

การพยากรณ์ปริมาณออกซิเจนละลายน้ำของแม่น้ำบางปะกง
ด้วยการวิเคราะห์การถดถอยเชิงเส้นพหุคูณ

วารี สีดาคำ

โครงการสถิตินี้เป็นส่วนหนึ่งของการศึกษาตามหลักสูตรวิทยาศาสตรบัณฑิต
สาขาวิชาสถิติ ปีการศึกษา 2561
ภาควิชาคณิตศาสตร์ คณะวิทยาศาสตร์ มหาวิทยาลัยบูรพา

ชื่อเรื่อง

การพยากรณ์ปริมาณออกซิเจนละลายน้ำของแม่น้ำบางปะกงด้วยการวิเคราะห์การถดถอยเชิงเส้นพหุคูณ
PREDICTION OF DISSOLVED OXYGEN DEMAND IN BANGPAKONG RIVER BY MULTIPLE LINEAR
REGRESSION ANALYSIS

ชื่อนิสิต วารีย์ สีตาคำ

รหัสประจำตัวนิสิต 58030628

อาจารย์ที่ปรึกษาหลัก ผู้ช่วยศาสตราจารย์ ดร.จตุภัทร เมฆพ่ายัพ

อาจารย์ที่ปรึกษาร่วม ผู้ช่วยศาสตราจารย์ปรียารัตน์ นาคสุวรรณ

โครงการสถิตินี้เป็นส่วนหนึ่งของการศึกษาตามหลักสูตรวิทยาศาสตรบัณฑิต สาขาวิชาสถิติ
ปีการศึกษา 2561

คณะกรรมการควบคุมโครงการสถิตินี้

.....อาจารย์ที่ปรึกษาหลัก
(ผู้ช่วยศาสตราจารย์ ดร.จตุภัทร เมฆพ่ายัพ)

.....อาจารย์ที่ปรึกษาร่วม
(ผู้ช่วยศาสตราจารย์ปรียารัตน์ นาคสุวรรณ)

คณะกรรมการสอบโครงการสถิตินี้

.....ประธานกรรมการ
(ผู้ช่วยศาสตราจารย์ ดร.จตุภัทร เมฆพ่ายัพ)

.....กรรมการ
(รองศาสตราจารย์ ดร.กิตติการ สายธนู)

.....กรรมการ
(ผู้ช่วยศาสตราจารย์ปรียารัตน์ นาคสุวรรณ)

คณะกรรมการสอบโครงการสถิตินอุมัติให้โครงการสถิตินี้เป็นส่วนหนึ่งของการศึกษาตาม
หลักสูตรวิทยาศาสตรบัณฑิต สาขาวิชาสถิติ ของมหาวิทยาลัยบูรพา

.....ประธานหลักสูตรวิทยาศาสตรบัณฑิต
สาขาวิชาสถิติ

(ผู้ช่วยศาสตราจารย์ปรียารัตน์ นาคสุวรรณ)

วันที่.....เดือน.....พ.ศ.

ประกาศคุณูปการ

โครงการสถิตินี้สำเร็จลุล่วงด้วยดีเนื่องจากผู้ศึกษางานวิจัยได้รับความเมตตาช่วยเหลืออย่างดียิ่งจากผู้ช่วยศาสตราจารย์ ดร.จตุภัทร เมฆพ่ายัพ และผู้ช่วยศาสตราจารย์ปรียาร์ตน์ นาคสุวรรณ ซึ่งเป็นอาจารย์ที่ปรึกษาโครงการสถิตินี้ ที่เสียสละเวลา ให้คำปรึกษา ติดตาม ถ่ายทอดความรู้ และแก้ไขข้อบกพร่องด้วยความดูแลเอาใจจนโครงการสถิตินี้เสร็จสมบูรณ์ ผู้วิจัยขอกราบขอบพระคุณอย่างสูงไว้ ณ โอกาสนี้

นอกจากนี้ผู้ศึกษาวิจัยขอกราบขอบพระคุณรองศาสตราจารย์ ดร.กิตติการ สายธนู กรรมการสอบโครงการสถิตินี้ในครั้งนี้ที่กรุณาให้คำแนะนำเพิ่มเติมและตรวจสอบแก้ไขเพื่อให้โครงการสถิตินี้สมบูรณ์มากยิ่งขึ้น

ขอกราบขอบพระคุณบิดา มารดา และเพื่อนที่คอยให้กำลังใจให้ความช่วยเหลือคำปรึกษาแนะนำจนทำให้โครงการสถิตินี้สำเร็จลุล่วงไปได้ด้วยดี

วารี สีดาคำ

58030628: สาขาวิชา: สถิติ; วท.บ. (สถิติ)

คำสำคัญ: การวิเคราะห์การถดถอยเชิงเส้นพหุคูณ / ปริมาณออกซิเจนละลายน้ำ / แม่น้ำบางปะกง
วารี สื่อดาคำ: การพยากรณ์ออกซิเจนละลายน้ำของแม่น้ำบางปะกงด้วยการวิเคราะห์การ

ถดถอยเชิงเส้นพหุคูณ (PREDICTION OF DISSOLVED OXYGEN DEMAND IN BANGPAKONG RIVER BY
MULTIPLE LINEAR REGRESSION ANALYSIS)

คณะกรรมการควบคุมโครงการ: ผู้ช่วยศาสตราจารย์ ดร.จตุภัทร เมฆพ่ายัพ, ประ.ด. และผู้ช่วยศาสตราจารย์
ปรียารัตน์ นาคสุวรรณ, วท.ม. 40 หน้า. ปีการศึกษา 2561.

บทคัดย่อ

โครงการสถิตินี้ศึกษาปริมาณออกซิเจนละลายน้ำของแม่น้ำบางปะกงด้วยการวิเคราะห์การถดถอย
เชิงเส้นพหุคูณเพื่อสร้างสมการสำหรับพยากรณ์ปริมาณออกซิเจนละลายน้ำในแม่น้ำบางปะกงโดยใช้ปัจจัย
มลพิษทางน้ำ ได้แก่ ค่าความเป็นกรด-ด่าง (pH) ปริมาณฟอสฟอรัสทั้งหมด (TP) ปริมาณความสกปรกในรูป
สารอินทรีย์ (BOD) และอุณหภูมิ (WT) ผลการศึกษาพบว่า สมการถดถอยที่ใช้พยากรณ์ปริมาณออกซิเจน
ละลายน้ำของแม่น้ำบางปะกง คือ $\widehat{DO} = -12.57 + 2.347pH + 0.437BOD$ มีค่าคลาดเคลื่อนของการถดถอย
เท่ากับ 0.8710 และค่าสัมประสิทธิ์การตัดสินใจที่ปรับค่าแล้วเท่ากับ 54.15 จากนั้นทำการตรวจสอบ
ประสิทธิภาพของสมการถดถอยพบว่า มีค่าคลาดเคลื่อนสัมบูรณ์เฉลี่ยเท่ากับ 1.0742 และรากของค่า
คลาดเคลื่อนกำลังสองเฉลี่ยเท่ากับ 1.4469

58030628: MAJOR: STATISTICS; B.SC. (STATISTICS)

KEYWORDS: MULTIPLE LINEAR REGRESSION ANALYSIS

DISSOLVED OXYGEN DEMAND / BANGPAKONG RIVER

WAREE SEEDAKAM: PREDICTION OF DISSOLVED OXYGEN DEMAND IN
BANGPAKONG RIVER BY MULTIPLE LINEAR REGRESSION ANALYSIS

ADVISORS: ASSISTANT PROFESSOR JATUPAT MEKPARYUP, Ph.D. AND ASSISTANT
PROFESSOR PREEYARAT NARKSUWAN, M.Sc. 40 P. ACADEMIC YEARS 2018.

