33]:

$$V_{th} = V_{gsmaxS} - 0.5 V_{ds} - (I_{dsmaxS}/S_{max}) \quad (9)$$

where V_{gsmaxS} is gate voltage at maximum slope, I_{dsmaxS} is current at maximum slope and S_{max} is the calculated maximum slope. The information at the point of maximum slope for the V_{th} calculation is presented in Table 1. From the values, it can be concluded that the presence of higher ionic activity electrochemically induces higher voltage at the membrane/electrolyte interface thus lowers the threshold voltage of the sensor.

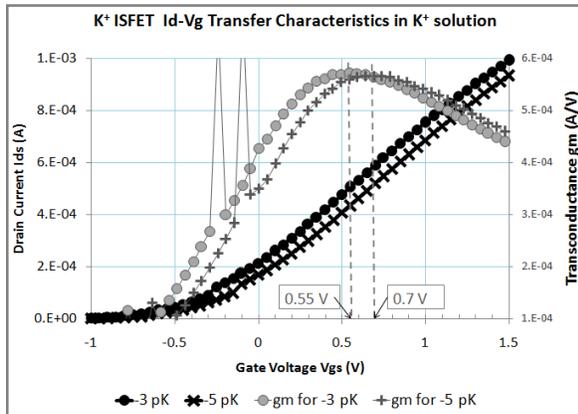


Fig. 7: K⁺ ISFET transfer characteristics in 10⁻³ and 10⁻⁵ M K⁺ solutions

Table 1: V_{th} calculation based on values extracted from transfer characteristics.

Parameters	In 10 ⁻³ K ⁺	In 10 ⁻⁵ K ⁺
V_{gsmaxS} (V)	0.55	0.7
I_{dsmaxS} (A)	0.000504	0.000520
S_{max} (S)	0.000529	0.000502
V_{th} (V)	-0.653	-0.586

Data points that are 3 standard deviations away from the mean are labelled as outliers and removed using standard mathematical programming in Matlab. Data is also cleaned using 1st order lowpass digital Chebyshev type I filter with 1 dB ripple in the passband with normalized passband edge frequency of 0.005. Reasons for the low edge frequency value are based on the nature of data requiring only the capture of step response dc voltage associated with concentration change. Sensor rise is found to have a maximum of 40 s from the readings thus subsequent measurements for neural network training data collection should allow

40 s for sensor response settling time. It is also found that the average sensitivity is 40 mV per order of K⁺ molarity. Graphs in Fig.8 to 11 provide visual representation of measured sensor response towards main and interfering ionic activity changes across the training data sample solutions. The four figures present overall K⁺ or NH₄⁺ sensor response towards K⁺ and NH₄⁺ ionic activity collectively, each from 6 titrations sets and each with 11 consecutive additions. Fig.8 and 9 present sensor response to interfering ions while Fig.10 and 11 present sensor response to the intended ion. All the figures mentioned show that the sensors are cross-sensing to a certain extent depending on the interfering ion. Sensors in Fig. 8 and 9 are seen to be responsive toward interfering ion when interfering ionic activity values begin to exceed the main ion by approximately an order of magnitude. For the sake of discussion here, the simple method of intersection between linear response and horizontal line is used to determine lower limit of detection from the graph. In Fig.8, it can be found that in 10⁻⁴ M K⁺, lower limit of NH₄⁺ detection for linear response is -3.5 log of ionic activity. As the main ion concentration gets lower to 10⁻⁶ M K⁺, lower limit of NH₄⁺ detection is also lower at -5 log of ionic activity indicating a more responsive environment towards the interfering ion. At a high main ion concentration of 10⁻¹ M there is no response towards the comparatively lower NH₄⁺ ranges introduced during titration hence the non-responsive curve. The NH₄⁺ sensor towards K⁺ interfering ion is likewise shown in Fig.9.

