# Fuser design for thick film pH sensor electrodes using empirical data

S.J. Wellington[*a], J.K. Atkinson[**b] and R.P. Sion[***c]
[a]School of Computing and Digital Communication, Southampton Institute; [b]Dept. of Engineering Sciences, University of Southampton; [c]C-Cubed Limited, St Mary Bourne, Andover.

## ABSTRACT

An array of thick film pH sensor electrodes has been fused using two separate fuser designs: the feedforward neural network and Nadaraya-Watson kernel estimator. In both cases the fuser is based on empirical data rather than analytical sensor models. Complementary sensor responses have been obtained by fabricating sensors using different metal oxides. This approach provides some immunity to interference caused by the ionic composition of the solution being sensed. The Nadaraya-Watson estimator is shown to provide a useful alternative to the feedforward neural network for multisensor fusion where sensor distributions are unknown. Indicative test results are provided for the measurement of pH in printing ink. The results confirm that the fused results are more accurate than those obtained using the single best sensor, or simple fusion schemes such as averaging or majority voting.

**Keywords:** pH electrode, thick film sensor, multisensor fusion, Nadaraya-Watson, artificial neural network

## 1. INTRODUCTION

The measurement of pH is very important in a wide range of application areas, however the conventional glass pH electrode is notoriously sensitive to process conditions and maintenance practices. Consequently a variety of thick film pH electrodes have been proposed, with promising results obtained using sensors based on the oxides of various electrically semiconducting metals. Materials such as ruthenium, iridium and palladium have been utilised with encouraging results [1]. An ideal pH electrode will sense only the activity of hydrogen ions, however the metal oxide pH sensor has worse selectivity than the conventional glass electrode [2]. Thick film pH sensors generally exhibit higher redox sensitivity than the glass electrode and are also susceptible to interference from particular substances [3]. The response of the thick film device will therefore be influenced by the ionic composition of the solution being sensed and this can present problems, particularly in situations where the ionic composition of the determinand is dynamic, for example sensing the pH level of river water. Here sensor fusion can be usefully employed to reduce measurement uncertainty and increase reliability by combining readings from multiple pH sensors having complementary responses. The sensor array can be conveniently fabricated by printing sensors using different metal oxides for the active sensing element onto a single substrate.

Many approaches to sensor fusion require closed-form analytical expressions for sensor distributions [4]. These sensor distributions can be arbitrarily complex and consequently Gaussian distributions are often assumed, generally with a detrimental effect on overall system performance. In the case of the thick film pH sensor array it is relatively easy to collect data by experimentation, hence it is possible to exploit empirical data in the design of a suitable fusion scheme. Rao demonstrates that a feedforward neural network may be employed to realise a fuser with performance guarantees based on a finite sample. The design of the empirical 'best' neural network is still difficult, however good results have been obtained using the back propagation algorithm to provide an estimate of the optimal fuser [5]. The classical Nadaraya-Watson (N-W) estimator is also shown to solve the generic sensor fusion problem where the underlying sensor distributions are not known, but a sample is available. The use of Haar kernels has been shown to yield finite

[*]sean.wellington@solent.ac.uk; phone +44 (0)23 8031 9826; fax (0)23 8033 4441; School of Computing and Digital Communications, Southampton Institute, East Park Terrace, Southampton, SO14 0RD, United Kingdom; [**]jka@soton.ac.uk; Thick Film Unit, School of Engineering Sciences, University of Southampton, Highfield, Southampton, SO17 1BJ, United Kingdom; [***]russ@c-cubed.co.uk; C-Cubed Ltd., Pioneer House, St. Mary Bourne, Andover, SP11 6BL, United Kingdom.

sample guarantees, and also to be efficiently computable [6]. This paper describes the design and implementation of two separate fusers [7]:

- Data in – Data out: Individual sensor outputs are combined to estimate the actual value of pH.
- Data in – Decision out: Individual sensor outputs are combined to indicate whether some pre-defined condition exists; for example whether the value of pH is above or below some threshold level.

In each case a parsimonious solution is required that provides results better than the single 'best' sensor alone.

