ELSEVIER

Sensors and Actuators B 69 (2000) 336–341

SENSORS AND ACTUATORS B CHEMICAL

www.elsevier.nl/locate/sensorb

An electronic nose system for monitoring the quality of potable water

Julian W. Gardner [a,*], Hyun Woo Shin [a], Evor L. Hines [a], Crawford S. Dow [b]

[a] *School of Engineering, University of Warwick, Gibbet Hill Road, Coventry CV4 7AL, UK*
[b] *Department of Biological Sciences, University of Warwick, Gibbet Hill Road, Coventry CV4 7AL, UK*

Received 20 September 1999; accepted 22 February 2000

Abstract

A measurement system has been developed for the testing of cyanobacteria in water, and it consists of three main stages: the odour sampling system, an electronic nose (e-nose) and a CellFacts instrument that analyses liquid samples. The e-nose system, which employs an array of six commercial odour sensors, has been used to monitor not only different strains but also the growth phase of cyanobacteria (i.e. blue-green algae) in water over a 40-day period. Principal components analysis (PCA), multi-layer perceptron (MLP), learning vector quantisation (LVQ) and Fuzzy ARTMAP were used to analyse the response of the sensors. The optimal MLP network was found to classify correctly 97.1% of the unknown nontoxic and 100% of the unknown toxic cyanobacteria. The optimal LVQ and Fuzzy ARTMAP algorithms were able to classify 100% of both strains of cyanobacteria samples. The accuracy of MLP, LVQ and Fuzzy ARTMAP in terms of predicting four different growth phases of toxic cyanobacteria was 92.3%, 95.1% and 92.3%, respectively. These results show the potential application of neural network based e-noses to test the quality of potable water as an alternative to instruments, such as liquid chromatography or optical microscopy.

Keywords: Cyanobacteria; Gas sensor array; Electronic nose system; Neural network; Fuzzy ARTMAP

1. Introduction

The diversity of cyanobacteria species is becoming both a severe problem in the quality of potable water and a common source of odour pollution in freshwater reservoirs and local environmental water [1,2]. It has been associated with sewage effluent, industrial effluent, waste products from agriculture, and animal wastes from intensive farming. Cyanobacteria, which are usually called blue-green algae, grow in lakes and reservoirs and can be a serious nuisance due to their unpleasant odour and taste, in the case of reservoirs. The cyanobacteria is the largest group of photosynthetic prokaryotes and contain chlorophyll that differs from the bacteriochlorophylls of the photosynthetic eubacteria. The main problem associated with certain cyanobacteria is that they can produce toxins that are poisonous to cattle, wildfowl, fish and people. Many species of cyanobacteria have been observed to produce these toxins. The toxins can be divided into three groups: peptide hepatotoxins, neurotoxins and lipopolysaccharides. Thus, appropriate methods to detect and quantify these toxins in natural waters are very important. Here, we report on the use of an electronic nose (e-nose) [3], based on an array of six metal oxide semiconductor (MOS) sensors, to analyse cyanobacteria cultures grown in water, with the intent to provide a simple tool to test the quality of potable water as an alternative to analytical instruments that are based on liquid chromatography or optical microscopy. Earlier work on microbial detection has been reported on *Escherichia coli* grown in blood medium [4], biopharmaceutical process [5] and vagina infection [6], and shows that an e-nose has potential application within the fields of bioprocess monitoring and medicine.

2. Experimental

We have constructed a measurement system for the testing of the cyanobacteria over a period of 40 days. The system consisted of three main parts: the odour sampling unit, the Warwick-modified Fox 2000 unit (Alpha MOS, France) and a CellFacts I instrument (Microbial System, Coventry). The number of cells and cell size are sensitive indicators of the physiological status of algal cells: they

* Corresponding author. Tel.: +44-24-76523695; fax: +44-24-76418922.
E-mail address: j.w.gardner@warwick.ac.uk (J.W. Gardner).