ABSTRACT

The project is to study of Oxygen demand in Bangpakong River by multiple liner regression analysis for predicting Oxygen demand in Bangpakong River using factors affecting water pollution as follows pH-value (pH), Total phosphorus (TP), Biochemical oxygen demand (BOD), Water temperature (WT). The results of this study found that the regression equation for predict Oxygen demand in Bangpakong River is

$\widehat{DO} = -12.57 + 2.347pH + 0.437BOD$ with 0.8710 standard error of regression and 54.15 adjusted coefficient of determination. Then the performance of the regression equation was evaluated with mean absolute error of of 1.0742 and root mean square error 1.4469.

สารบัญ

	หน้า
บทคัดย่อภาษาไทย.....	ง
บทคัดย่อภาษาอังกฤษ.....	จ
สารบัญ.....	ฉ
สารบัญตาราง.....	ช
สารบัญภาพ.....	ฅ
บทที่ 1 บทนำ.....	1
1.1 ความเป็นมาและความสำคัญของปัญหา.....	1
1.2 วัตถุประสงค์ของการศึกษา.....	3
1.3 ประโยชน์ที่คาดว่าจะได้รับจากการศึกษา.....	3
1.4 ขอบเขตการศึกษา.....	3
1.5 นิยามศัพท์เฉพาะ.....	3
บทที่ 2 เอกสารและงานวิจัยที่เกี่ยวข้อง.....	4
1. แนวคิดและทฤษฎีที่เกี่ยวข้อง.....	4
1.1 การวิเคราะห์สหสัมพันธ์.....	4
1.2 การวิเคราะห์การถดถอยเชิงเส้นพหุคูณ.....	4
1.3 การตรวจสอบข้อสมมติเบื้องต้นของการวิเคราะห์การถดถอย.....	5
1.4 การตรวจสอบค่าสังเกตมีอิทธิพล.....	6
1.5 ค่าซีพีของมอลลิโลว์.....	7
1.6 การคัดเลือกตัวแปรอิสระและตัวแบบที่เหมาะสม.....	7
2. งานวิจัยที่เกี่ยวข้อง.....	9
บทที่ 3 วิธีดำเนินการศึกษา.....	13
ตอนที่ 1 การหาความสัมพันธ์ระหว่างปริมาณออกซิเจนละลายน้ำกับปัจจัยมลพิษทางน้ำ.....	13
ตอนที่ 2 การสร้างสมการถดถอยเพื่อพยากรณ์ปริมาณออกซิเจนละลายน้ำ.....	13
ตอนที่ 3 การตรวจสอบประสิทธิภาพของสมการถดถอย	14
บทที่ 4 ผลการศึกษา.....	15
ตอนที่ 1 ผลการหาความสัมพันธ์ระหว่างปริมาณออกซิเจนละลายน้ำกับปัจจัยมลพิษทางน้ำ.....	15
ตอนที่ 2 ผลการสร้างสมการถดถอยเพื่อพยากรณ์ปริมาณออกซิเจนละลายน้ำ.....	16
ตอนที่ 3 ผลการตรวจสอบประสิทธิภาพของสมการถดถอย	32

สารบัญ (ต่อ)

	หน้า
บทที่ 5 สรุปและอภิปรายผลการศึกษา.....	34
5.1 สรุปผลการศึกษา.....	34
5.2 อภิปรายผลการศึกษา.....	35
5.3 ข้อเสนอแนะ.....	35
บรรณานุกรม.....	36
ภาคผนวก.....	38

สารบัญตาราง

ตารางที่	หน้า
1. ค่าสัมประสิทธิ์สหสัมพันธ์ระหว่าง DO กับปัจจัยมลพิษทางน้ำ.....	16
2. สมการถดถอยที่เป็นไปได้ทั้งหมดสำหรับพยากรณ์ DO.....	16
3. การวิเคราะห์ความแปรปรวนของสมการถดถอยสมการที่ 11.....	17
4. สัมประสิทธิ์การถดถอยและการทดสอบสมการถดถอยสมการที่ 11.....	17
5. ค่า $ DFFITS_i $ ของสมการถดถอยสมการที่ 11.....	20
6. การวิเคราะห์ความแปรปรวนของสมการถดถอยสมการที่ 12.....	21
7. สัมประสิทธิ์การถดถอยและการทดสอบสมการถดถอยสมการที่ 12.....	21
8. ค่า $ DFFITS_i $ ของสมการถดถอยสมการที่ 12.....	24
9. การวิเคราะห์ความแปรปรวนของสมการถดถอยสมการที่ 13.....	24
10. สัมประสิทธิ์การถดถอยและการทดสอบสมการถดถอยสมการที่ 13.....	25
11. ค่า $ DFFITS_i $ ของสมการถดถอยสมการที่ 13.....	27
12. การวิเคราะห์ความแปรปรวนของสมการถดถอยสมการที่ 14.....	28
13. สัมประสิทธิ์การถดถอยและการทดสอบสมการถดถอยสมการที่ 14.....	28
14. การวิเคราะห์ความแปรปรวนของสมการถดถอยเมื่อค่าคลาดเคลื่อนกำลังสองเป็นตัวแปรตาม	31
15. ค่าวิกฤตไคกำลังสองด้วยองศาเสรีเท่ากับ 1.....	31
16. ค่า VIF ของสมการถดถอยสมการที่ 14.....	31
17. การตรวจสอบประสิทธิภาพของสมการถดถอย.....	33
18. ค่าวิกฤตของตัวสถิติแอนเดอร์สัน-ดาร์ลิง.....	39
19. ค่าวิกฤตของตัวสถิติเดอ์บิน-วัตสัน ที่ระดับนัยสำคัญ 0.05.....	40

