

APPLICATION OF HYBRID EVOLUTIONARY ALGORITHM (HEA) TO  
DISCOVER THE BEST RULE SET TO EXPLAIN DISSOLVED OXYGEN (D.O.)  
DYNAMICS IN 2 FRESHWATER LAKES

AWANIS AZIZAN

FACULTY OF SCIENCE  
UNIVERSITY OF MALAYA  
KUALA LUMPUR

2012

APPLICATION OF HYBRID EVOLUTIONARY ALGORITHM (HEA) TO  
DISCOVER THE BEST RULE SET TO EXPLAIN DISSOLVED OXYGEN  
(D.O.) DYNAMICS IN 2 FRESHWATER LAKES

AWANIS AZIZAN

(SGJ100003)

SUBMITTED TO  
INSTITUTE OF BIOLOGICAL SCIENCES  
FACULTY OF SCIENCE  
UNIVERSITY OF MALAYA

IN PARTIAL FULLFILLMENT  
OF THE REQUIREMENTS FOR  
THE DEGREE OF MASTER OF BIOINFORMATICS

2012

UNIVERSITI MALAYA

**ORIGINAL LITERARY WORK DECLARATION**

Name of Candidate: AWANIS BINTI AZIZAN (I.C/Passport No:) 871126-06-5648

Registration/Matric No: SGJ100003

Name of Degree: MASTER OF BIOINFORMATICS

TITLE (“APPLICATION OF HYBRID EVOLUTIONARY ALGORITHM (HEA) TO  
DISCOVER THE BEST RULE SET TO EXPLAIN DISSOLVED OXYGEN (D.O.)  
DYNAMICS IN 2 FRESHWATER LAKES”):

Field of Study:

I do solemnly and sincerely declare that:

- (1) I am the sole author/writer of this Work;
- (2) This Work is original;
- (3) Any use of any work in which copyright exists was done by way of fair dealing and for permitted purposes and any excerpt or extract from, or reference to or reproduction of any copyright work has been disclosed expressly and sufficiently and the title of the Work and its authorship have been acknowledged in this Work;
- (4) I do not have any actual knowledge nor do I ought reasonably to know that the making of this work constitutes an infringement of any copyright work;
- (5) I hereby assign all and every rights in the copyright to this Work to the University of Malaya (“UM”), who henceforth shall be owner of the copyright in this Work and that any reproduction or use in any form or by any means whatsoever is prohibited without the written consent of UM having been first had and obtained;
- (6) I am fully aware that if in the course of making this Work I have infringed any Copyright whether intentionally or otherwise, I may be subject to legal action or any other action as may be determined by UM.

Candidate’s Signature

Date:

Subscribed and solemnly declared before,

Witness’s Signature

Date:

Name:

Designation:

## **ABSTRACT**

This project was initiated to study the ability of Hybrid Evolutionary Algorithms (HEA) in predicting the best rule sets to explain the dynamics of dissolved oxygen pattern in 2 freshwater lakes, Tasik Bera (Bera Lake) and Putrajaya Lake. In this study, we would like to observe the correlation between dissolved oxygen and other water quality parameters of the respected lakes that have been generated by the training of the algorithm.

After each data training, analysis on rule sets generated was done and comparison was made against a set of testing data. Relations between each parameter were individually examined on how they reflect to the dynamics of oxygen concentration in the water bodies. The result obtained is compared to the existing research or literature to support the findings.

## **ABSTRAK**

Projek ini telah dimulakan untuk mengkaji keupayaan Hybrid Evolutionary Algorithms (HEA) untuk meramal set peraturan terbaik yang boleh digunakan untuk menerangkan corak kepekatan oksigen terlarut di 2 tasik air tawar, Tasik Bera dan Tasik Putrajaya. Dalam kajian ini, pemerhatian dilakukan untuk mengetahui kaitan antara kepekatan oksigen terlarut dan beberapa parameter kualiti air yang telah dijana oleh algoritma ini untuk kedua-dua tasik berkenaan.

Analisis dilakukan ke atas set peraturan yang telah dijana dan dibandingkan dengan satu set data untuk percubaan. Kaitan antara setiap parameter terhadap kepekatan oksigen terlarut dikenalpasti untuk mengetahui sejauh mana setiap parameter ini mempengaruhi kepekatan oksigen terlarut di dalam air. Keputusan yang diperolehi akan dibandingkan dengan kajian sedia ada untuk menyokong keputusan tersebut.

## ACKNOWLEDGEMENT

First and foremost, I send my gratitude to Allah for the lending me the strength to go through all the challenges and hardship from the very beginning until the completion of this project.

My thanks and heartiest appreciation goes to my project supervisor, Dr. Sorayya Bibi Malek, and her master's students, for giving me enormous amount of assistance and guidance throughout the process of making this project successful.

To my parents and family, thank you very much for your endless encourage, support and sacrifices. Also to my friends, who were very understanding upon my absence in most of our occasionally activities.