0925-4005/00/$ - see front matter
PII: S0925-4005(00)00482-2

Table 1
Examples of signal processing algorithms used

Signal processing algorithms	Formula
Difference signal (signal from cyanobacteria a – signal from medium b)	$x_{ij} = x'_{a,ij} - x'_{b,ij}$
Relative signal	$x_{ij} = x'_{a,ij} / x'_{b,ij}$
Fractional difference	$x_{ij} = (x'_{a,ij} - x'_{b,ij}) / x'_{b,ij}$
Normalisation of sensor range	$k_{ij} = (x_{ij} - x_{ij,\min}) / (x_{ij,\max} - x_{ij,\min})$
Autoscaling ($\bar{x}$ is mean, and σ = population standard deviation)	$k_{ij} = (x_{ij} - \bar{x}_{ij}) / \sigma_i$

change as cells go through the different stages of growth. Cell size and numbers can be used to determine optimum culture conditions and growth phase. Small samples of liquid were extracted and analysed using a commercial CellFacts instrument, which measures the size and distribution of the bacteria. An e-nose based on an early Fox 2000 instrument (comprising an array of six commercial odour sensors) has been used to monitor not only different strains, but also the growth phase of the different strains of cyanobacteria. A series of experiments were carried out to analyse the nature of two closely related cyanobacteria, *Microcystis aeruginosa* PCC 7806, which produces a toxin, and PCC 7941, which does not. The sampling system was operated in a cyclic fashion, whereby a set sequence of timed valve actuations was repeated for a pre-determined number of times. It consisted of sampling the headspace (1.5 l) of three identical vessels (5 l), one of which contained just 3.5 l of growth medium (nutrient water) while the other two were inoculated with a small number of cells in 100 ml of inoculum. The growth medium (BG-11) was made using double distilled water and various analytical grade chemicals [8]. The reference vessel was sampled for most (80%) of the duty cycle and this set the output of the sensor array to the headspace of the pure growth medium and not ambient air. Each of the two inoculated vessels was selected in turn and the change in sensor responses observed and recorded by a PC running a custom virtual instrument under Labview 6.0 (National Instruments). See the next section for typical plots of the sensor signals recorded.

3. Results and discussions

Several data pre-processing techniques were investigated, having been selected following the results from earlier work [2] (see Table 1). These algorithms condition the input signals from the sensor array prior to pattern recognition that, like the mammalian olfactory system, performs data compression and noise rejection. Any systematic bias of the sensor signals can be reduced at this stage. Each sensor i produces a time-dependent signal, $x'_{ij}(t)$, in response to odour j, and it is often convenient to remove the time dependence of the signal output. Autoscaling and normalisation give equal weighting to each sensor and thus compensate for differences in the magnitudes of the signals. The resulting pre-processed time-independent parameters were then used for the principal components analysis (PCA), and for the analysis using the multi-layer perceptron neural network (MLP), learning vector quantisation (LVQ) and Fuzzy ARTMAP neural networks. All the neural network analysis was performed using the software package, NeuralWorks Professional II/Plus (NeuralWare, USA).

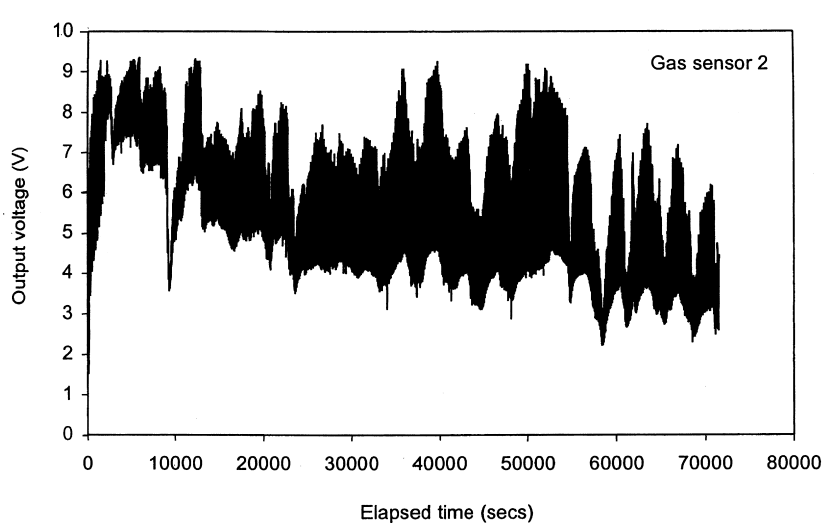

Fig. 1. Plot of the voltage signal produced by gas sensor 2 over the lifetime of the cyanobacteria. The broad line is due to the compression of hundreds of sampling sequences as seen in Fig. 2. The observed periodic variation in (a) baseline is attributed to diurnal variation in ambient air quality and (b) pulse height is attributed to ambient temperature producing diurnal variation in the cell metabolism on top of that associated with cell growth phase.

