

# Application of Machine Learning For Real-time Evaluation of Salinity (or TDS) in Drinking Water Using Photonic Sensor

Sandip Kumar Roy<sup>1</sup>, Dr. Preeta Sharan<sup>2</sup>

<sup>1</sup>Department of ECE, AMET University, Chennai, 603112, India

5 <sup>2</sup>Department of ECE, Bangalore, 560068, India

*Correspondence to:* Sandip Kumar Roy (sandipr@hotmail.com)

**Abstract.** World is facing unprecedented problem to safeguard 0.4 percent of potable water, which is eventually depleting day-by-day. From literature survey it has been observed that the refractive index (RI) of water changes with change in salinity or total dissolved solids (TDS). In this paper we have proposed an automatic system that can be used for real-time evaluation of salinity or TDS in drinking water. Photonic Crystal (PhC) based ring resonator sensor has been designed and simulated using MEEP tool and Finite Difference Time Domain (FDTD) algorithm. The modelled and designed sensor is highly sensitive to the changes in RI of water sample. This work includes a real-time based natural sequence follower which is a machine learning algorithm of Naïve Bayesian type. A sequence of statistical algorithm implemented in MATLAB with reference to training data to analyse the sample water. Further interfacing has been done using Raspberry Pi device to provide an easy display to show the result of water analysis. The main advantage of the designed sensor with interface is to check whether the salinity or TDS in drinking water is less than 1000 ppm or not.

## 1 Introduction

Drinking water (or potable water) is considered to be safe enough to consume by humans or use it for domestic and medical purposes with low risk of immediate or long term harm. In most of the countries, the salinity of drinking water is restricted to be less than 1000 ppm. Salinity is the measure of concentration of salts in water. Greater concentration of salts in water not only affects the taste of the water but also causes health hazards. TDS includes inorganic salts and organic matter dissolved in water and TDS level between 300 and 600 mg/litre is considered to be good (Fawell et al, 1996). Hence there is necessity for evaluation of water before it is allowed to be consumed. PhC are periodic structures and consist of band gap that restrict propagation of specific frequency range of light. This property enables one to control light and produce effects that are impossible with conventional optics. Behaviour of light propagation in PhC is described by Maxwell's Equations (Yee, 1996).

## 2 Theory

### 2.1 Light propagation in PhC

In a PhC, RI is periodically modulated where periodicity is in the order of wavelength. PhC are periodic structure of dielectric material which allow to propagate a certain frequency range of light (Joannopoulos et al, 1995, 1997) and stop others (forbidden band gap). This unique behaviour of a PhC is used to control the propagation of light (Meade et al, 1992). The deviation of light in a lattice structure can be controlled by defect engineering. The following Eq. (1) explains the movement of light in PhC by solving Maxwell's Electromagnetic Equation.

$$\nabla_x \left( \frac{1}{\epsilon} \nabla_x H \right) = \left( \frac{\omega}{C} \right)^2 H \quad (1)$$

H - photon's magnetic field,  $\epsilon$  - permittivity, C speed of light

$$\epsilon = n^2 \quad (2)$$

n- RI,  $\omega$  - angular frequency.

As in Eq. (1), permittivity of a medium ( $\epsilon$ ) changes as the angular frequency of resonance ( $\omega$ ) changes. Eq. (2) shows  $\epsilon$  is dependent on RI and is the basis to use PhC as sensor (Liu and Salemin, 2012). Methods like photonic band gap method, effective RI method, spectroscopy, optical imaging are available (Fan et al, 2008). Since input variations are significantly low, sensitivity of these methods are less (Nguyen et al, 2011).

The design and simulation of sensor is done by MEEP tool. This is a FDTD simulation software to model electromagnetic systems. To compute transmission flux at each frequency ' $\omega$ ', sampling of continuous electromagnetic field in a finite volume of space is done and is determined by Eq. (3).

$$P(\omega) = Re \int \dot{n} \cdot \int E_{\omega}(X) * \times H_{\omega}(X) d^2X \quad (3)$$

To calculate P( $\omega$ ), following steps used in MEEP tool:

- 1 Compute the integral of the Poynting vector P(t) for each time
- 2 Fourier transform the value in #1.
- 3 Compute flux at the specified regions and frequencies.

### 2.2 Machine Learning Algorithm

Machine learning is an automated in which improvement is done in future action based on learning from past. The key element of this is to devise learning algorithms that do the learning automatically with minimum human actions. The algorithm in machine learning allows the developed application to come up with its own assessment based on supplied training data (Haug et al, 2010).