Fig.10 and 11 show sensor behaviour towards the intended ion designed to be detected. Referring to Fig.10, in 10⁻⁶ M interfering ion NH₄⁺, lower limit of detection is -5 log of K⁺ ionic activity. At a higher level of 10⁻² M interfering ion, the limit of detection is -4 log of K⁺ ionic activity. For both the K⁺ and NH₄⁺ sensors, interfering ion is seen to increase the linear response lower detection limit. In other words, interfering ion causes the sensor to lose sensitivity towards the intended ion by increasing the minimum level of detectable main ion. This behaviour conforms to the model equation (3) where the electrochemically induced potential depends on the term $\left(a_i + \sum_j K_{ij}^{pot} a_j^{z_i/z_j} \right)$. When the interfering ionic activity a_j is larger than the main ion a_i , changes in a_i is not significant over the value of a_j to the overall expression of E_{chem} .

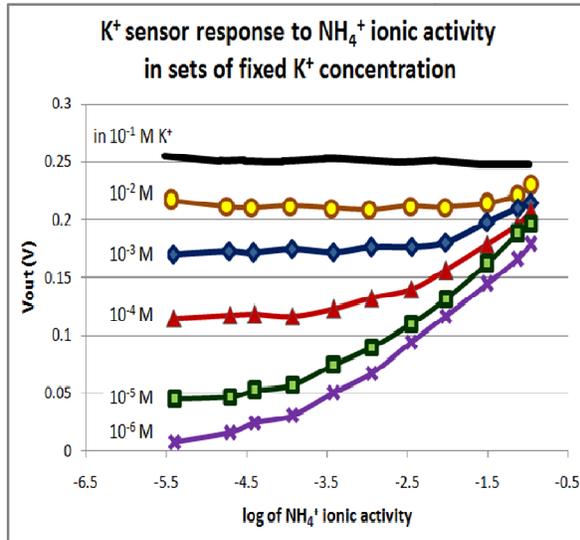


Fig.8: K⁺ ISFET collective response vs. interfering NH₄⁺ ionic activity, from 6 titration sets of initial solutions 10⁻⁶ to 10⁻¹ M K⁺.

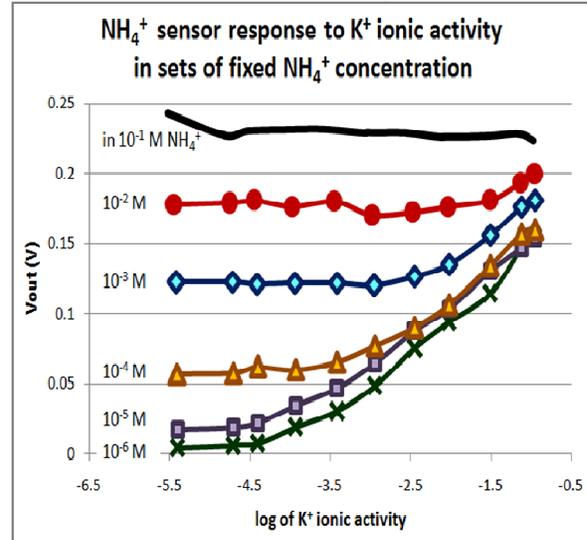


Fig.9: NH₄⁺ ISFET collective response vs. interfering K⁺ ionic activity, from 6 titration sets of initial solutions 10⁻⁶ to 10⁻¹ M NH₄⁺.