สารบัญญภาพ

ภาพที่	หน้า
1. แผนภาพค่าดัชนีคุณภาพน้ำจากจุดตรวจวัดคุณภาพน้ำในแม่น้ำบางปะกง.....	2
2. แผนภาพการกระจายระหว่าง DO กับปัจจัยมลพิษทางน้ำ.....	15
3. แผนภาพความน่าจะเป็นปรกติของส่วนเหลือจากสมการถดถอยสมการที่ 11.....	18
4. แผนภาพอนุกรมเวลาของส่วนเหลือจากสมการถดถอยสมการที่ 11.....	19
5. แผนภาพความน่าจะเป็นปรกติของส่วนเหลือจากสมการถดถอยสมการที่ 12.....	22
6. แผนภาพอนุกรมเวลาของส่วนเหลือจากสมการถดถอยสมการที่ 12.....	23
7. แผนภาพความน่าจะเป็นปรกติของส่วนเหลือจากสมการถดถอยสมการที่ 13.....	25
8. แผนภาพอนุกรมเวลาของส่วนเหลือจากสมการถดถอยสมการที่ 13.....	26
9. แผนภาพความน่าจะเป็นปรกติของส่วนเหลือจากสมการถดถอยสมการที่ 14.....	29
10. แผนภาพอนุกรมเวลาของค่าประมาณของส่วนเหลือจากสมการถดถอยสมการที่ 14.....	30
11. แผนภาพการกระจายระหว่างค่าประมาณของส่วนเหลือและค่าประมาณจาก สมการถดถอยสมการที่14.....	30
12. แผนภาพอนุกรมเวลาระหว่างค่าพยากรณ์และค่าจริงของปริมาณออกซิเจนละลายน้ำ.....	32

บทที่ 1

บทนำ

1. ความเป็นมาและความสำคัญของปัญหา

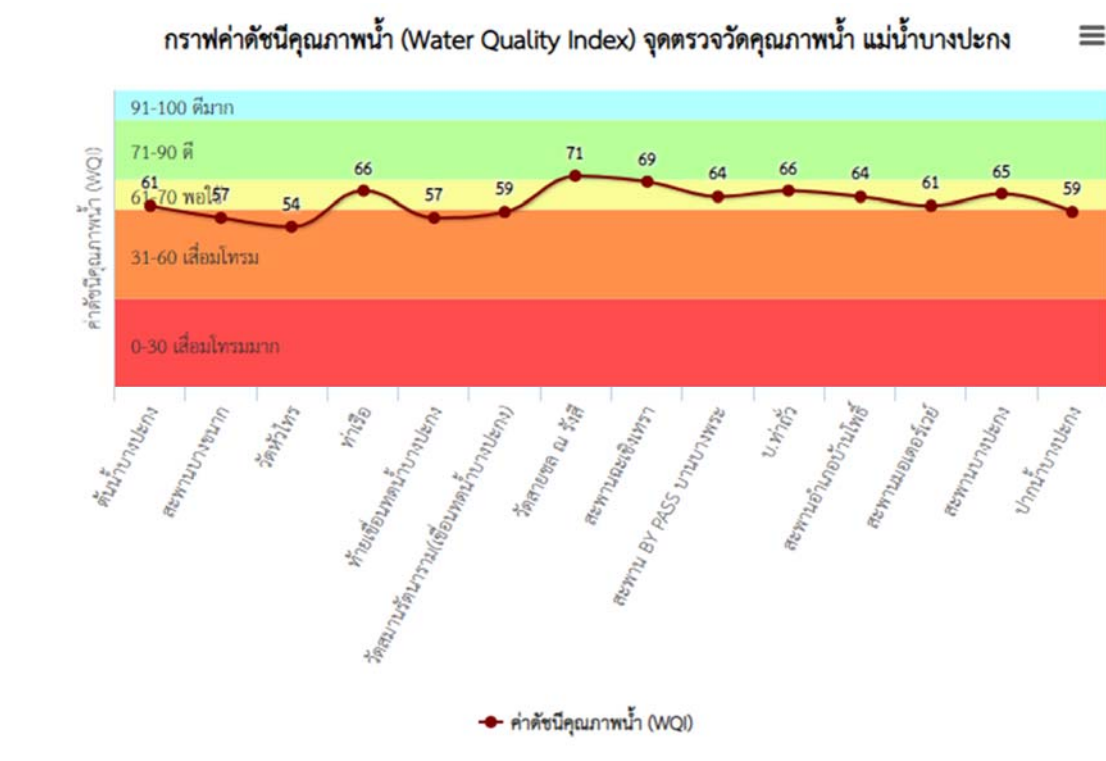
น้ำเป็นปัจจัยที่สำคัญของสิ่งมีชีวิตในการใช้อุปโภคและบริโภคซึ่งมีความหลากหลายในการใช้ประโยชน์ น้ำที่มีความสำคัญกับสิ่งมีชีวิตมากที่สุดคือ น้ำผิวดินซึ่งประกอบด้วยแหล่งน้ำเค็มและแหล่งน้ำจืดซึ่งเป็นแหล่งน้ำที่มนุษย์สามารถใช้ประโยชน์ในการดำรงชีวิตได้มากที่สุด แหล่งน้ำจืดที่สำคัญคือ แม่น้ำ ลำคลอง ฝาย หนอง คลอง บึง (จันทน์ เบ็ญจวรรณ, 2562) ซึ่งในอดีตธรรมชาติสามารถปรับสมดุลของแหล่งน้ำไม่ให้เกิดการเน่าเสียได้แต่ในปัจจุบันมีการเพิ่มขึ้นของประชากรจนกลายเป็นแหล่งชุมชนซึ่งมีความหนาแน่นทำให้มีการเจริญทั้งการเกษตรกรรมและอุตสาหกรรมซึ่งเป็นต้นเหตุสำคัญของการปล่อยสารมลพิษลงสู่แหล่งน้ำธรรมชาติทำให้แหล่งน้ำธรรมชาติเกิดการเน่าเสีย (กรมควบคุมมลพิษ, 2550)

ในพื้นที่ภาคตะวันออกเฉียงเหนือของประเทศไทยมีแหล่งน้ำที่สำคัญหลายแห่ง ได้แก่ แม่น้ำระยอง แม่น้ำประแสร์ แม่น้ำจันทบุรี แม่น้ำพังรัต แม่น้ำตราด แม่น้ำเวฬุ และแม่น้ำบางปะกงซึ่งเป็นหนึ่งในแม่น้ำที่มีความสำคัญมากของภาคตะวันออกเฉียงเหนือทั้งยังเป็นแม่น้ำสายที่ยาวที่สุดในภาคตะวันออกเฉียงเหนือ แม่น้ำบางปะกงมีจุดตรวจสอบวัดคุณภาพน้ำทั้งหมด 14 สถานี ได้แก่