## 2. BACKGROUND

### 2.1 Feedforward Neural Network

Artificial neural networks are useful for function approximation because they are efficient universal approximators. A feedforward neural network with at least one hidden layer is also capable of approximating any non-linear function once enough training data are provided [8]. Various network architectures and learning algorithms have been employed for sensor fusion, for example Broten and Wood demonstrate the use of a feedforward neural network with sigmoidal activation functions to fuse an array of partially selective sensors [9]. The neural network is shown to generalise adequately with both complex sensor responses and noisy sensor data. A multilayer perceptron (MLP) was therefore selected as the candidate solution due to the simplicity of the architecture. The number of neurons in the input and hidden layers was made equal to the number of inputs (sensors). Sigmoidal activation functions were used throughout.

### 2.2 Nadaraya-Watson Kernel Estimator

The N-W estimator is a non-parametric kernel estimator (smoother) applicable to univariate and multivariate problems [10]. If n data points $\{X_i, Y_i\}_{i=1}^n$ have been collected, the regression relationship can be modelled as:

$$X_i = m(Y_i) + \varepsilon_i \tag{1}$$

where m in the unknown regression function and $\varepsilon_i$ are observation errors[1]. In most cases X denotes the (scalar) response variable and Y is a vector. Estimation of the function at a particular point y is performed using the mean of observations $X_i$ that correspond to $Y_i$ in the region of y. A kernel function is employed to weight the observations and the N-W estimator is given by:

$$\hat{f}_{h,n}(y) = \frac{\sum_{i=1}^{n} K_h(y - Y_i) X_i}{\sum_{i=1}^{n} K_h(y - Y_i)} \tag{2}$$

where K is the kernel function and h is the bandwidth parameter. The value of h controls the weighting of the observations with respect to their distance to y. In a system with N sensors the Haar kernel function is an N-dimensional hypercube, J centred on y, and having volume $h^N$. The N-W estimator based on Haar kernels is therefore:

$$\hat{f}_{h,n}(y) = \frac{\sum_{Y_i \in J} X_i}{\sum 1_J(Y_i)} \tag{3}$$

where $1_J(Y_i)$ denotes the indicator function of $1_J(Y_i) = 1$ if $Y_i \in J$, and $1_J(Y_i) = 0$ otherwise. Computation of $\hat{f}_{h,n}(y)$ at any given y involves obtaining the local sum of $X_i$'s in J. If no $Y_i$'s lie in J then $\hat{f}_{h,n}(y)$ is taken to be zero. In order to reduce the complexity of the estimator it is desirable to minimise the number of *a-priori* observations employed. Using

[1] Some authors (i.e. [10]) define the regression relationship as $Y_i = m(X_i) + \varepsilon_i$ however the definition given in Eqn. 1 (i.e. [5,6]) has been preferred for this paper.

few observations suggests that a large value of h will be needed to ensure at least one $Y_i$ lies in J for all y and this will increase the approximation bias.

## 3. EXPERIMENTAL METHOD

The training and test data were obtained by experimentation. The experimental procedure was intended to simulate an ink bath on a typical offset printer. The ink was pumped from a tank up into a tray approx 300mm by 200mm to a depth of approx 50mm. In order to simulate ink usage this was then drained into a further tray approx 0.5m below the first. This then in turn drained back into the ink tank.

The sensors were placed in the second tray as the ink in the first tray was prone to foaming, which made taking readings difficult. The sensors were mounted on a floating platform, which could hold approx 20 sensors at any one time. All readings were measured against a standard BDH reference electrode. The pH level was modulated by adding either ammonia to increase it, or acetic acid to reduce it. The additions were made in the ink tank as the pump also had the effect of mixing the solutions.