Naive Bayes Algorithm is a classification technique based on "Bayes" theorem with an assumption of independence among predictors (Rish, 2001). Naive Bayes classifier assumes that the presence of a particular feature in a class is unrelated to the

presence of any other feature. Naive Bayes is known for its simplicity to do better than other existing classification methods. Bayes theorem provides a way of calculating posterior probability  $P(c|x)$  from  $P(c)$ ,  $P(x)$  and  $P(x|c)$ . This is defined at Eq. (4):

$$P(c|x) = \frac{P(x|c)P(c)}{P(x)}$$

$$P(c|x) = P(x_1|c) \times P(x_2|c) \times \dots \times P(x_n|c) \times P(c) \quad (4)$$

5  $P(c|x)$  is the Posterior Probability of class (c, target) given predictor (x, attributes).

$P(c)$  is the Prior Probability of Class.

$P(x|c)$  is the likelihood which is the probability of predictor given class.

$P(x)$  is the prior probability of predictor.

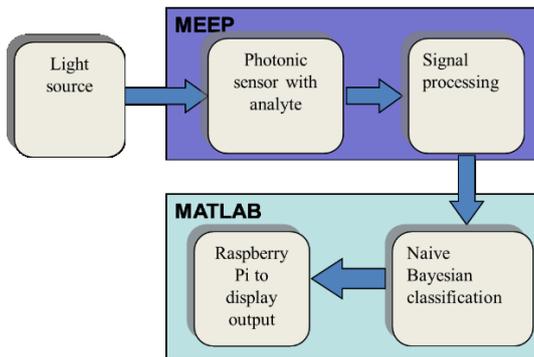
Depending on various attributes, algorithm based on Naive Bayes theorem predicts the probability of different class. This

10 algorithm is used to solve problems having multiple classes.

### 2.3 Methodology

Gaussian light pulse is considered as source for simulation (Oskooi et al, 2010). The simulated data obtained and ready reference data available (training data) are given as input to MATLAB program. The output results are displayed on a LCD screen along with a voice message, using Raspberry Pi kit.

15

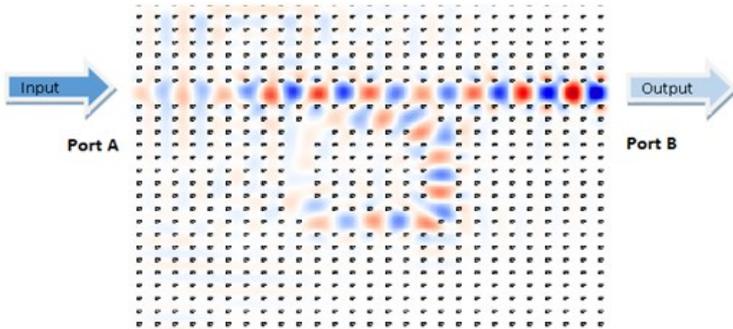


**Figure 1: Evaluation of salinity/TDS in water**

### 3 Sensor design

The objective is to design a 2D PhC based sensor (Akahane et al, 2003) for water analysis. The refractive indices of water with different salinity/TDS were used and simulations were carried out for the variations in properties of the sample for each constituents (Sharan et al, 2013, Lavanya et al, 2014). Shift in output transmitted power and frequency is observed. Figure 1 shows PhC based sensor design and light propagation.

20



**Figure 2: Design of two-dimensional PhC line defect**

The design specifications are:

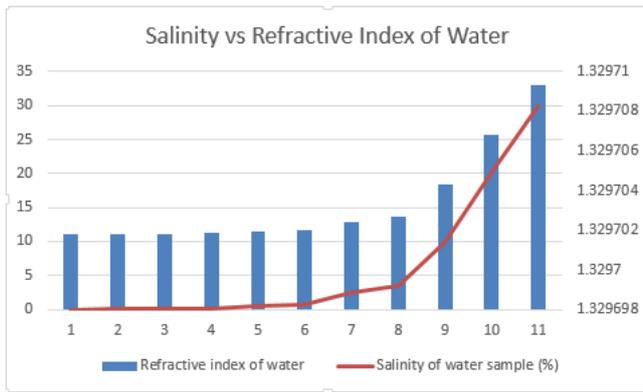
1. Rods in air configuration
- 5 2. Lattice Constant: 'a' = 1.
3. Rod's radius 'r' = 0.2 $\mu$ m
4. Silicon slab's di-electric constant ' $\epsilon$ ' =12
5. Di-electric constant of sample used for simulation in place of air.
6. Light source type used, Gaussian pulse (centre frequency 0.295 and width 0.1).
- 10 7. Wavelength of light 1350nm.
8. Height of rods considered as infinity

#### 4 Sensor Simulation result analysis

Table 1 shows RI changes in the order of  $10^{-5}$  with change in % of salinity of water.