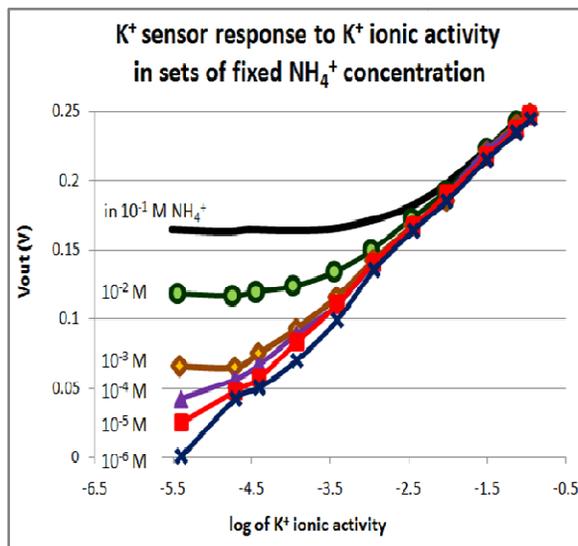


Fig.11: K⁺ ISFET collective response vs. intended K⁺ ionic activity, from 6 titration sets of initial solutions interfering NH₄⁺ 10⁻⁶ to 10⁻¹ M.

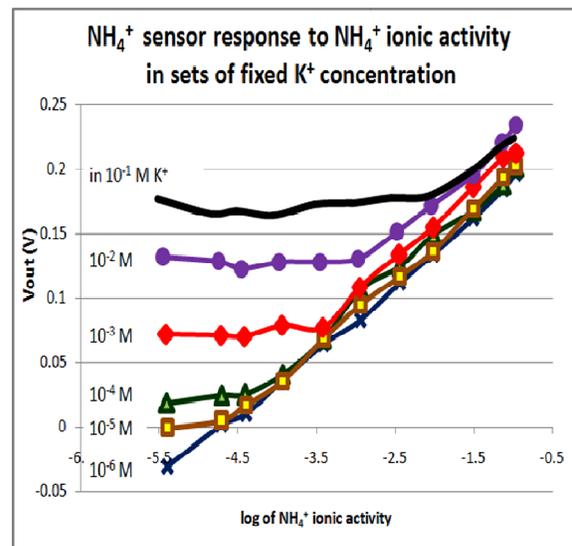


Fig.10: NH₄⁺ ISFET collective response vs. intended NH₄⁺ ionic activity, from 6 titration sets of initial solutions interfering K⁺ 10⁻⁶ to 10⁻¹ M.

3.2 Training Data Quality

Repeated measurements show that the sensor response demonstrates slight variation due to many possible sources such as light, temperature, device fabrication, membrane characteristics, membrane lifetime, reference electrode behaviour and ambient electrical noise. It is also found that despite keeping

the controllable factors such as light and temperature constant, the sensor demonstrates sudden DC voltage shifts as large as 0.7 V with subsequent readings maintained at this level. Repeatability and reproducibility are badly affected without pre-processing the values. Repeatability refers to successive runs made with the same sensor;

reproducibility refers to sensor dissimilarity between similarly fabricated sensors. Readings are taken from identically prepared different samples, sensors, and days of measurement. The sensors are assumed to be from the same recipe of fabrication process and membrane preparation.

As a preprocessing step on the sensor signal prior as inputs to the network, the voltage values as input data are referenced to the sensor response in DIW taken prior to the measurement instead of direct values of the sensor in the ionic solution [34]. To negate the unaccounted sudden DC offset, the voltage values as input data in the training set are taken with respect to the response of the sensor in DIW, V_{DIW} , measured prior to the measurement instead of raw measured values in ionic solution, $V_{solution}$. The training data input values, $V_{inputdata}$, are pre-processed by referencing as follows:

$$V_{inputdata} = V_{solution} - V_{DIW} \quad (10)$$

Any variations due to noise and behaviour of the sensor and reference electrode included would exist in both values hence would be negated in (10). This is done assuming that sensitivity is not affected in the DC shifts, as shown as far the measurements were done, and that offsets are limited to DC voltage shifts as observed in the measurements taken.

Fig.12 illustrates the improvement by referencing on repeated sensor response that had exhibited DC shifts. 6 repetitions were carried out on the same titration setup, consisting of $NH_4^+ 10^{-3}$ M initial solution with K^+ as titrant, using K^+ sensors of the same fabrication recipe. The sensors had been conditioned sufficiently in 10^{-2} M with bias but had shifted its DC response during the third repetition. Between the 6 repetitions, there is a maximum difference of 0.78 V between the readings. However, with referencing, the difference due to the shift is omitted. The maximum difference is now reduced to 0.08 V, reducing a significant amount of 89.7% of the variation.

