Measurement 74 (2015) 78–86

Contents lists available at ScienceDirect

Measurement

journal homepage: www.elsevier.com/locate/measurement

Application of multiple linear regression, central composite design, and ANFIS models in dye concentration measurement and prediction using plastic optical fiber sensor

Su Sin Chong [a], A.R. Abdul Aziz [a,*], Sulaiman W. Harun [b], Hamzah Arof [b], Shahaboddin Shamshirband [c]

[a] Department of Chemical Engineering, Faculty of Engineering, University of Malaya, 50603 Kuala Lumpur, Malaysia
[b] Department of Electrical Engineering, Faculty of Engineering, University of Malaya, 50603 Kuala Lumpur, Malaysia
[c] Department of Computer Science, Faculty of Computer Science and Information Technology, University of Malaya, 50603 Kuala Lumpur, Malaysia

ARTICLE INFO

Article history:
Received 25 November 2014
Received in revised form 8 February 2015
Accepted 4 June 2015
Available online 4 July 2015

Keywords:
Remazol black B
Central composite design
ANFIS
Prediction
Multiple linear regression
Environmental sensing

ABSTRACT

The measurement and prediction of dye concentration is important in the design, planning and management of wastewater treatment. Soft computing techniques can be used as a support tool for analyzing data and making prediction. In this study, Central Composite Design (CCD) and adaptive neuro-fuzzy inference system (ANFIS) are employed to identify and predict the output intensity ratio of light that passes through a plastic optical fiber (POF) sensor in Remazol Black B (RBB) dye solution of different concentrations. The predictive performances of these models are compared to that of the traditional Multiple Linear Regression (MLR). The accuracies of MLR, CCD and ANFIS models are evaluated in terms of square correlation coefficient (R^2), root mean square error (RMSE), value accounted for (VAF), and mean absolute percentage error (MAPE) against the empirical data. It is found that the ANFIS model exhibits higher prediction accuracy than the MLR and CCD models.

1. Introduction

Pollution to watersheds, lakes, streams and rivers is contaminating clean water resources and causing serious problems to the environments. Annually, around 1.3 million tonnes of wastewater is discharged with textile industrial wastewater contributing about 20% of the amount [1]. Rapid monitoring of color wastewater is an essential process and a necessity to textile industry. Recently, many fiber optic sensors have been introduced for various applications that require high sensitivity and fast response. An important type of fiber optic sensors manipulates the interaction between evanescent field and the environment. The sensitivity of these sensors can be enhanced by increasing the amount of the evanescent field generated [2]. For sensors that employ tapered fiber probe, the amount of evanescent field produced can be influenced by the diameter of the tapered section. By adopting thinner fiber, a stronger evanescent field is made available which in turn increases the probe's sensitivity [3]. Compared to sensors made of silica fiber, polymer optical fiber (POF) based sensors are more ductile and easier to handle. Thus, they are suitable for sensing chemical properties such as pH [4], salinity [5], refractive index [6] and others. Dye concentration measurement using evanescent field POF sensor has been reported in studies [7,8]. The use of empirical methods for determining dye concentration does not account for multiple factors involved which are time-consuming and require a large number of experiments and analysis. These limitations impede their widespread use in real applications. Reliable predictive

* Corresponding author. Tel.: +60 3 7967 5300; fax: +60 3 7967 5319.
E-mail address: azizraman@um.edu.my (A.R. Abdul Aziz).

http://dx.doi.org/10.1016/j.measurement.2015.06.019
0263-2241/

models that could correlate the sensed output with the relevant input parameters in a quick and nondestructive manner are useful at the preliminary stage of designing a structure. The use of Multiple Linear Regression (MLR), Central Composite Design (CCD) and adaptive neuro-fuzzy inference system (ANFIS) as a support tool to improve the data analysis and prediction of the dye concentration is first time has been reported to the best of the author's knowledge.