Salinity of water sample (%)	Refractive index of water
0.01	1.329701796
0.05	1.329701806
0.1	1.329701818
0.2	1.329701843
0.5	1.329701918
0.75	1.329701981
2.5	1.329702418
3.5	1.329702668
10	1.329704293
20	1.329706793
30	1.329709293

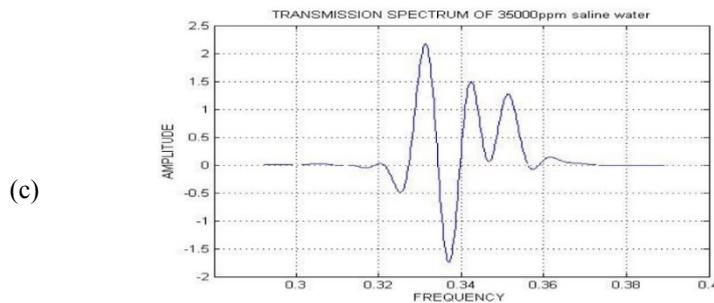
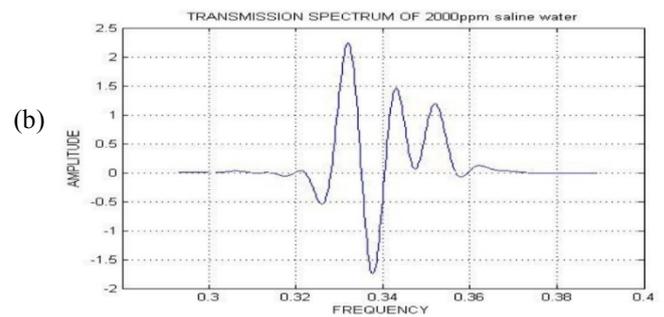
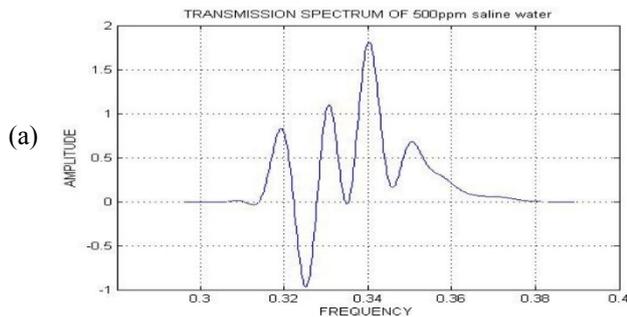
15 **Table 1. Shows variation of RI with %of salinity of water.**



**Figure 3: Salinity vs RI of water.**

Figure 3 shows variation of RI with change in salinity of water. Figure 4(a) indicates that a nascent curve of transmission spectrum until 20 percent of the outcome is exited immediately with a dip being followed. These curves indicate that the light intensity drops abruptly is around 25 percent of the frequency range of the input light wave. These lines of frequency are again in tendency to achieve maxima and do so at exact 50 percent of the frequency spectrum. This is indication of highest possible absorption of intensity of input light wave that may be a cause of polarization on the vicinity of the waveguide of the proposed structure. This velocity of intensity increase will again tend to become sluggish and abruptly embraces an exponential decay, for which the trapping of light in the waveguide begins to throw off certain frequency of harmonic waves that tend to create a disruptive interference of the travelling pulse of Gaussian mode.

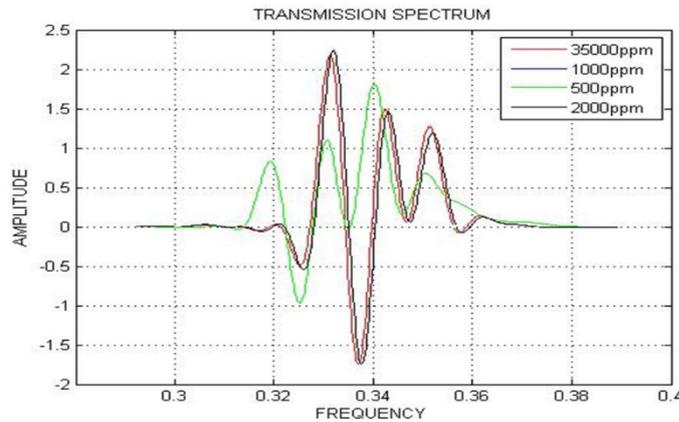
Figure 4 (a) – (c) depicts transmission spectra with distinct shifts in peak frequencies for different samples.