Tables 2 and 3 compare the effect of referencing the voltage value by comparing the mean square error and regression factor respectively. Both tables compare values with and without referencing and lists percentage of improvement the better approach offers. MSE in Table 2 refers to the error between identical setups that should be resulting equal values of response, with an ideal value of 0. Regression in Table 3 refers to the similarity in data between repeated sets with one set acting as a reference set and the other as a test set, with an ideal value of 1. It is found that the mean square error (MSE) is improved for all sets of comparison when voltage

values are referenced to sensor response in DIW. For sensors demonstrating most of the shifts, repeatability is improved by an average of 98.3% for mean square error and pushing correlation to above 0.9 for every repetition. Even for sensors that did not demonstrate sudden DC shifts, the voltage variation is still improved by an order of magnitude.

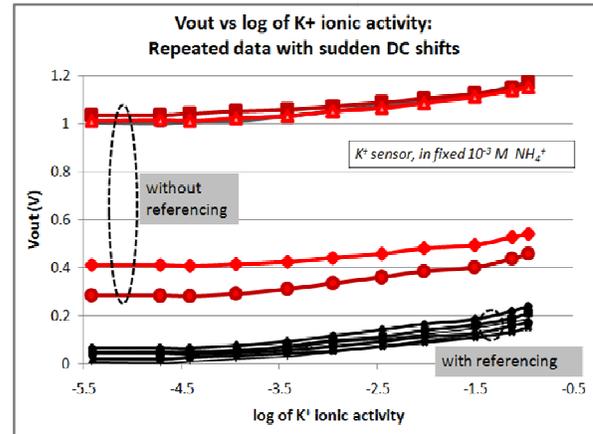


Fig.12: K^+ ISFET response to increasing K^+ ionic activity in fixed $NH_4^+ 10^{-3}$ M solution, performed in 6 repetitions.

Table 2: Effects of referencing to repeatability by comparing MSE of repeated sets.

Samples	MSE Direct Values	MSE with DIW response	Improvement
Separate samples	0.37	0.0011	99.7%
Different K^+ sensors	0.0044	0.0009	79.5%
Different NH_4^+ sensors	0.0074	0.0005	93.2%
Different days	0.24	0.008	96.7%
Average	0.16	0.0026	98.3%

Table 3: Effects of referencing to repeatability by comparing regression factor of repeated sets.

Comparison	R-factor Direct Values	R-factor Referenced value	Improvement
Different samples	0.52	0.92	40%
Different K^+ sensors	0.99	0.99	-
Different NH_4^+ sensors	0.98	0.98	-
Different days	0.70	0.92	22%
Average	0.80	0.95	15.5%

For the purpose of comparing measured data and simulated data, a single classifier is subjected to two sets of data as compared in Table 4. Simulated data is based on equation (4) above. In this case, the neural network is to approximate the potassium level by fitting function based on an array of sensor response. Neural network performance with simulated data is much better compared to actual data. This shows that the noise content in measured data affects the network ability to interpret sensor response and degrades its ability to relate it to ionic activity levels affectively. As expected, the single classifier can only learn weakly based on measured data due to background ion in the ionic solution, behaviour of reference electrode and membrane as well as noise from the environment. On top of the chemical environment, the transistor is mass produced in a semiconductor technology fabrication process that would itself have up to 5% of allowed device variation across the wafer.

Table 4: Network performance with measured and simulated data

	Measured Data	Simulated Data
Epoch	42	20
Regression factor	0.8065	0.98
MSE	0.15	0.015

3.3 Improving Interpretation of Sensor Response with Neural Network

Table 5 and 6 present performance of a feedforward network trained with gradient descent backpropagation with momentum and adaptive learning rate in estimating main ion concentration from a 4-sensor array from 317 samples. The 317 samples consist of response in various mixed-ion environments as shown in Figures 8 to 11. The network ability to interpret sensor response in terms of ionic concentration levels indicates the effectiveness of the network in overcoming the selectivity issue. It also shows the ability of the network to learn the sensor response without having to rely on neither semiconductor device characteristic parameters, as required by standard MOSFET current/voltage expressions, nor graphical approach, as required by standard potentiometric sensor electrochemistry approach.