Central Composite Design (CCD) is a useful statistical-based experimental design tool which has been widely used to identify and optimize the performance of complex systems [9,10]. CCD models allow researchers to visualize interaction among independent factors under different experimental conditions. Compared to the conventional approach which considers an individual parameter separately, CCD is able to evaluate the interaction of a few parameters simultaneously. Besides that, it can also yield good estimations of output variable with the help of the established model. Another common modeling tool is neural network type called adaptive neuro-fuzzy inference system (ANFIS). ANFIS is a hybrid intelligent system that possesses good learning and prediction capabilities to deal with uncertainties in many different systems [11,12]. This tool has been used by researchers for identification and real-time prediction of various engineering systems due to its adaptability to a wide range of uncertainties. It has the capability of mapping input and output variables in mathematical form.

In this paper, POF based evanescent field sensor for measuring Remazol Black B (RBB) dye solution is demonstrated. Data analysis and prediction are executed with the help of CCD and ANFIS and compared to that of the traditional MLR model. A plastic optical fiber (POF) with an overall diameter of around 1 mm was subjected to chemical etching process to fabricate three sensors of different diameters (0.35, 0.45 and 0.65 mm). Then light is penetrated into the sensors, while they were immersed in Remazol Black B dye solution of different concentrations (10–70 ppm with 10 ppm increment). The intensity ratio of the transmitted light over the base intensity obtained when the sensor was immersed in distilled water was calculated. Multiple Linear Regression (MLR), Central Composite Design (CCD) and adaptive neuro-fuzzy inference system (ANFIS) models were developed using the data accumulated. The relationships between the output intensity ratios of the different RBB concentrations using tapered POF sensors of different diameters were established. The models treat the intensity ratio as a dependent variable while the sensor's diameter and RBB concentration are regarded as independent variables. Moreover, a performance comparison of MLR, CCD and ANFIS models is carried out to evaluate their prediction accuracy based on root mean square error (RMSE), value accounted for (VAF), mean absolute percentage error (MAPE) and square correlation coefficient (R^2).

2. Materials and methods

2.1. Preparation of sensor probe

The sensor was prepared by etching the POF chemically as demonstrated in reference [13,14]. The deployed POF was sourced from Edmund Optics (Model 02-534, USA). Its refractive indices for core and cladding are 1.492 and 1.402, respectively and its numerical aperture is 0.51. First, acetone was applied to the surface of the POF. It turned the surface into milky white foam which was then removed by the sand paper. This procedure was repeated until the tapered fiber shrank in size to the desired waist diameter. Lastly, the tapered POF fiber was neutralized and cleansed using de-ionized water. In this work, three tapered POFs sensors with stripped region waist diameters of 0.65 mm, 0.45 mm and 0.35 mm were prepared. The length of the tapered section for these sensors was approximately 10 mm. Fig. 1 shows the microscopic image of these three samples wherein a ruler with a scale of 0.5 mm is included for reference. As shown in the figure,

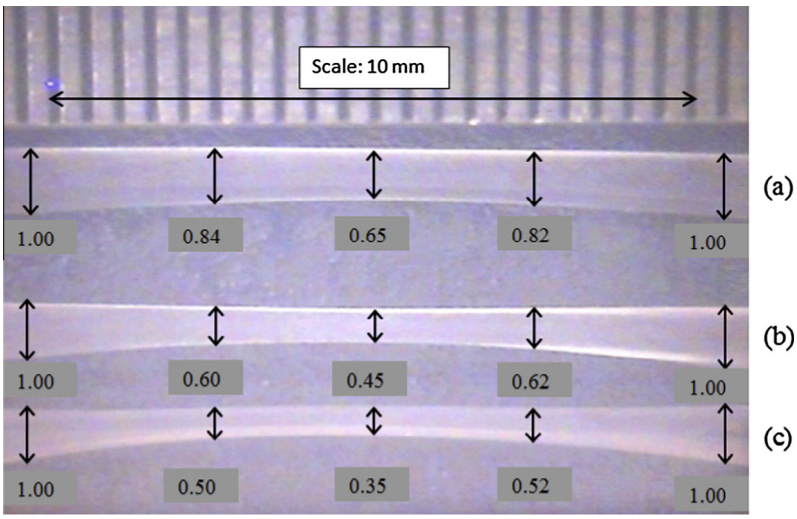

Fig. 1. Profiles of three tapered POF based sensors with diameters of (a) 0.65 mm (b) 0.45 mm and (c) 0.35 mm.