**Figure4 Transmitted spectrum for (a) 500 ppm saline water, (b) <=2000 ppm saline water, (c) 35000 ppm saline water**

In Figure 4 (b) we can observe that the salinity of water being an analyte as compared against Figure 4 (a) is increased in concentration by 300 percent. Here the entire spectra is exactly the reciprocal of Figure 4(a) in that the dip has happened in the first phase of frequency shift, while here the same has happened in the second. Also as against Figure 4(a) the light intensity abruptly decrease before the centre of frequency spectrum is achieved. Thereby the absorption and reflection taken place before the identification of light intensity at the output become noticeable only after does the frequency of spectra are over the central frequency of operation will the light intensity become noticeable twice. The curve remains nascent for around 30 percent of the applied frequency and vigorously excites until 60 percent. This excitement is immediately damped with a scattering time of under 0.5 units of intensity and remains nascent throughout. This signature curve of transmission spectrum is incorporated in the database of the application.

Figure 4(c) is a replica of Figure 4(a) of 4 (b). The only distinguishing factor for the current scenario is the fact that the settling time at the tail and at the horn are squeezed but remain stable for a very long range of frequency of the applied intensity of light. This signature curve of transmission spectrum is incorporated in the database of the application.



**Figure. 5 Transmitted spectrum for water with various salinity levels.**

The Figure 5 shows the overlapping of all the previous spectra to highlight the shift in the frequency and amplitude.

## 25 5 Machine Learning Application design and development

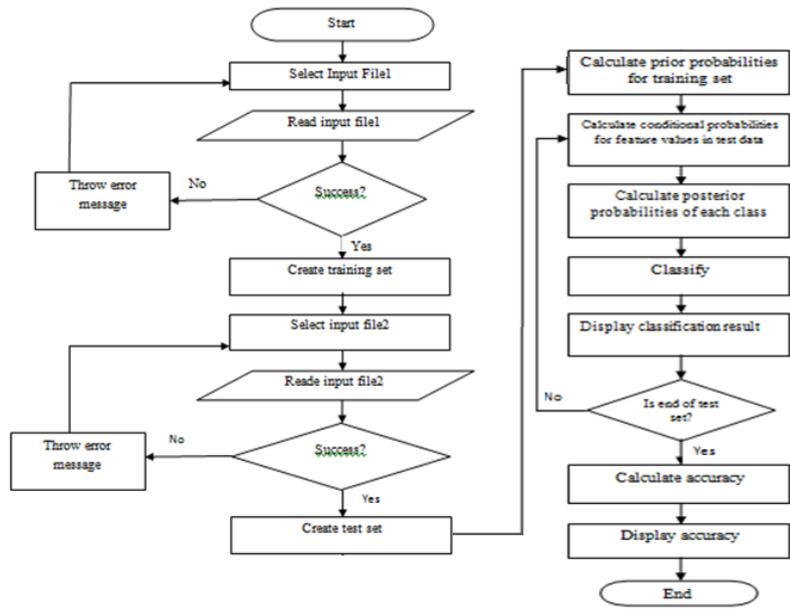
The Naïve Bayes classifier algorithm is developed as MATLAB based desktop application (Garg, 2013). The classifier is designed for using unconditional data provide by the user and is made generalized to read any dataset with unconditional data. Microsoft Excel file is used as input file categorical feature values (non-numerical continuous data). The system is intended to read two input files (.xlsx file) which contains the data set provided by the user. One file contains training set and other test set.

30 Using the training set, the prior probabilities of each class is calculated. Using a single instance from the test set, the conditional probabilities for each feature value is calculated. These values are then used to calculate the posterior probabilities for each class. The class with the highest posterior probability is assigned as the class for that test instance. This process is done on each instance

in the test set. The accuracy of the algorithm is calculated by performing comparison of the class values that are assigned to the class with original class values of that class.

The workflow of the system is shown in the flowchart below in Figure 6.

5  
10  
15  
20



## 6 Workflow of the developed system.

## 25 6 Application of Machine Learning output and result

The MATLAB based application developed was used to detect, analyse and classify the outputs obtained by the PhC based sensors simulated result and were used to evaluate the ppm level of salinity/TDS in drinking water. The algorithm selects the class with highest posterior probability and assigns it to the test data. The accuracy of the algorithm can be obtained by performing comparison of the class assigned to the test data with the actual class of the test data. The accuracy of the classifier is calculated by the number of correct classifications made / the total number of classifications made. The simulated result of salinity/TDS and training data used from the selected USB drives by the developed application (Figure 7a). Based on the salinity check done the observed result is shown in display of Raspberry Pi. If it is greater than or equal to 2000 ppm the display shows “High Salinity/TDS Observed” (Figure 7b) and if ppm is less than equal to 1000 ppm then the display shows “Low salinity/TDS Observed” (Figure 7c).