Table 5 presents the regression factor between network output and target values; an ideal case

would be close to 1. Table 6 presents the mean square error between the output and target values across the set; the smaller the value, the better. Both tables allow comparison of three approaches: (i) application of neural network post-processing stage vs. lookup table (ii) pre-processing with reference to sensor performance in DIW prior to test sample reading vs. direct values (iii) varying number of hidden neurons.

Referring to the performances of networks with different numbers of hidden neurons between 5 to 25 and layers 1 to 3, the MLP 2 layer with 15 hidden neurons is a safe choice without increasing too much complexity. The lookup table entry provides a reference point for neural network post-processing stage to prove its effectiveness in improving the estimation of ion concentration.

For the construction of the lookup table, a set of values relating sensor voltages to known K⁺ ion concentration is generated based on the IUPAC standard with low interfering selectivity. The matching K⁺ concentration is selected based on the lookup sensor array voltage that has the least difference between the tested sample and the lookup sensor array voltages.

Table 5: Performance by regression factor in estimating main ion concentration in the presence of interfering ions

Approach		Output-Target R-Factor	
		no pre-processing	with pre-processing
With neural network	Single layer	0.22	0.68
	MLP 2 layer- 5	0.18	0.62
	MLP 2 layer- 10	0.57	0.71
	MLP 2 layer- 15	0.66	0.71
	MLP 2 layer- 25	0.58	0.72
	MLP 3 layer- 5-10	0.478	0.70
	MLP 3 layer - 10-15	0.43	0.71
	MLP 3 layer - 15-10	0.63	0.72
	MLP 3 layer - 30-25	0.58	0.73
Lookup table		0.33	0.72

Table 6: Mean-square error in estimating main ion concentration in the presence of interfering ions

Approach		Output-Target MSE	
		no pre-processing	with pre-processing
With neural network	Single layer	1.31	0.542
	MLP 2 layer- 5	0.951	0.473
	MLP 2 layer- 10	0.672	0.469
	MLP 2 layer- 15	0.651	0.485
	MLP 2 layer- 25	0.624	0.527
	MLP 3 layer - 5-10	0.638	0.459
	MLP 3 layer - 10-15	0.723	0.455
	MLP 3 layer - 15-10	0.857	0.438
	MLP 3 layer - 30-25	0.66	0.444
Lookup table		5.712	1.276

As seen in Table 5 values for regression factor between predicted values and target values are improved by 44.7% (from 0.33 to 0.66) with the application of neural network of 2 layers with 15 hidden neurons. A further improvement of 58.2% (from 0.33 to 0.72) is achieved with the referencing to sensor response in DIW which is as good as the neural network performance.

However, referring to Table 6, the mean square error of lookup table is 62% (1.276 compared to 0.485) larger than neural network performance. Mean square error of lookup table estimation without pre-processing is unacceptably larger by 88.6% (5.712 compared to 0.651). Comparing effects of preprocessing by referencing the sensor voltages for the case of 2-layer 15 hidden neuron performance, it is clearly evident that referencing the sensor voltages improve regression factor by 14.7% (from 0.66 to 0.71) and reduces the mean-square error by 25.5% (from 0.651 to 0.485).

Table 7 presents the performance of different back-propagation algorithms in the Neural Net toolbox in Matlab on the measured data. The architecture is fixed at 2 layer with 15 hidden

neurons. The Levenberg Marquadt algorithm clearly provides faster learning with lesser number of epochs required to reach the same specified goal.