- สถานีปากน้ำบางปะกง อำเภอบางปะกง จังหวัดฉะเชิงเทรา
- สถานีสะพานบางปะกง อำเภอบางปะกง จังหวัดฉะเชิงเทรา
- สถานีสะพานมอเตอร์เวย์ อำเภอบางปะกง จังหวัดฉะเชิงเทรา
- สถานีสะพานอำเภอบ้านโพธิ์ อำเภอบางปะกง จังหวัดฉะเชิงเทรา
- สถานีบ้านท่าถั่ว อำเภอบางปะกง จังหวัดฉะเชิงเทรา
- สถานีสะพาน BY PASS อำเภอมือง จังหวัดฉะเชิงเทรา
- สถานีสะพานฉะเชิงเทรา อำเภอมือง จังหวัดฉะเชิงเทรา
- สถานีวัดสายชล ณ รังสี อำเภอมือง จังหวัดฉะเชิงเทรา
- สถานีวัดสมานรัตนาราม อำเภอมือง จังหวัดฉะเชิงเทรา
- สถานีท้ายเขื่อนทดน้ำบางปะกง อำเภอมือง จังหวัดฉะเชิงเทรา
- สถานีท่าเรือ อำเภอบางคล้า จังหวัดฉะเชิงเทรา
- สถานีวัดหัวไทร อำเภอบางคล้า จังหวัดฉะเชิงเทรา
- สถานีสะพานบางขนาก อำเภอบางน้ำเปรี้ยว จังหวัดฉะเชิงเทรา
- สถานีต้นน้ำบางปะกง อำเภอบ้านสร้าง จังหวัดปราจีนบุรี

ปัจจุบันคุณภาพน้ำของแม่น้ำบางปะกงที่ตรวจวัดจากดัชนีคุณภาพน้ำ (Water Quality Index: WQI) พบว่า คุณภาพน้ำของแม่น้ำบางปะกงอยู่ในเกณฑ์เสื่อมโทรมถึงร้อยละ 52.08 (สำนักจัดการน้ำคุณภาพน้ำ กรมควบคุมมลพิษ, 2561) ปัจจัยสำคัญที่ทำให้คุณภาพน้ำของแม่น้ำบางปะกงมีความเสื่อมโทรมได้แก่ น้ำเสีย

จากครัวเรือน น้ำเสียจากภาคเกษตรกรรม และน้ำเสียจากภาคอุตสาหกรรม น้ำเสียเหล่านี้ถูกระบายลงสู่แม่น้ำบางปะกงก่อนได้รับการบำบัด (กลุ่มงานกลยุทธ์ศาสตร์และข้อมูลเพื่อการพัฒนาจังหวัด, 2560) จากที่ได้กล่าวมาข้างต้นการตรวจสอบคุณภาพน้ำพิจารณาจากค่า WQI ซึ่งคำนวณมาจากสารเคมีทั้งหมด 5 ตัว ได้แก่ ความสกปรกในรูปแบบสารอินทรีย์ (Biochemical Oxygen Demand: BOD) ออกซิเจนละลายน้ำ (Dissolved Oxygen: DO) แบคทีเรียกลุ่มโคลิฟอร์ม (Fecal Coliform Bacteria: FCB) แบคทีเรียในรูปโคลิฟอร์มทั้งหมด (Total Coliform Bacteria: TCB) และแอมโมเนีย-ไนโตรเจน ($\text{NH}_3 - \text{N}$) จะเห็นได้ว่าหนึ่งในสารเคมีที่ใช้ในการคำนวณค่าดัชนีคุณภาพน้ำคือ DO เนื่องจาก DO เป็นสารเคมีที่ใช้วัดปริมาณออกซิเจนละลายน้ำ หากปริมาณออกซิเจนในน้ำน้อยสิ่งมีชีวิตไม่สามารถอาศัยในแหล่งน้ำได้เนื่องจากไม่มีออกซิเจนไปหล่อเลี้ยง และเมื่อไม่มีออกซิเจนจะเกิดการย่อยสลายสารอินทรีย์โดยไม่ใช้ออกซิเจนจนเกิดเป็นก๊าซไฮโดรเจนซัลไฟด์ซึ่งมีกลิ่นเหม็นและทำให้คุณภาพของแม่น้ำเสื่อม (วรชาดา บัวไพร, 2555)



ภาพที่ 1 ค่าดัชนีคุณภาพน้ำจากจุดตรวจวัดคุณภาพน้ำของแม่น้ำบางปะกง

จากภาพที่ 1 แสดงค่า WQI จากสถานีตรวจวัดในแม่น้ำบางปะกงทั้ง 14 สถานี มีสถานีที่จัดอยู่ในเกณฑ์คุณภาพน้ำเสื่อมโทรมซึ่งมีค่า WQI อยู่ในช่วง 31 – 60 คะแนน จำนวน 5 สถานี (สำนักจัดการน้ำคุณภาพน้ำ กรมควบคุมมลพิษ, 2561) ดังนี้

1. สถานีวัดสมานรัตนารามมีค่า WQI เท่ากับ 59
2. สถานีปากน้ำบางปะกงมีค่า WQI เท่ากับ 59
3. สถานีท้ายเขื่อนทดน้ำบางปะกงมีค่า WQI เท่ากับ 57

4. สถานีสะพานบางขนากมีค่า WQI เท่ากับ 57

5. สถานีวัดหัวไทรมีค่า WQI เท่ากับ 54

จากเหตุผลที่กล่าวมาข้างต้นผู้วิจัยจึงสนใจที่จะศึกษาการพยากรณ์ปริมาณออกซิเจนละลายน้ำในแม่น้ำบางปะกงด้วยการถดถอยพหุคูณ (Multiple linear regression: MLR) ซึ่งผลการศึกษาอาจทำให้ทราบปัจจัยที่ใช้พยากรณ์ปริมาณออกซิเจนละลายน้ำของแม่น้ำบางปะกงเพื่อใช้เป็นแนวทางในการควบคุมและแก้ไขปัญหามลพิษทางน้ำในเบื้องต้นต่อไปได้

2. วัตถุประสงค์ของการศึกษา

2.1 เพื่อศึกษาความสัมพันธ์ของปริมาณออกซิเจนละลายน้ำกับปัจจัยมลพิษทางน้ำของแม่น้ำบางปะกง

2.2 สร้างสมการถดถอยเพื่อพยากรณ์ปริมาณออกซิเจนละลายน้ำของแม่น้ำบางปะกง

2.3 ตรวจสอบประสิทธิภาพของสมการถดถอยที่ใช้พยากรณ์ปริมาณออกซิเจนละลายน้ำของแม่น้ำบางปะกง

3. ประโยชน์ที่คาดว่าจะได้รับจากการศึกษา

3.1 ทำให้ทราบถึงปัจจัยมลพิษทางน้ำที่มีผลต่อปริมาณออกซิเจนละลายน้ำของแม่น้ำบางปะกง

3.2 ได้สมการถดถอยสำหรับพยากรณ์ปริมาณออกซิเจนละลายน้ำของแม่น้ำบางปะกง

3.3 ใช้เป็นแนวทางให้กับหน่วยงานท้องถิ่นในการควบคุมและแก้ไขปัญหามลพิษทางน้ำในเบื้องต้น