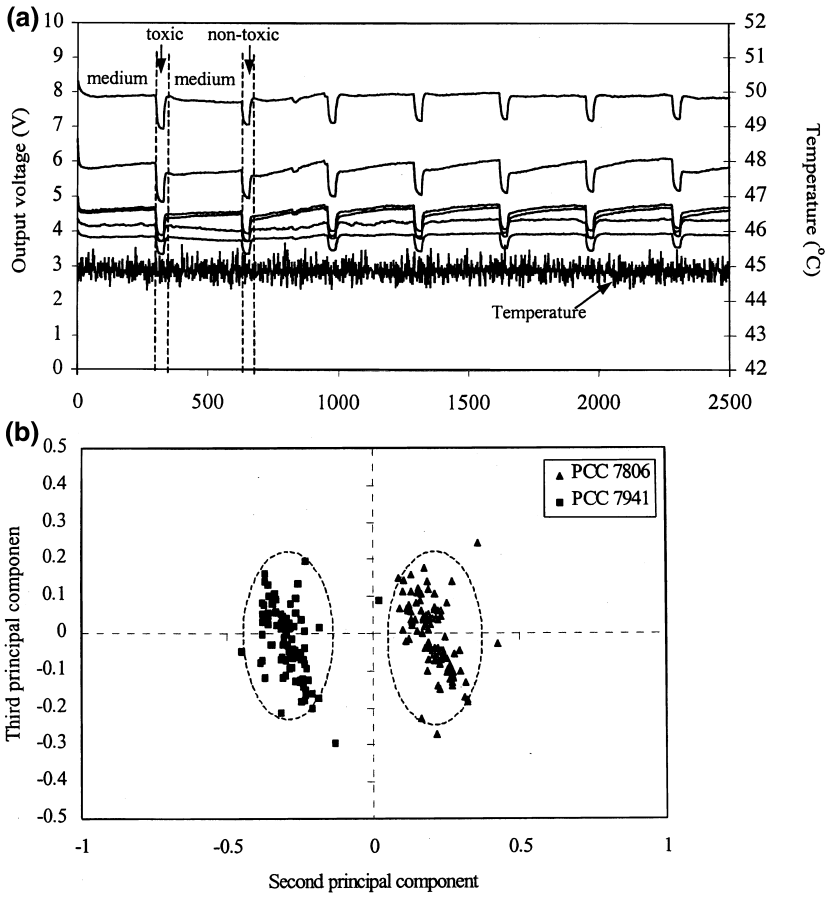

Fig. 2. (a) Plot of voltage signal produced by gas sensor 2 over three sampling sequences. The baseline signal appears stable over this short time period (42 min) while the toxic and nontoxic pulses heights are similar. (b) PCA results of second and third principal components on cyanobacteria samples, PCC 7806 (toxic) and PCC 7941 (nontoxic). The first component contains intensity information and helps in the discrimination of growth phase rather than strain. However, it still shows good separation between clusters [8]. The raw data are pre-processed using the normalised fractional difference model.

Fig. 1 shows the response of a single sensor over the entire experiment. There is considerable long-term variation in the baseline value of some 50%, and diurnal variation of some 15%. We attribute these phenomena to changes in the properties of the ambient air, such as pollutants, and perhaps some slight aging of the growth

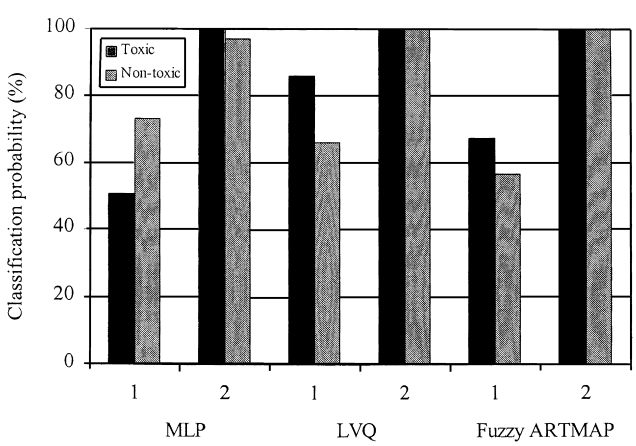

Fig. 3. A bar chart showing the MLP, LVQ and Fuzzy ARTMAP classification probability (%) of correctly classified toxic (black) and nontoxic (grey) bacteria for two representative pre-processing algorithms; (1) difference normalisation, (2) fractional difference normalisation.