Table 7: Comparison of backpropagation algorithm variations performance in ion estimation

	MSE	Epoch	R
Batch Gradient descent with momentum.	0.51	89	0.68
Adaptive learning rate.	0.60	54	0.69
Resilient bp.	0.37	20	0.73
Scaled Conjugate Gradient	0.553	13	0.69
Quasi newton. One Step Secant Algorithm	0.436	23	0.40
Levenberg Marquadt.	0.44	9	0.76

3.4 Multiple Decision by Bagging and Voting

Table 8 compares the performance of classification between weak and strong K⁺ molarity in the presence of varying NH₄⁺ weak to strong molarity between 3 cases; using lookup table (no neural network), single classifier system and multiple classifier system. Multilayer perceptron feed-forward neural network with single hidden layer was able to estimate test data with 15% improvement over direct estimation without neural network post-processing. Further consideration of the best-case performance from multiple classifier voting gives a further 4% increase in performance.

Table 8: Performance based on percentage of correct classification.

	% of correct classification	% of improvement
Lookup-table	70.362%	-
Single Classifier	80.935%	15.02%
Multiple Classifier	83.897%	19.24%

Table 9 shows that the system output performs better than the worst of classifiers in the system all the time. 80% of the 15 runs results in the voted system output giving an average of 3% improvement compared to individual classifier average performance. 13.3% of the 15 runs results in the system output performing better than the best of the individual classifiers. This is the effect of seeking the opinion of multiple classifiers.

Table 9: Performance of ensemble compared to individual classifier

Multiple classifiers compared to single classifiers	%
Better than worst?	100
Better than average?	80
Better than best of all single classifiers?	13.33

A multiple classifier system performance with 10 single hidden layer MLP with 15 hidden neurons each is shown in Fig. 13. The graph demonstrates the performance of the ensemble with variation solely on initial weight and bias value randomness of single hidden layer MLP in comparison to single classifier performance across 15 runs of test data. The ensemble performance is seen to be able to perform better than the average of the classifiers. The ensemble is able to avoid unpredictable weak estimations in regression by a single classifier thus improving performance stability.

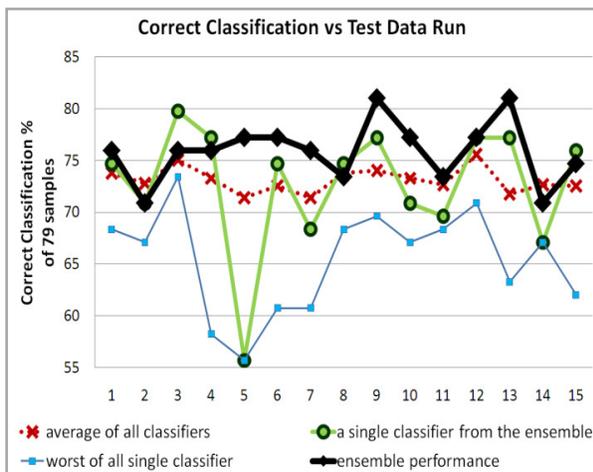


Fig. 13: Stability of performance across all samples

4 Conclusion

Neural network post-processing stage is shown to perform classification of main ion concentration in the presence of interfering ion from weak to strong from ISFET voltage response. Results corroborate the implementation of neural network post-processing towards improving the accuracy of device sensor reading interpretation as compared to estimation based on lookup table. It is also found that referencing sensor voltage signal to response in DIW is able to improve quality of training data in terms of repeatability and reproducibility. Thus the learning and performance of neural network is also improved with the pre-processing of the sensor signal. Further improvement is achieved by a multiple classifier system consisting of single classifier variation based on bagging from initial value randomness. Additionally, it is found that the multiple classifier system voted output reduces the risk of relying on unexpected poor classifications.