4. ขอบเขตการศึกษา

ข้อมูลที่ใช้ในการศึกษาเป็นข้อมูลจากสถานีจุดตรวจวัดคุณภาพน้ำของแม่น้ำบางปะกงตั้งแต่เดือนกุมภาพันธ์ พ.ศ. 2551 ถึงเดือนสิงหาคม พ.ศ. 2561 ซึ่งเก็บรวบรวมข้อมูลจากสำนักจัดการคุณภาพน้ำ กรมควบคุมมลพิษ

5. นิยามศัพท์เฉพาะ

ออกซิเจนละลายน้ำ (Dissolved Oxygen: DO) คือค่าที่บ่งบอกถึงปริมาณก๊าซออกซิเจนที่ละลายในน้ำ หากปริมาณออกซิเจนในน้ำมากบ่งบอกได้ว่าคุณภาพน้ำดีสิ่งมีชีวิตสามารถอาศัยอยู่ได้ ส่วนใหญ่ปริมาณออกซิเจนในน้ำมีค่าอยู่ระหว่าง 5-7 มิลลิกรัม/ลิตร ถ้าหากมีค่าน้อยกว่า 3 มิลลิกรัมต่อลิตร บ่งบอกได้ว่าคุณภาพน้ำเน่าเสีย และหากมีค่าน้อยกว่า 2 มิลลิกรัม/ลิตร สิ่งมีชีวิตไม่สามารถอยู่ในแหล่งน้ำได้ (มงคล ต๊ะอุ้น และสุทธิพงศ์ เปรื่องคำ, 2546)

บทที่ 2

เอกสารและงานวิจัยที่เกี่ยวข้อง

การศึกษาครั้งนี้มีวัตถุประสงค์เพื่อพยากรณ์ปริมาณออกซิเจนละลายน้ำ (Dissolved Oxygen: DO) ของแม่น้ำบางปะกงโดยมีแนวคิด ทฤษฎี และงานวิจัยที่เกี่ยวข้องดังนี้

1. แนวคิดและทฤษฎีที่เกี่ยวข้อง

1.1 การวิเคราะห์สหสัมพันธ์ (Correlation analysis)

ค่าวัดทางสถิติที่นิยมใช้อธิบายความสัมพันธ์เชิงเส้นระหว่างตัวแปรสองตัวแปรโดยพิจารณาจากค่าสัมประสิทธิ์สหสัมพันธ์มีสัญลักษณ์คือ ρ สามารถบอกถึงทิศทางของความสัมพันธ์ได้ การประมาณค่า ρ ประมาณโดยสัมประสิทธิ์สหสัมพันธ์ตัวอย่างมีสัญลักษณ์คือ R สามารถคำนวณได้ดังนี้

$$R = \frac{S_{xy}}{\sqrt{S_{xx}S_{yy}}} \quad (1)$$

เมื่อ S_{xy} คือ ความแปรปรวนของตัวอย่างระหว่างตัวแปร x กับตัวแปร y

S_{xx} คือ ความแปรปรวนของตัวอย่างของตัวแปร x

S_{yy} คือ ความแปรปรวนของตัวอย่างของตัวแปร y

คุณสมบัติของ R มีดังนี้

1. มีค่าระหว่าง -1 ถึง 1

2. หาก $R = -1$ หรือ 1 แสดงว่าตัวแปรทั้งสองมีความสัมพันธ์เชิงเส้นกันอย่างสมบูรณ์

3. หาก R มีค่าใกล้ -1 หรือ 1 แสดงว่าตัวแปรทั้งสองมีความสัมพันธ์เชิงเส้นกันมาก

4. หาก $R = 0$ แสดงว่าตัวแปรทั้งสองอาจไม่มีความสัมพันธ์เชิงเส้นกัน หรือ อาจมีความสัมพันธ์กันใน

รูปแบบที่ไม่ใช่เส้นตรง

5. หาก R มีค่าเป็นบวกหมายความว่า ตัวแปรทั้งสองมีความสัมพันธ์ไปในทิศทางเดียวกัน เมื่อตัวแปรหนึ่งค่าเพิ่มขึ้น ค่าของอีกตัวแปรก็เพิ่มขึ้นด้วย

6. หาก R มีค่าเป็นลบหมายความว่า ตัวแปรทั้งสองมีความสัมพันธ์ไปในทิศทางตรงกันข้าม เมื่อตัวแปรหนึ่งค่าเพิ่มขึ้น ค่าของอีกตัวแปรหนึ่งจะลดลง

1.2 การวิเคราะห์การถดถอยเชิงเส้นพหุคูณ (Multiple linear regression analysis)

การวิเคราะห์การถดถอยเชิงเส้นพหุคูณเป็นการวิเคราะห์เพื่อแสดงความสัมพันธ์ระหว่างตัวแปรอิสระและตัวแปรตามอย่างน้อยหนึ่งตัว สำหรับตัวแบบเชิงเส้นพหุคูณแสดงได้ดังนี้

$$y_i = \beta_0 + \beta_1 x_{i1} + \beta_2 x_{i2} + \dots + \beta_k x_{ik} + \varepsilon_i \quad (2)$$

เมื่อ y_i คือค่าสังเกตที่ i ของตัวแปรตามโดยที่ $i = 1, 2, \dots, n$

$\beta_0, \beta_1, \dots, \beta_k$ คือ พารามิเตอร์ของสัมประสิทธิ์การถดถอย

$x_{i1}, x_{i2}, \dots, x_{ik}$ คือ ค่าสังเกตที่ i ของตัวแปรอิสระตัวที่ $1, 2, 3, \dots, k$ เป็นตัวแปรที่ทราบค่าหรือมีค่าคงที่

ε_i คือ ค่าคลาดเคลื่อนของค่าสังเกตที่ i

1.3 การตรวจสอบข้อสมมติของการวิเคราะห์การถดถอย

1.3.1 ตรวจสอบการแจกแจงปกติของค่าคลาดเคลื่อน

การตรวจสอบการแจกแจงปกติของค่าคลาดเคลื่อนสามารถตรวจสอบการทดสอบของแอนเดอร์สัน-ดาร์ลิง (Anderson-Darling test) นำเสนอโดย Lewis (1961) สำหรับสถิติทดสอบสามารถคำนวณได้ดังนี้

$$AD = -n - \sum_{i=1}^n \left(\frac{2i-1}{n} \right) \left\{ \ln F(y_i) + \ln [1 - F(y_{n+1-i})] \right\} \quad (3)$$

พิจารณาปฏิเสธสมมติฐานหลักเมื่อค่า AD ที่คำนวณได้มีค่ามากกว่าค่าวิกฤต หรือค่าพี (P-value) มีค่าน้อยกว่าระดับนัยสำคัญที่กำหนด