medium. Considerable effort was made to make sure that we did not learn sensor drift. First, a control experiment was run in which all three vessels contained the growth medium and no discernible signal was observed while switching between vessels. Second, we repeated the experiment with the alternative vessels containing toxic and nontoxic inocula. Thirdly, we grew separate cultures of bacteria at different times to sample out of chronological order. Finally, we also looked at the mass spectra and found that it again correlated to headspace. Fig. 2(a) shows the actual signal from a single sensor over a short period of time with the fall in signal on introducing an inoculated vessel's headspace clearly visible. Fig. 2(b) shows a PCA plot of the entire data set using the normalised fractional difference algorithm (defined in Table 1). The PCA plot of the second and third principal components displays the two distinct strains of cyanobacteria; the first component (97% of variance) gives discrimination with elongated clusters [8], because it contains information about the odour intensity (concentration) and this correlates better with the growth phase (see Fig. 4). The best MLP set of parameters (based on back propagation using the delta learning rule) was found to classify correctly 97.1% of the unknown nontoxic bacteria samples and 100% of the unknown toxic cyanobacteria on the basis of a set of 378 training vectors and 202 test vectors from total a 580 of vectors. The best LVQ and Fuzzy ARTMAP were able to classify 100% of both types of cyanobacteria. The results for MLP, LVQ and Fuzzy ARTMAP are summarised in Fig. 3. Further details may be found in Ref. [7].