References:

- [1] J. Janata, *Principles of Chemical Sensors*, pp.156-159, 2nd ed.: Springer, 2009.
- [2] P. Bergveld, "Thirty years of ISFETOLOGY: What happened in the past 30 years and what may happen in the next 30 years," *Sensors and Actuators B: Chemical*, vol. 88, pp. 1-20, 2003.
- [3] P. Bergveld, "ISFET, Theory and Practice," in *IEEE Sensor Conference*, Toronto, 2003, pp. 9-10.
- [4] W.-Y. Chung, C.-H. Yang, Y.-F. Wang, Y.-J. Chan, W. Torbicz, and D. G. Pijanowska, "A signal processing ASIC for ISFET-based chemical sensors," *Microelectronics Journal*, vol. 35, pp. 667-675, 2004.
- [5] P. Bergveld, "The impact of MOSFET-based sensors," *Sensors and Actuators*, vol. 8, pp. 109-127, 1985.
- [6] A. vandenBerg, P. Bergveld, D. N. Reinhoudt, M. Elwenspoek, and J. H. J. Fluitman, "Miniaturized chemical analysis systems," in *Micro Machine and Human Science, 1994. Proceedings., 1994 5th International Symposium on*, 1994, p. 181.
- [7] J. Goldman, N. Ramanathan, R. Ambrose, D. A. Caron, D. Estrin, J. C. Fisher, R. Gilbert, M. H. Hansen, T. C. Harmon, J. Jay, W. J. Kaiser, G. S. Sukhatme, and Y.-C. Tai, "White Paper: Distributed Sensing Systems for Water Quality Assessment and Management," Center for Embedded Networked Sensing (CENS) at UCLA, February 2007 2007.