1.3.2 ตรวจสอบความแปรปรวนคงที่ของค่าคลาดเคลื่อน

การตรวจสอบความแปรปรวนคงที่ของค่าคลาดเคลื่อนสามารถตรวจสอบด้วยการทดสอบของบรูช-เพแกน (Breush-Pagan's test) นำเสนอโดย Breush and Pagan (1979) สำหรับสถิติทดสอบสามารถคำนวณได้ดังนี้

$$BP = \frac{SSR^* / 2}{(SSE/n)^2} \quad (4)$$

เมื่อ SSR^* แทนผลบวกกำลังสองของการถดถอยโดยตัวแปรตามคือ $\hat{\varepsilon}_i^2$ และตัวแปรอิสระคือ x

SSE แทนผลบวกกำลังสองของค่าคลาดเคลื่อนระหว่างตัวแปรตาม y และตัวแปรอิสระ x

พิจารณาปฏิเสธสมมติฐานหลักเมื่อค่า $BP \geq \chi_{1-\alpha; (1)}^2$ หรือค่าพิน้อยกว่าระดับนัยสำคัญที่กำหนด

1.3.3 ตรวจสอบความเป็นอิสระกันของค่าคลาดเคลื่อน

การตรวจสอบความเป็นอิสระกันของค่าคลาดเคลื่อนสามารถตรวจสอบด้วยการทดสอบเดอร์บิน-วัตสัน (Durbin-Watson test) นำเสนอโดย Durbin and Watson (1950, 1951, 1971) สำหรับสถิติทดสอบสามารถคำนวณได้ดังนี้

$$d = \frac{\sum_{t=2}^T (\hat{\varepsilon}_t - \hat{\varepsilon}_{t-1})^2}{\sum_{t=1}^T \hat{\varepsilon}_t^2} \quad (5)$$

เกณฑ์การพิจารณาค่าสถิติทดสอบ d พิจารณาจากค่าขอบเขตบน d_L และขอบเขตล่าง d_U จากตารางค่าวิกฤตของสถิติทดสอบ Durbin-Watson

ถ้า $d < d_L$ แล้วจะปฏิเสธสมมุติฐานหลัก

ถ้า $d > d_U$ แล้วจะยอมรับสมมุติฐานหลัก

ถ้า $d_L \leq d \leq d_U$ แล้วไม่สามารถสรุปผลการทดสอบได้

1.3.4 ตรวจสอบความเป็นอิสระกันของตัวแปรอิสระ

การตรวจสอบความเป็นอิสระกันของตัวแปรอิสระสามารถตรวจสอบได้โดยพิจารณาจากค่า VIF (Variance Inflation Factors) สามารถคำนวณได้ดังนี้

$$VIF = C_{jj} = \frac{1}{1 - R_{j/\text{others}}^2}; j = 1, 2, \dots, k \quad (5)$$

เมื่อ $R_{j/\text{others}}^2$ แทนค่าสัมประสิทธิ์การถดถอยระหว่างตัวแปรอิสระ x_j และสมมติอื่นของเมทริกซ์ $X_{n \times p}$ ทั้งหมด เกณฑ์การพิจารณาค่า VIF มีดังนี้

$VIF \leq 5$ หมายความว่า มีความสัมพันธ์เชิงเส้นแบบพหุน้อย

$5 < VIF < 10$ หมายความว่า มีความสัมพันธ์เชิงเส้นแบบพหุปานกลาง

$VIF > 10$ หมายความว่า มีความสัมพันธ์เชิงเส้นแบบพหุมาก

1.4 การตรวจสอบค่าสังเกตมีอิทธิพล (Influential observation)

การตรวจสอบค่าสังเกตที่มีอิทธิพลนำเสนอโดย Welsch and Kuh (1977) แสดงได้ดังนี้

$$DFFITS_i = \frac{\hat{y}_i - \hat{y}_{(i)}}{\sqrt{MSE_{(i)} h_{ii}}} \quad (6)$$

เมื่อ \hat{y}_i คือค่าประมาณของ y_i ที่ได้จากการถดถอยของค่าสังเกตทั้งหมด n

$\hat{y}_{(i)}$ คือค่าประมาณของ y_i ที่ได้จากการถดถอยของค่าสังเกตทั้งหมด $n-1$ ค่าโดยไม่รวมค่า

สังเกตที่ i

$MSE_{(i)}$ คือค่า MSE ที่ได้จากการถดถอยของค่าสังเกตทั้งหมด $n-1$ ค่าโดยไม่รวมค่าสังเกตที่ i

h_{ii} คือค่าเลเวอร์เรจ (Leverage) ของค่าสังเกตที่ i ของค่าสังเกตทั้งหมด n ค่า

เกณฑ์การพิจารณาสำหรับ n ที่มีขนาดใหญ่ของค่า $DFITS_i$ คือ $|DFITS_i| > 2\sqrt{\frac{k+1}{n}}$ หากค่า $DFITS_i$ มีลักษณะดังกล่าวแสดงว่าค่าสังเกตที่ i เป็นค่าสังเกตที่มีอิทธิพล

1.5 ค่าซีพีของมอลส์โลว์

ค่าซีพีของมอลส์โลว์ (Mallows' Cp) เป็นเกณฑ์การพิจารณาการหาเซตย่อยของตัวแปรอิสระที่เหมาะสมเพื่อให้ได้ตัวแบบการถดถอยที่ดีที่สุด สามารถคำนวณได้ดังนี้

$$C_p = \frac{SSE_p}{MSE} - n + 2p \quad (7)$$

เมื่อ SSE_p คือ ผลบวกกำลังสองของค่าคลาดเคลื่อนที่ได้จากตัวแบบการถดถอยที่มีพารามิเตอร์ p ตัว

MSE คือ ค่าคลาดเคลื่อนกำลังสองเฉลี่ยที่ได้จากการถดถอยที่มีตัวแปรอิสระ k ตัว

n คือ จำนวนข้อมูลทั้งหมด

เกณฑ์การพิจารณามีดังนี้

1. เลือกตัวแบบการถดถอยที่มีค่า C_p น้อย
2. เลือกตัวแบบการถดถอยที่มีค่า C_p ใกล้เคียงกับจำนวนพารามิเตอร์

1.6 การคัดเลือกตัวแปรอิสระและตัวแบบที่เหมาะสม

ในการวิเคราะห์การถดถอยอาจมีตัวแปรอิสระมากมายที่เกี่ยวข้องกับตัวแปรตาม ตัวแปรอิสระบางตัวอาจมีความสัมพันธ์เชิงเส้นกับตัวแปรตามหรืออาจไม่มีความสัมพันธ์เชิงเส้นกัน ดังนั้นการคัดเลือกตัวแปรอิสระเป็นขั้นตอนที่สำคัญเพื่อให้ได้ตัวแบบการถดถอยที่ดีที่สุด

1.6.1 การถดถอยเซตย่อยที่ดีที่สุด (Best subsets regression)