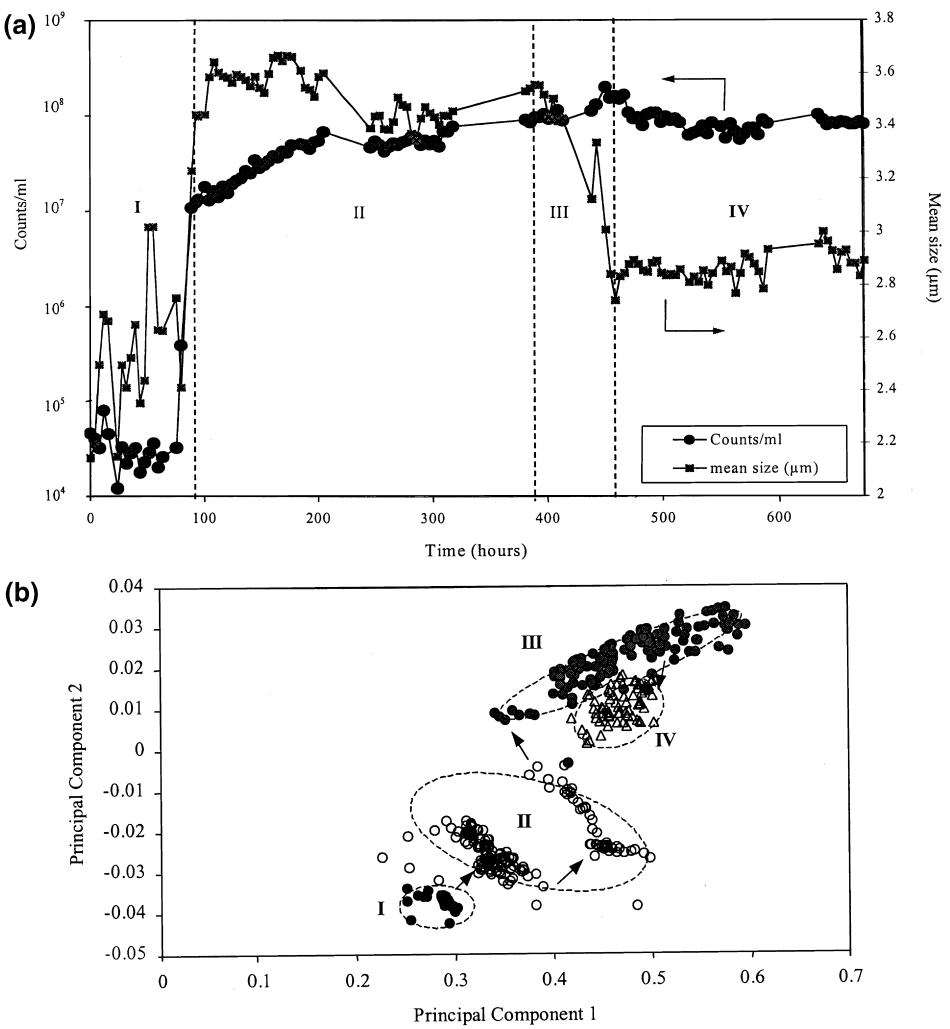

Fig. 4. (a) A growth phase plot, using a CellFacts instrument, showing the number of cell and cell size for cyanobacteria over a 700 h period. (b) PCA results of the response of a six-element gas sensor based electronic nose to the headspace of cyanobacteria. The four growth phases are lag, growth, stationary and late stationary (labelled I to IV).

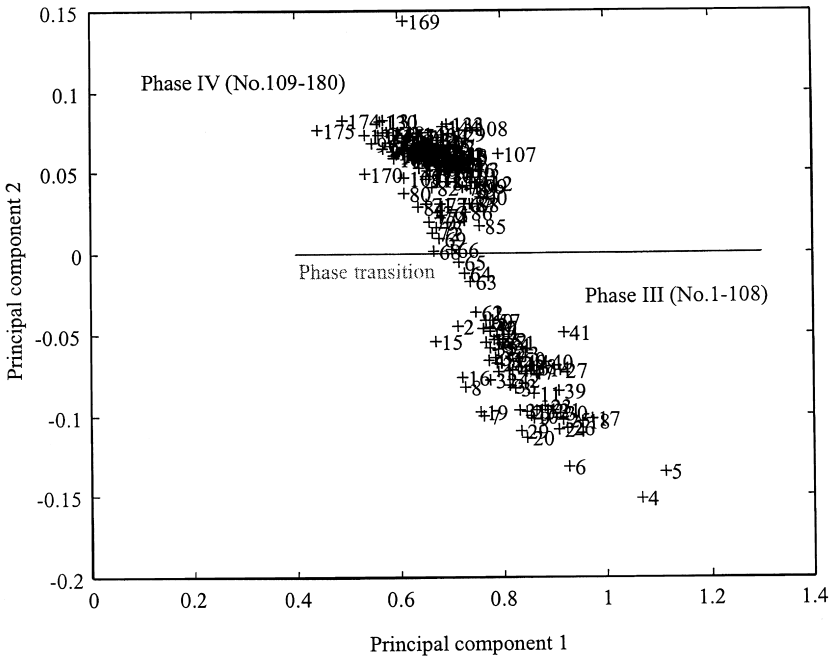

Fig. 5. PCA results of the response of a six-element gas sensor based electronic nose to the headspace of cyanobacteria. The two growth phases are stationary and late stationary (labelled III to IV). The response data were pre-processed using the normalised fractional difference model.

Fig. 4 shows (a) the results of a growth plot of Cell-Facts instrument and (b) a PCA plot for four different growth phases (lag, growth, stationary and late stationary) of cyanobacteria samples. There is some overlap of the response vectors, corresponding to the transition periods between phases. Fig. 5 shows the PCA plot for two growth phases, stationary and late stationary, from another series of measurements where point 1 and point 180 represent $t = 0$ and $t = 9000$ h, respectively. Each point is a representation of response from all of he sensors in the six-sensor array after pre-processing with the normalised fractional difference and a time interval of 50 min. The data points move from Phase III into Phase IV, and hence result in a transition of the response pattern.

The cyanobacteria data set was divided into three test folds containing 48 measurements each (12 measurements per phase category) and each neural network was trained using 144 vectors for each fold. Each network has six inputs and four outputs since a one-of-four code was used to code the four different phases (labelled I to IV). The classification rates of MLP, LVQ and Fuzzy ARTMAP were similar (92.3%, 95.1% and 92.4%, respectively), but Fuzzy ARTMAP was the fastest and judged to perform best because it self-organises and selects its own " hidden neurones".

4. Conclusions

A six-element metal oxide based e-nose system has been used for the continuous monitoring of the growth of cyanobacteria over a period of 40 days. Several pre-processing techniques were explored in order to remove the variation associated with growing the bacteria at ambient temperature and running the e-nose in ambient air. The normalised fractional difference method gave the best PCA result. There is some overlap of the response vectors in each PCA plot, corresponding to the transition periods of the cyanobacteria cultures.