- [8] O. Leistiko, "The Selectivity and Temperature Characteristics of Ion Sensitive Field Effect Transistors," *Physica Scripta*, vol. 18, pp. 445-450, 1978.
- [9] F. Deyhimi, "A method for the determination of potentiometric selectivity coefficients of ion selective electrodes in the presence of several interfering ions," *Talanta*, vol. 50, pp. 1129-1134, 1999.
- [10] A. Bratov, N. Abramova, and C. Domínguez, "Lowering the detection limit of calcium selective ISFETs with polymeric membranes," *Talanta*, vol. 62, pp. 91-96, 2004.
- [11] Y. Umezawa, K. Umezawa, and H. Sato, "Selectivity coefficients for ion-selective electrodes: Recommended methods for reporting $K_{A,B}^{pot}$ values," *Pure Appl. Chem.*, vol. 67, No. 3, , pp. 507-518, 1995.
- [12] W. Jatmiko, A. Nugraha, R. Effendi, W. Pambuko, R. Mardian, K. Sekiyama, and T. Fukuda, "Localizing multiple odor sources in a dynamic environment based on modified niche particle swarm optimization with flow of wind," *WTOS*, vol. 8, pp. 1187-1196, 2009.
- [13] S. Bermejo and J. Sole-Casals, "Blind source separation for solid-state chemical sensor arrays," in *Proceedings of Sensor Array and Multichannel Signal Processing Workshop , 2004*, 2004, pp. 437-440.
- [14] M. Janicki, M. Daniel, and A. Napieralski, "Application of Inverse Problem Algorithm for Estimation of Ion Mixture Composition," in *MIXDES 2004, 11th International Conference*, , Szczecin, Poland, 2004.
- [15] G. Bedoya, C. Jutten, S. Bermejo, and J. Cabestany, "Improving semiconductor-based chemical sensor arrays using advanced algorithms for blind source separation," in *Sensors for Industry Conference, 2004. Proceedings the ISA/IEEE*, 2004, pp. 149-154.
- [16] H. C. de Sousa, A. C. P. L. F. Carvalho, A. Riul, Jr., and L. H. C. Mattoso, "Using MLP networks to classify red wines and water readings of an electronic tongue," in *Neural Networks, 2002. SBRN 2002. Proceedings. VII Brazilian Symposium on*, 2002, pp. 13-18.
- [17] R. F. Pierret, *Semiconductor Device Fundamentals*, pp.525-550, 2nd edition ed.: Addison Wesley, 1996.
- [18] L. Ingemar, B. Albert van den, H. v. d. S. Bartholomeus, H. v. d. V. Hendrik, A. Mårten, and I. N. Claes, "Field Effect Chemical Sensors," in *Sensors*, P. J. H. D. J. N. Z. Prof. Dr. W. Göpel, Ed., 2008, pp. 467-528.
- [19] R. P. Buck and E. Lindner, "Recommendations for nomenclature of ion-sensitive electrodes (IUPAC Recommendations 1994)," *Pure Appl. Chem.*, vol. 66, pp. 2527-2536, 1994.
- [20] C.-M. Y. Daniel Tomaszewski, Bohdan Jaroszewicz, Michał Zaborowski, Piotr Grabiec, and Dorota G. Pijanowska "Electrical Characterization of ISFETs," *Journal of Telecommunications and Information Technology (JTIT)* pp. 55-60, 2007.
- [21] M. Gotoh, S. Oda, I. Shimizu, A. Seki, E. Tamiya, and I. Karube, "Construction of amorphous silicon ISFET," *Sensors and Actuators*, vol. 16, pp. 55-65, 1989.
- [22] W. F. H. O. Abdullah, M.; Ali, M.A.M., "Chemical field-effect transistor with constant-voltage constant-current drain-source readout circuit," in *2009 IEEE Student Conference on Research and Development (SCORED)*, UPM Serdang 2009, pp. 219 - 221
- [23] Y. Umezawa, P. Buhlmann, K. Umezawa, K. Tohda, and S. Amemiya, "Potentiometric Selectivity Coefficients of Ion-Selective Electrodes, Part I: Inorganic Cations (IUPAC Technical Report)," *Pure Appl. Chemosphere*, vol. 72, pp. 1851-2082, 2000.
- [24] D. A. Skoog, F. J. Holler, and D. M. West, *Fundamentals of analytical chemistry*, pp.100-118, 8th ed. New York: Brooks Cole, 2003.
- [25] A. Wan Fazlida Hanim, "Chemical Field Effect Transistor Response with Post Processing Supervised Neural Network," in *International Conference of Soft Computing and Pattern Recognition (SOCPAR)*, Melaka, 2009, pp. 250-253.
- [26] M. T. Hagan, H. B. Demuth, and M. H. Beale, *Neural network design*, 1st ed. Boston: PWS Pub. Co., 1996.
- [27] M. Riedmiller and H. Braun, "A direct adaptive method for faster backpropagation learning: the RPROP algorithm," in *IEEE International Conference on Neural Networks*, 1993, pp. 586-591 vol.1.
- [28] M. T. Hagan and M. B. Menhaj, "Training feedforward networks with the Marquardt algorithm," *IEEE Transactions on Neural Networks*, vol. 5, pp. 989-993, 1994.
- [29] H. B. Demuth, M. H. Beale, and MathWorks Inc., *Neural network toolbox for use with MATLAB : user's guide, version 5*. Natick, Mass.: MathWorks, Inc., 2006.
- [30] L. Breiman, "Bagging Predictors," in *Machine Learning*, 1996, pp. 123-140.
- [31] R. Polikar, "Bootstrap - Inspired Techniques in Computation Intelligence," *IEEE Signal*

- Processing Magazine*, , vol. 24, pp. 59-72, 2007.
- [32] R. Polikar, D. Parikh, and S. Mandayam, "Multiple classifier systems for multisensor data fusion," 2006, pp. 180-184.
- [33] N. Arora, "Mosfet modeling for VLSI simulation : theory and practice," in *International series on advances in solid state electronics and technology* New Jersey: World Scientific, 2007, pp. 402-494.
- [34] M. O. Wan Fazlida Hanim Abdullah, Mohd Alaudin Mohd Ali and Md Shabiul Islam, "Knowledge Representation of Ion-Sensitive Field-Effect Transistor Voltage Response for Potassium Ion Concentration Detection in Mixed Potassium/Ammonium Ion Solutions," *American Journal of Applied Sciences* vol. 7, pp. 81-88, 2010.