35

5

(a)



10

15

(b)



20

25

(c)



**Fig. 7 (a) Integrated Device display with Raspberry Pi (red circle), (b) 2000 ppm salinity result (yellow circle) (c) 1000 ppm salinity water result (blue circle)**

## 7 Conclusion

- 30 Proposed paper concludes a design and implementation of an automatic system that can be used for real-time evaluation of potable water. This developed system includes a PhC based ring resonator sensing application interface with LCD display. The result shows, the performance of the sensor is optimum as it can detect RI change in the order of  $10^{-5}$  in drinking water. Even a small % of change in salinity of water can be detected. The application is based on statistical algorithm implemented. Further interfacing has been done using Raspberry Pi device to provide an easy display to show the ppm level of salinity/TDS in water.
- 35 This application is more accurate and do continuous measurement than traditional methods. Because of use of machine learning algorithm the accuracy can be further enhanced by use of further sub classification of TDS. As a future work this approach can be extended to detect whether water can be used for other purposes like farming and industrial use.

## References

- Yee, K.S.: Numerical solution of initial boundary value problems involving Maxwell's equations in isotropic media, *IEEE Trans. Antennas Propag.* AP-14, 302–307, 1996
- Rish, I.: An empirical study of the naive Bayes classifier, IBM Research Report, Computer Science RC 22230 (W0111-014), 2001
- 5 Garg, B.: Design and Development of Naive Bayes Classifier. Diss., North Dakota State University, 2013.
- Liu, Y. and Saleh, H. W. M.: Photonic crystal-based all-optical on-chip sensor, *Opt. Express*, 20, 19912-19920, 2012
- Meade, R.D., Rappe, A.M., Brommer, K.D., Joannopoulos, J.D.: Existence of a photonic band gap in two dimensions, *Appl. Phys., Lett.* 61, 495, 1992
- Joannopoulos, J.D., Meade, R.D., Winn, J. N.: *Photonic Crystal: Modeling of Flow of Light*, Princeton University Press, Princeton, NJ, 1995.
- Fawell, J.K., Lund, U. Mintz, B.: *Guidelines for drinking-water quality*, 2nd ed. Vol. 2. Health criteria and other supporting information. World Health Organization, Geneva, 1996
- 10 Joannopoulos, J.D., Villeneuve, P.R., Fan, S.: Photonic Crystals: Putting a new twist of light, *Nature* 386, 143–149, 1997
- Akahane, Y., Asano, T., Song, B.S., Noda, S.: High-Q photonic nanocavity in a two-dimensional photonic crystal, *Nature*, 425 (6961), 944–947, 2003.
- Fan, X., White, I.M., Shopova, S.I., Zhu, H., Suter, J.D., Sun, Y.: Sensitive optical biosensors for unlabeled targets: a review, *analytica chimica acta* 620 (1), 8–26, 2008
- 15 Haug, Y., Lan, Y., Thomson, S. J., Fang, A., Hoffmann, W. C., Lacey, R. E.: Development of Soft Computing and Application in Agriculture and Biological Engineering, *Computers and Electronics in Agriculture*, 71, 107-127, 2010.
- Oskooi, A.F., Roundy, D., Ibanescu, M., Bermel, P., Joannopoulos, J.D., Johnson, S.G.: MEEP: A flexible free-software package for electromagnetic simulations by the FDTD method, *Comput. Phys. Commun.* 181, 687–702, 2010
- 20 Nguyen, L. V., Vasiliev, M., Alameh, K.: Three-Wave Fiber Fabry-Perot Interferometer for Simultaneous Measurement of Temperature and Water Salinity of Seawater, *IEEE Photonics Technology Letters*, 23(7), 450 – 452, doi:10.1109/LPT.2011.2109057, 2011
- Sharan, P., Deshmukh, P., Roy, S.K.: Mapping of aqua constituents using photonic crystal. Proceedings of the IEEE R10-HTC Conference, 320 – 325, doi: 10.1109/R10-HTC.2013.6669063, 2013
- Lavanya, J., Roy, S.K., Sharan, P.: Design of optical sensor for detection of brininess of water, Proceedings of the IEEE GHTC-SAS Conference, 99-104, doi: 10.1109/GHTC-SAS.2014.6967566, 2014