Three supervised neural networks, MLP, LVQ and Fuzzy ARTMAP, were used and compared for the classification of both two strains and four different growth phases of cyanobacteria (lag, growth, stationary and late stationary). Our best results showed that the toxic strain of cyanobacteria grown in laboratory conditions was correctly predicted with an accuracy of 100% using MLP, LVQ and Fuzzy ARTMAP. The growth phase of the toxic cyanobacteria was correctly predicted for 95.1% of all unknown samples using LVQ. The LVQ was shown to perform a few percentage points better than MLP and Fuzzy ARTMAP, but the training iterations of Fuzzy ARTMAP was found to be typically more than an order of magnitude less than those for the MLP and the LVQ network. This work shows the potential application of an e-nose for monitoring the quality of potable water.

Acknowledgements

The authors thank Dr. Eduard Llobet, Mr. Frank Courtney, Dr. Uthaya Swoboda and Mrs. Jenny Flint for their help. Thanks are also due to Alpha MOS (France) for the supply of the odour sensors, and the British Embassy and LG Corporate Institute of Technology in Seoul for the UK–Korea Science and Technology Collaboration Fund.

References

[1] W.W. Carmichael, The toxins of cyanobacteria, Sci. Am. (1994) 64–72, January.
[2] W.W. Carmichael, Cyanobacteria secondary metabolites — the cyanotoxins, J. Appl. Bacteriol. 72 (1992) 445–459.
[3] J.W. Gardner, P.N. Bartlett, Electronic Noses: Principles and Applications, Oxford Univ. Press, 1999.
[4] J.W. Gardner, M. Craven, C. Dow, E.L. Hines, The prediction of bacteria type and culture growth phase by an electronic nose with a multi-layer perceptron network, Meas. Sci. Technol. 9 (1998) 120–127.
[5] C. Mandenius, A. Hagman, F. Dunas, H. Sundgren, I. Lundström, A multisensor array for visualizing continuous state transition in biopharmaceutical processes using principal component analysis, Biosens. Bioelectron. 13 (1998) 193–199.
[6] S. Chandiok, B.A. Crawley, B.A. Oppenheim, P.R. Chadwick, S. Higgins, K.C. Persaud, Screening for bacteria vaginosis: a novel application of artificial nose technology, J. Clin. Pathol. 50 (1997) 790–791.
[7] H.W. Shin, E. Llobet, J.W. Gardner, E.L. Hines, C.S. Dow, The classification of the strain and growth phase of cyanobacteria in potable water using an electronic nose system, IEE-Sci. Meas. Technol., submitted for publication.
[8] H.W. Shin, PhD Thesis, 1999, University of Warwick.

Biographies

Julian W. Gardner (BSc, PhD, DSc, CEng, FIEE, MIEEE) joined the School of Engineering at Warwick University in 1987 and is now Professor of Electronic Engineering. He has published over 200 technical papers and is an author of books on Nanotechnology and Instrumentation (1991), Electronic Noses (1992 and 1999) and Microsensors (1994). He currently heads the Sensors Research Laboratory in the Centre for Nanotechnology and Microengineering at Warwick University and Chairman of the IEE Professional Group J1 committee on Measurements and Instrumentation.

Hyun Woo Shin received his MSc in metallurgical engineering from Seoul National University in 1989. He joined LG Electronic Research Centre (LG-ERC) in 1989. He is currently a PhD student in the School of Engineering, University of Warwick.

Evor L. Hines (BSc, CNAA, PhD) joined the School of Engineering at Warwick University in 1984. His main research interest is concerned with Intelligent Systems and their applications. Most of his work has focused on Artificial Neural Networks, but also includes Genetic Algorithms, Fuzzy Logic, Neurofuzzy Systems. Typical application areas include intelligent sensors (e.g. electronic nose); nondestructive testing of, for example, composite materials; computer vision, among others. He has been involved in the publication of about 100 papers in these areas. He currently heads the Intelligent Systems Engineering Laboratory.

Crawford S. Dow (BSc, PhD) joined the Department of Biological Sciences at Warwick University in 1975. His main research interests lie in aspects of microbial biofilms, microbial adhesion and the prevalence and importance of quiescent vegetative cells. Cyanobacterial toxin research is closely allied to the water industry and involves the regulation of expression of peptide toxins, their toxicology, and environmental monitoring and management of cyanobacterial populations. He currently heads the Microbial Physiology Group (MPG).