Finally, to my supportive colleagues whom were always there for me through thick and thin, thank you so much. Hopefully this journey and experiences we shared would lead us to a successful career and a bright future ahead.

*Awanis Azizan*

## TABLE OF CONTENT

<b>ABSTRACT</b>	<b>iii</b>
<b>ABSTRAK</b>	<b>iv</b>
<b>ACKNOWLEDGMENT</b>	<b>v</b>
<b>LIST OF FIGURES</b>	<b>ix</b>
<b>LIST OF TABLES</b>	<b>xi</b>
<b>LIST OF ABBREVIATIONS</b>	<b>x</b>
<b>Chapter 1: INTRODUCTION</b>	<b>1</b>
1.1 Project Overview	2
1.2 Statement of Problems	2
1.3 Objectives	3
1.4 Project Scope	3
1.5 Limitation and Constraint	4
<b>Chapter 2: LITERATURE REVIEW</b>	<b>5</b>
2.1 Introduction	6
2.2 Evolutionary Algorithms	6
2.2.1 Process of Evolutionary Algorithms	7
2.2.2 Families of Evolutionary Algorithms	8
2.3 Hybrid Evolutionary Algorithm	10
2.3.1 Structure Optimization by GP	11
2.3.2 Parameter Optimization by GA	15

2.4 Water Quality Parameters	17
2.4.1 Dissolved oxygen (DO)	17
2.4.2 pH	17
2.4.3 Water temperature	18
2.4.4 Salinity	18
2.4.5 Turbidity	19
2.4.6 Ammonia (NH <sub>3</sub> N)	19
2.4.7 Nitrate (NO <sub>3</sub> -)	20
2.4.8 Biological Oxygen Demand (BOD)	20
2.4.9 Chemical Oxygen Demand (COD)	21
2.4.10 Chlorophyll-a	21
2.4.11 Conductivity	22
<b>Chapter 3: MATERIALS AND METHODS</b>	<b>23</b>
3.1 Study Sites and Data	24
3.2 Parameter Settings and Measures	28
<b>Chapter 4: RESULTS AND DISCUSSIONS</b>	<b>30</b>
4.1 RULE SET 1 (ELSE-BRANCH)	34
4.1.1 <i>DO vs. Conductivity</i>	34
4.1.2 <i>DO vs. pH</i>	35
4.1.3 <i>DO vs. Chlorophyll-a</i>	36
4.2 RULE SET 2 (THEN-BRANCH)	37
4.2.1 <i>DO vs pH</i>	37
4.2.2 <i>DO vs. Temperature</i>	38
4.2.3 <i>DO vs. E.coli abundance</i>	39
4.3 Comparison with Artificial Neural Network	40



<b>Chapter 5: CONCLUSION</b>	<b>42</b>
5.1 Concluding Statement	43
<b>REFERENCES</b>	<b>44</b>
<b>APPENDIX</b>	<b>47</b>

## LIST OF FIGURES

Figure 2.1: Evolutionary approach to optimization	7
Figure 2.2: Examples of complex representations	8
Figure 2.3: General flowchart of HEA	13
Figure 2.4: Example of vector-level crossover	14
Figure 2.5: Example of tree-level crossover	15
Figure 2.6: Process of Parameter Optimization by GA	16
Figure 3.1 Location Map of Bera Lake	25
Figure 3.2 Location of wetland cells at Putrajaya Wetlands	26
Figure 4.1 Observed and predicted dissolved oxygen value for Putrajaya Lake in 2009	32
Figure 4.2 Observed and predicted dissolved oxygen value for Bera Lake in 2009	33
Figure 4.3 Dissolved oxygen vs. Conductivity	34
Figure 4.4 Dissolved oxygen vs. pH	35
Figure 4.5 Dissolved oxygen vs. Chlorophyll- <i>a</i>	36
Figure 4.6 Dissolved oxygen vs. pH	37
Figure 4.7 Dissolved oxygen vs. water temperature	38
Figure 4.8 Dissolved oxygen vs. abundance of <i>E.coli</i>	39
Figure 4.9 Comparison of actual DO value and predicted DO value from the ANN training	41

## LIST OF TABLES

Table 3.1 Limnological properties of Bera Lake and Putrajaya Lake	27
Table 3.2 Size and storage capacity of Putrajaya Wetlands	28
Table 3.3 Parameter settings of HEA for rule set discovery	29
Table 4.1 Best rule set generated for both Putrajaya and Bera Lake	41

## LIST OF ABBREVIATIONS

HEA: Hybrid Evolutionary Algorithms

EA: Evolutionary Algorithm

EP: Evolutionary Programming

ES: Evolution Strategies

GA: Genetic Algorithms

GP: Genetic Programming

MA: Memetic Algorithms

DO: Dissolved Oxygen

NH<sub>3</sub>N: Ammonia

NO<sub>3</sub><sup>-</sup>: Nitrate

BOD: Biological Oxygen Demand

COD: Chemical Oxygen Demand