Temperature was kept reasonably constant by placing the ink tank in a temperature controlled water bath. All sensors were logged using a Keithley 2700 data logger with a sampling frequency of $\frac{1}{60}$ Hz. A calibrated glass pH sensor electrode was also used to monitor pH level in the second tray, again relative to the commercial reference electrode. Results are shown in Fig. 2 for the control pH level and five thick film sensors fabricated using the following metal oxides:

Sensor 1 Iridium oxide
Sensor 2 Iridium oxide
Sensor 3 Ruthenium oxide
Sensor 4 Titanium oxide
Sensor 5 Ruthenium oxide

The network architecture employed for the fusion of five sensors is illustrated in Fig. 1. The neural network was trained using the back-propagation learning algorithm. For the purpose of direct comparison, the N-W estimator used the same training and test data. In the case of the N-W estimator, fuser design involves selecting the *a-priori* observations $\{X_i, Y_i\}_{i=1}^{n}$ and then choosing a value for the bandwidth parameter, h. With a fixed sample the size of bandwidth determines the degree of smoothing and is therefore of crucial importance for the properties of the final estimate. Data-driven cross-validation techniques are sometimes used for automatic bandwidth selection, however the following empirical design procedure has been adopted:

- A set of observations was selected that covered the full operating range of the measurement system;
- The estimator was implemented and a suitable value for h chosen such that at least one $Y_i$ lies in J for all y;
- The process was repeated until acceptable performance achieved.

## 4. EXPERIMENTAL RESULTS

### 4.1 Data in – Data out Fusion

The mean squared error value for each sensor may be calculated (Table 1). It can be seen from this data, and by inspection of Fig. 2 that there is significant variance in the performance of individual sensors.

| Sensor | 1 | 2 | 3 | 4 | 5 |
|---|---|---|---|---|---|
| MSE×$10^6$ | 3.4562 | 3.7751 | 6.8783 | 5.6976 | 8.1311 |

Table 1: Mean square error of thick film pH sensors

Initially all five sensors were fused using both the feedforward neural network (Fig. 1) and the N-W kernel estimator. The results are shown in Fig. 3 and the mean squared error for each fuser is given in Table 2. Increasing the number of observations reduced the mean squared error, however in all cases the fused result is more accurate than the 'best sensor'. The performance of the neural network is slightly better than the N-W fuser in this configuration.

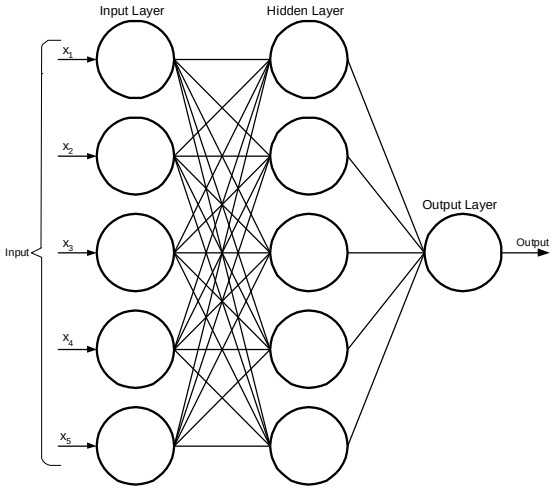


Figure 1: Architecture of the feedforward neural network used to fuse five pH sensors