เป็นการคัดเลือกตัวแปรอิสระโดยการสร้างตัวแบบการถดถอยที่มีตัวแปรอิสระเพียงหนึ่งตัวไปจนถึงตัวแบบการถดถอยที่มีตัวแปรอิสระ k ตัว จากนั้นทำการเลือกตัวแบบที่ดีที่สุดโดยพิจารณาจากค่าซีพีของมอลส์โลว์ (Mallows' Cp) และค่าคลาดเคลื่อนมาตรฐานของการถดถอย (Standard error of regression: S) ที่มีค่าน้อย

1.6.2 การกำจัดแบบถดถอยหลัง (Backward elimination)

เป็นการคัดเลือกตัวแปรอิสระโดยการสร้างตัวแบบการถดถอยที่มีตัวแปรอิสระจำนวน k ตัว แล้วพิจารณาค่าพีของตัวแปรอิสระที่มากที่สุดและมีความมากกว่าระดับนัยสำคัญ จากนั้นทำการตัดตัวแปรอิสระดังกล่าวออกจากตัวแบบการถดถอยและทำการสร้างตัวแบบการถดถอยใหม่ซึ่งประกอบด้วยตัวแปรอิสระที่เหลือ $k - 1$ ตัว ทำซ้ำเช่นนี้ไปจนกว่าไม่มีตัวแปรอิสระที่มีความมากกว่าระดับนัยสำคัญ

1.6.3 การเลือกแบบไปข้างหน้า (Forward selection)

เป็นการคัดตัวแปรอิสระโดยการสร้างตัวแบบที่ไม่มีตัวแปรอิสระใดอยู่ในตัวแบบการถดถอยเลยแล้วพิจารณาเพิ่มตัวแปรอิสระหนึ่งตัวจากตัวแปรอิสระทั้งหมด k เข้าไปในตัวแบบโดยพิจารณาค่าพีของตัวแปรอิสระที่น้อยที่สุดและมีความน้อยกว่าระดับนัยสำคัญ จากนั้นทำการเพิ่มตัวแปรอิสระดังกล่าวเข้าไปในตัวแบบการถดถอย ทำซ้ำเช่นนี้ไปจนกว่าไม่มีตัวแปรอิสระที่มีความน้อยกว่าระดับนัยสำคัญ

1.6.4 การถดถอยทีละขั้น (Stepwise regression)

เป็นการเลือกตัวแปรอิสระโดยการสร้างตัวแบบที่ไม่มีตัวแปรอิสระใดอยู่ในตัวแบบการถดถอยเลยแล้วพิจารณาเพิ่มตัวแปรอิสระหนึ่งตัวโดยพิจารณาจากค่าพีของตัวแปรอิสระที่น้อยที่สุดและมีความน้อยกว่าระดับนัยสำคัญเข้าไปในตัวแบบการถดถอย ขั้นตอนต่อมาทำการเพิ่มตัวแปรอิสระตัวที่สองเข้าไปในตัวแบบการถดถอยเช่นเดียวกับวิธีการเพิ่มตัวแปรอิสระตัวแรก จากนั้นใช้การกำจัดแบบถอยหลังเพื่อพิจารณาตัดตัวแปรอิสระออกจากตัวแบบการถดถอยโดยพิจารณาจากค่าพีของตัวแปรอิสระที่มีความมากกว่าระดับนัยสำคัญ จากนั้นทำการสร้างตัวแบบการถดถอยจากตัวแปรอิสระที่เหลืออยู่ในตัวแบบการถดถอยใหม่อีกครั้ง ทำซ้ำเช่นนี้ไปเรื่อยๆจนกว่าไม่สามารถเพิ่มหรือลดตัวแปรอิสระได้อีก

2. งานวิจัยที่เกี่ยวข้อง

Chen, W. B., & Liu, W. C. (2015) ศึกษาการเปรียบเทียบดัชนีคุณภาพน้ำปริมาณออกซิเจนละลายน้ำ (Dissolved Oxygen: DO) ปริมาณฟอสฟอรัสทั้งหมด (Total Phosphorus: TP) ปริมาณคลอโรฟิลล์ เอ (Chlorophyll *a*: Chl *a*) และความโปร่งใสของน้ำ (Secchi disk depth: SD) โดยใช้ 1) โครงข่ายประสาทเทียมเรเดียลเบสฟังก์ชัน (Radial basis function neural network: RBFN) 2) ระบบอนุมานฟัซซี่โครงข่ายปรับตัวได้ (Adaptive neural fuzzy inference systems: ANFIS) และ 3) ตัวแบบถดถอยเชิงเส้นพหุคูณ (Multiple linear regression model: MLR) ใช้พารามิเตอร์คุณภาพน้ำทั้งหมด 8 ตัว คือ อุณหภูมิ (Water temperature: WT) ค่าความเป็นกรด-ด่าง (pH value: pH) ปริมาณสารแขวนลอย (Suspended solids: SS) การนำไฟฟ้า (Electrical conductivity: EC) ความขุ่นในน้ำ (Turbidity: TB) ความกระด้างของน้ำ (Total hardness: TH) ความเป็นด่าง (Total alkalinity: TA) ไนเตรต (Nitrate nitrogen: NO₃) ความสกปรกในรูปแบบสารอินทรีย์ (Biochemical oxygen demand: BOD) ศึกษาในพื้นที่อ่างเก็บน้ำ Mingder ตั้งอยู่ในภาคกลางของประเทศไทยได้หวน ข้อมูลที่เข้าแบ่งเป็นข้อมูลสำหรับการเรียนรู้ทั้งหมด 272 ค่า และข้อมูลสำหรับตรวจสอบทั้งหมด 117 ค่า ได้สมการดังนี้

$$\begin{aligned} \text{สมการที่ 1} \quad \widehat{DO} &= 2.8083\text{pH} + 0.0887\text{Chla} - 0.074\text{WT} - 0.0045\text{TP} - 0.4928\text{NO}_3 \\ &\quad - 0.0055\text{BOD} - 13.851 \end{aligned}$$

$$\text{สมการที่ 2} \quad \widehat{TP} = 0.039\text{TB} + 0.07\text{SS} - 4.4717\text{pH} - 0.2463\text{DO} - 0.0553\text{EC} + 70.2402$$

$$\begin{aligned} \text{สมการที่ 3} \quad \widehat{\text{Chla}} &= 1.1423\text{pH} + 0.05528\text{DO} + 0.3228\text{WT} + 0.1079\text{BOD} - 1.7734\text{NO}_3 \\ &\quad - 0.0103\text{TH} - 12.8466 \end{aligned}$$

$$\begin{aligned} \text{สมการที่ 4} \quad \widehat{SD} &= -0.0119\text{SS} - 0.0025\text{EC} + 0.0042\text{TB} - 0.0016\text{TH} - 0.0010\text{TA} \\ &\quad - 0.0017\text{TP} + 2.4449 \end{aligned}$$