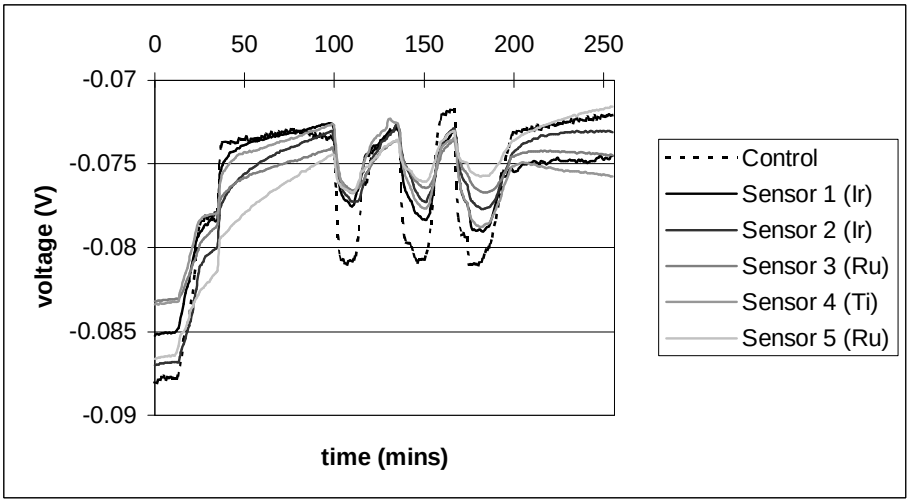


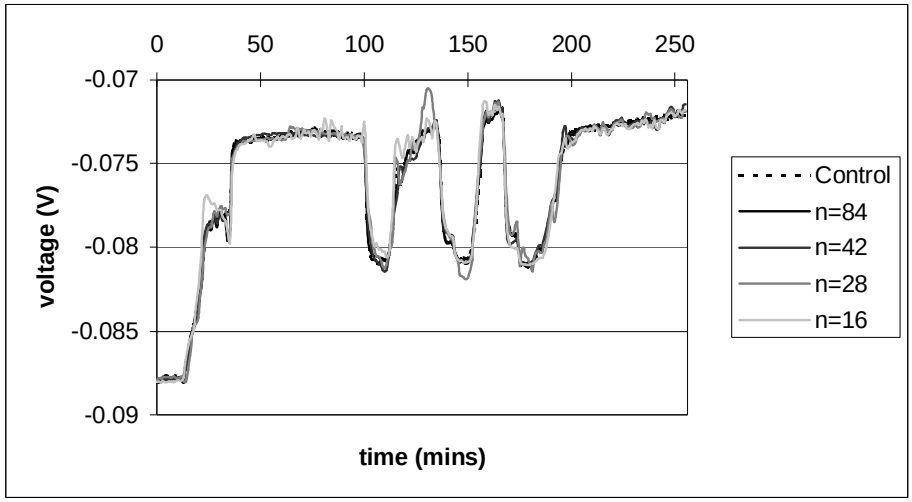
Figure 2: Test results obtained from five thick film pH sensors

| Number of observations | MSE×$10^6$ | |
|---|---|---|
| | Neural Network | Nadaraya-Watson |
| 16 | 0.4654 | 1.2724 |
| 28 | 0.3336 | 0.7826 |
| 42 | 0.1501 | 0.5912 |
| 84 | 0.0643 | 0.3674 |

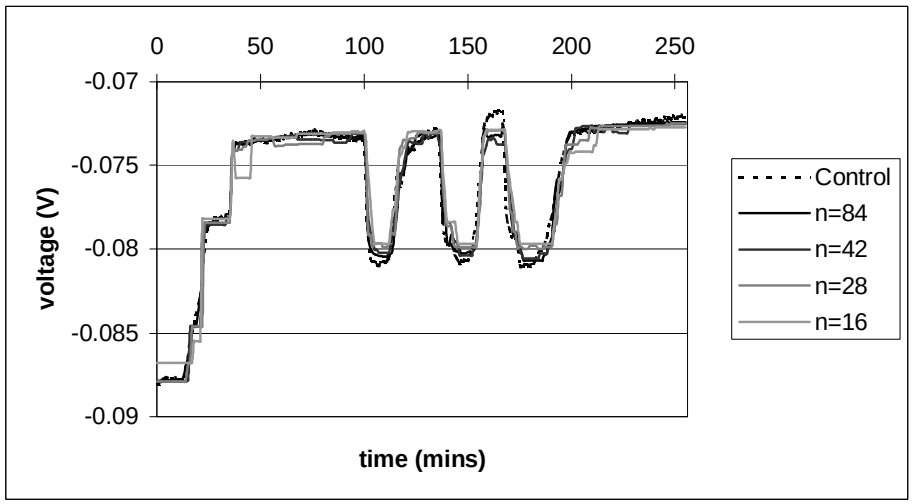
Table 2: Effect of altering the number of observations used to realise the fuser

| Bandwidth, h | MSE×$10^6$ |
|---|---|
| 0.0029 | 0.5912 |
| 0.0040 | 0.7493 |
| 0.0050 | 1.1228 |
| 0.0060 | 2.0655 |

Table 3: Effect of altering the bandwidth of the kernel estimator




(a) Artificial neural network




(b) Nadaraya-Watson estimator

Figure 3: Effect of altering the number of observations used to realise the fuser

The effect of varying the bandwidth parameter for the N-W estimator is shown in Fig. 4 and Table 3 for a fuser based on 42 observations. Reducing the bandwidth led to a reduction in the mean squared error of the estimate, however for the particular test data employed, h=0.0029 was the minimum value to give at least one $Y_i$ in J for all y. This apparent correlation between bandwidth and mean squared error suggests that bias errors contribute significantly to the total measurement uncertainty, and this inference is confirmed when the measurement errors are evaluated using simple statistical techniques.

Tests were repeated to evaluate the results of fusing 4, 3 and 2 sensors. Indicative results are summarised in Table 4. The network topology employed in case used N neurons in each of the input and hidden layers. The results indicate that both fusion techniques produce a fused result that is more accurate the best single sensor. Simply averaging the sensor readings did not generally improve the accuracy of the estimate due to the bias errors of individual sensors, while increasing the number of sensors tended to improve the accuracy of the fused result.

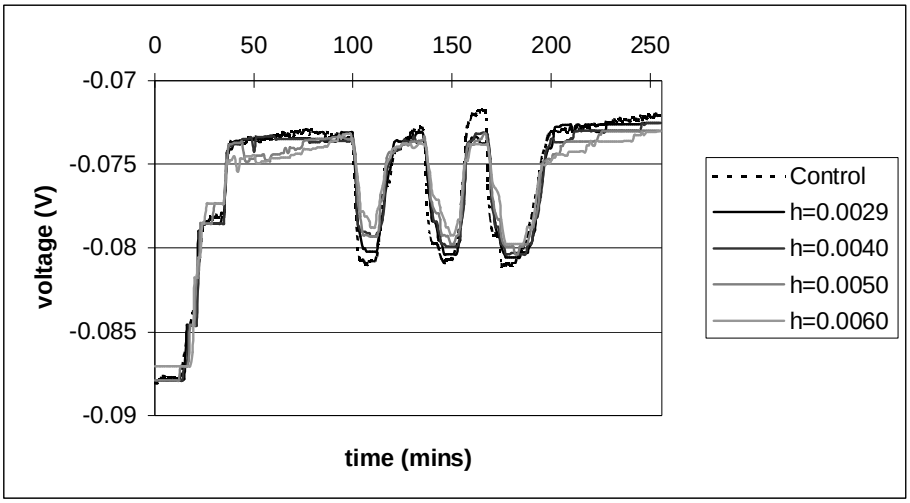



Figure 4: Effect of altering the bandwidth of the kernel estimator

| MSE×$10^6$ | 5 Sensors | 4 Sensors | 3 Sensors | | 2 Sensors | | |
|---|---|---|---|---|---|---|---|
| | 1,2,3,4,5 | 1,2,3,4 | 1,2,3 | 2,3,4 | 1,2 | 2,3 | 3,4 |
| Best Sensor | 3.4562 | 3.4562 | 3.4562 | 3.7751 | 3.4562 | 3.7751 | 5.6976 |
| Average | 4.4865 | 4.3199 | 4.1076 | 4.7826 | 3.1505 | 4.8925 | 5.9539 |
| Neural Network | 0.1501 | 0.3421 | 0.4875 | 0.7441 | 0.9493 | 1.0391 | 1.7278 |
| Nadaraya-Watson | 0.5912 | 1.1532 | 1.0614 | 1.8251 | 1.0661 | 2.8628 | 2.3417 |

Table 4: Summary of sensor fusion results

**4.2 Data in – Decision out Fusion**

The training data are configured such that:
y = +1 if hypothesis Ho is True
y = -1 if hypothesis Ho is False

A multilayer perceptron was again found to give good results, with sigmoidal activation functions employed throughout. It is, of course, necessary to ensure that the training data includes sufficient examples of input-output pairs for the case Ho is True and also Ho is False.

Consider the hypothesis Ho, where Ho is True if the value of pH sensor output is greater than –0.075V using the sensor data shown in Fig. 2. The neural network was trained using 42 input-output pairs with the back-propagation learning algorithm and tested using 256 input-output pairs. The N-W estimator used the same training and test data and the results have been compared with those obtained using a simple majority voting scheme [11]. The performance of the individual sensors is summarised in Table 5, while the results obtained from the fusion schemes are given in Table 6.

Systematic errors contribute significantly to the relatively poor performance of the majority voting scheme, even when five separate sensors are employed. It is, however, apparent that both the neural network and N-W estimator are able to provide the non-linear mapping necessary to compensate for the deficiencies of individual sensors.

| Sensor | 1 | 2 | 3 | 4 | 5 |
|---|---|---|---|---|---|
| %Misclassification | 9.8% | 12.1% | 18.8% | 26.9% | 26.5% |

Table 5: Number of misclassifications for the hypothesis Ho is True if the pH sensor output is greater than –0.075V

| %Misclassification | 5 Sensors | 3 Sensors | |
|---|---|---|---|
| | 1,2,3,4,5 | 1,2,3 | 2,3,4 |
| Best Sensor | 9.8% | 9.8% | 12.1% |
| Majority Vote | 12.5% | 12.9% | 12.9% |
| Neural Network | 2.7% | 4.7% | 5.1% |
| Nadaraya-Watson | 3.1% | 5.5% | 7.4% |

Table 6: Comparison of decision fusion results obtained using majority voting and a feedforward neural network

## 5. DISCUSSION

The measurement of pH is very important in a wide range of application areas, however the conventional glass pH electrode has the highest sensitivity to process conditions and engineering and maintenance practices of the common industrial measurements. Consequently a variety of thick film pH electrodes have been proposed, with promising results obtained using sensors based on the oxides of various electrically semiconducting metals.

Uncertainty reduction by multisensor fusion is best achieved using sensor complementarity [12]. In this application complementary pH sensor responses can be conveniently obtained by screen-printing thick film sensor electrodes using different metal oxides onto a single substrate. Sensors based on the oxides of iridium, ruthenium and titanium have been fused using a feedforward neural network and the N-W kernel estimator. Fog and Buck have investigated the characteristics of these metal oxides as pH-sensitive electrodes [3]. The complementary nature of the sensors is evident from Table 7. All three devices exhibit near-Nernstian characteristics over the pH range 2 – 10, however the sensors also have some sensitivity to oxidixing and reducing agents, with two sensors affected by iodide and individual devices also affected by the ions of fluoride, chloride and bromide.

| Metal Oxide | Sensitivity MV/pH | Recommended pH range | Accuracy ±mV | Hysteresis pH 2→12 →2 ±mV | Redox interference ±mV | Other interferences |
|---|---|---|---|---|---|---|
| Titanium | 55.0 | 2 – 12 | 15 | 30 | 100 | $F^-$ |
| Ruthenium | 61.8 | 2 – 12 | 2 | 9 | 100 | $I^-$ |
| Iridium | 59.8 | 2 – 10 | 2 | 25 | 20 | $Cl^-, Br^-, I^-$ |

Table 7: Characteristics of metal oxides as pH-sensitive electrodes at 25°C (Source: [3])

The potential benefits of fusing multiple thick film pH sensors include:

- Thick film technology provides a robust, low-cost, low-maintenance alternative to the conventional glass pH sensor electrode;
- The system should continue to operate despite the failure of one or more sensors, although the performance may be degraded;
- Complementary sensor responses are obtained by fabricating sensors using different materials. The combined sensor array will therefore be less susceptible to errors caused by the ionic composition of the solution being sensed.

Feedforward neural networks can be used to provide data in – data out and data in – decision out fusion for thick film pH sensor arrays. The design of the required fuser is an iterative process that seeks to balance overall performance against computational cost and complexity. For an array with N sensors, a multilayer perceptron with N neurons in each of the input and hidden layers, and 1 neuron in the output layer was found to give consistently good results.

The N-W kernel estimator is a well-known non-parametric estimation technique. It has, however, found only limited use as a fuser. The use of Haar kernel functions yields an implementation that is efficiently computable, however the artificial neural network also provides a convenient and well-established method of approximating some non-linear function based on training data. An empirical design methodology is adopted for both the N-W and artificial neural network fuser; the latter is facilitated by the comprehensive design tools and literature that is now available. The N-W fuser provides good results and benefits from an implementation that is not reliant on topology or learning algorithm.

The fuser is a form of rule-based system, with rules based on the *a-priori* observations, and the rules that 'fire' determined by the measurement vector. There is an inherent trade-off involved in the choice of bandwidth: the effect of random errors will be reduced by increasing the bandwidth and hence including more observations in the computation of the estimate, while the effects of bias errors are reduced by reducing the bandwidth and hence including fewer observations. It is hypothesised that some automated optimisation technique, for example a genetic algorithm, may improve the performance of the fuser by evolving an optimal set of observations and value of the bandwidth parameter.

The use of the Haar kernel, while simplifying the computation of the estimate, imposes one significant constraint: the kernel bandwidth is equal for all N sensors. This implies that the sensors are commensurate and have similar calibration constants. The method is therefore most suited to the fusion of an array of similar sensors, although an alternative kernel function may overcome this limitation. A potential benefit of the N-W fuser is that it can be easily reconfigured in the event of a sensor fault. The artificial neural network cannot provide this degree of flexibility because reconfiguration of the network to accept one less sensor input would generally need to be accompanied by retraining. The N-W fusion algorithm is relatively straightforward and can be implemented easily using a routine commercial general-purpose digital signal processor or even a microcontroller for applications employing a low sampling frequency.

## 6. CONCLUSIONS

The N-W kernel estimator provides a useful alternative to the feedforward neural network for multisensor fusion where sensor distributions are unknown. In each case the fuser is implemented using data obtained by experimentation. These data are used explicitly in the case of the N-W estimator, and to adapt the weights of the neural network. In general, optimisation of the N-W fuser design is more difficult than the neural network, although the N-W estimator with Haar kernels can be easily reconfigured to accept one less input in the event of a sensor fault.

## 7. ACKNOWLEDGEMENTS

Experimental data were obtained using thick film pH sensors fabricated by the Thick Film Unit at the University of Southampton as part of the EU CRAFT project SpHINX – Sensors for measuring pH in inks (BRST-CT-98-5521).

## REFERENCES

1. J.A. Mihell and J.K. Atkinson, "Planar thick-film pH electrodes based on ruthenium dioxide hydrate," *Sensors and Actuators (B),* **48**, pp. 505-511, 1998.
2. J.K. Atkinson, A.W.J. Cranny and P.R. Siuda, "Organic materials for gas sensing and the construction of sensors for liquids," In: M. Prudenziati, ed. *Handbook of Sensors and Actuators: Vol. 1: Thick Film Sensors.* Elsevier Science, pp. 313-340, 1994.
3. A. Fog and R.P Buck, "Electronic Semiconducting Oxides as pH Sensors," *Sensors and Actuators,* **5**(2), pp. 137-146, 1984.
4. D.L. Hall, *Mathematical Techniques in Multisensor Data Fusion.* Artech House, 1992.
5. N.S.V. Rao, "Multiple sensor fusions under unknown distributions," *Journal of the Franklin Institute,* **336**, pp. 285-299, 1999.
6. N.S.V. Rao, "Nadaraya-Watson estimator for sensor fusion," *Optical Engineering,* **36**(3), pp. 642-647, 1997.
7. B.D. Dasarathy, "Sensor Fusion Potential Exploitation – Innovative Architectures and Illustrative Application," *Proc. IEEE,* **85**(1), pp. 6-23, 1997.
8. G. Cybenko, "Approximation by superposition of sigmoidal functions," *Mathematics of Control Signals and Systems,* **2**(4), pp. 303-314, 1989.
9. G.S. Broten and H.C. Wood, "A neural network approach to analysing multi-component mixtures," *Meas. Sci. Technol.*, **4**, pp. 1096-1105, 1993.
10. W. Hardle, *Applied Nonparametric Regression,* Cambridge University Press, 1990.
11. P.K. Varshney, "Multisensor Data Fusion," *IEE Computing and Control Engineering Journal,* December 1997, pp. 245-253, 1997.
12. P. Grossmann, "Multisensor Data Fusion," *GEC Journal of Technology,* **15**(1), pp. 27-37, 1998.