จากนั้นตรวจสอบประสิทธิภาพการพยากรณ์โดยพิจารณาจากค่าคลาดเคลื่อนสัมบูรณ์เฉลี่ย (Mean absolute error: MAE) ค่าสัมประสิทธิ์การตัดสินใจ (Coefficient of determination: R²) และค่าสัมประสิทธิ์สหสัมพันธ์ (Correlation coefficient: R) ผลการวิจัยแบ่งตามวิธีวิเคราะห์ดังนี้

ตัวแบบ RBFN

- สำหรับข้อมูลการทดสอบใช้เซตการเรียนรู้ทั้งหมด
 - สมการที่ 1: $R=0.77$, $RMSE=1.97$, $MAE=1.49$
 - สมการที่ 2: $R=0.75$, $RMSE=11.77$, $MAE=8.89$
 - สมการที่ 3: $R=0.79$, $RMSE=4.39$, $MAE=3.09$
 - สมการที่ 4: $R=0.68$, $RMSE=0.39$, $MAE=0.29$
- สำหรับข้อมูลชุดตรวจสอบทั้งหมด
 - สมการที่ 1: $R=0.79$, $RMSE=1.96$, $MAE=1.45$
 - สมการที่ 2: $R=0.73$, $RMSE=9.91$, $MAE=7.95$
 - สมการที่ 3: $R=0.75$, $RMSE=4.08$, $MAE=2.86$
 - สมการที่ 4: $R=0.76$, $RMSE=0.31$, $MAE=0.23$

ตัวแบบ ANFIS

- สำหรับข้อมูลการทดสอบใช้เซตการเรียนรู้ทั้งหมด
 - สมการที่ 1: $R=0.85$, $RMSE=1.74$, $MAE=1.30$
 - สมการที่ 2: $R=0.86$, $RMSE=9.92$, $MAE=6.45$
 - สมการที่ 3: $R=0.77$, $RMSE=4.67$, $MAE=3.19$
 - สมการที่ 4: $R=0.80$, $RMSE=0.31$, $MAE=0.21$
- สำหรับข้อมูลชุดตรวจสอบทั้งหมด
 - สมการที่ 1: $R=0.88$, $RMSE=1.32$, $MAE=0.92$
 - สมการที่ 2: $R=0.86$, $RMSE=7.42$, $MAE=5.44$
 - สมการที่ 3: $R=0.83$, $RMSE=3.05$, $MAE=2.25$
 - สมการที่ 4: $R=0.89$, $RMSE=0.24$, $MAE=0.15$

ตัวแบบ MLR

- สำหรับข้อมูลการทดสอบใช้เซตการเรียนรู้ทั้งหมด
 - สมการที่ 1: $R=0.67$, $RMSE=2.33$, $MAE=1.81$
 - สมการที่ 2: $R=0.55$, $RMSE=15.29$, $MAE=9.82$
 - สมการที่ 3: $R=0.59$, $RMSE=5.85$, $MAE=3.96$
 - สมการที่ 4: $R=0.43$, $RMSE=0.45$, $MAE=0.34$
- สำหรับข้อมูลชุดตรวจสอบทั้งหมด
 - สมการที่ 1: $R=0.64$, $RMSE=2.41$, $MAE=1.81$
 - สมการที่ 2: $R=0.31$, $RMSE=14.70$, $MAE=9.99$
 - สมการที่ 3: $R=0.55$, $RMSE=5.92$, $MAE=4.18$
 - สมการที่ 4: $R=0.36$, $RMSE=0.53$, $MAE=0.37$

Antanasijević, D., Pocajt, V., Povrenović, D., Perić-Grujić, A., & Ristić, M. (2013) ศึกษาการเปรียบเทียบระหว่างโครงข่ายประสาทเทียม (ANN) 3 โครงสร้าง ได้แก่ 1) โครงข่ายประสาทเทียมการถดถอยทั่วไป (General regression neural network: GRNN) 2) โครงข่ายประสาทเทียมแบบส่งค่าย้อนกลับ (Backpropagation neural network: BPNN) และ 3) โครงข่ายประสาทเทียมแบบวนซ้ำ (Recurrent neural network: RNN) กับการถดถอยเชิงพหุคูณ (Multiple linear regression: MLR) สำหรับพยากรณ์ปริมาณออกซิเจนละลายน้ำ (Dissolved Oxygen) โดยมีพารามิเตอร์คือ อุณหภูมิ (Water temperature) ความเป็นกรด-ด่าง (pH) ตัวนำไฟฟ้า (Conductivity) และการไหลของน้ำ (Flow) ข้อมูลที่ใช้ในการวิจัยมาจากแม่น้ำดานูบในทวีปยุโรปตั้งแต่ปี พ.ศ. 2547 ถึงปี พ.ศ. 2551 เป็นชุดการเรียนรู้ และใช้ข้อมูลปี พ.ศ. 2552 เป็นชุดตรวจสอบ จากนั้นทำการเปรียบเทียบประสิทธิภาพการพยากรณ์ด้วยค่าคลาดเคลื่อนสัมบูรณ์เฉลี่ย (Mean absolute error: MAE) รากที่สองของค่าคลาดเคลื่อนกำลังสองเฉลี่ย (Root mean square error: RMSE) และค่าสัมประสิทธิ์การตัดสินใจ (Coefficient of determination: R^2) ผลการวิจัยแบ่งตามวิธีวิเคราะห์ดังนี้

โครงข่ายประสาทเทียม (ANN)

ข้อมูลชุดการเรียนรู้

- ตัวแบบ GRNN: MAE= 0.07 , RMSE= 0.19
- ตัวแบบ BPNN: MAE= 0.55, RMSE= 0.74
- ตัวแบบ RNN: MAE= 0.44, RMSE= 0.65

ข้อมูลชุดการตรวจสอบ

- ตัวแบบ GRNN: MAE= 0.60, RMSE= 0.78, $R^2 = 0.85$
- ตัวแบบ BPNN: MAE= 0.72, RMSE= 0.83, $R^2 = 0.76$
- ตัวแบบ RNN: MAE= 0.49, RMSE= 0.59, $R^2 = 0.87$

ตัวแบบการถดถอยเชิงเส้นพหุคูณ (MLR)

ข้อมูลชุดการเรียนรู้

- ตัวแบบที่ MLR: MAE= 1.35, RMSE= 1.63

ข้อมูลชุดการตรวจสอบ

- ตัวแบบที่ MLR: MAE= 1.28, RMSE= 1.49, $R^2 = 0.66$

Nemati, S., Fazelifard, M. H., Terzi, Ö., & Ghorbani, M. A. (2015) ศึกษาเปรียบเทียบประสิทธิภาพระหว่างการถดถอยเชิงเส้นพหุคูณ (Multilinear linear regression: MLR) ระบบอนุมานฟัซซี่โครงข่ายปรับตัวได้ (Adaptive neural fuzzy inference systems: ANFIS) และโครงข่ายประสาทเทียม (Artificial neural networks: ANN) เพื่อพยากรณ์ปริมาณออกซิเจนละลายน้ำ (Dissolved Oxygen) โดยมีพารามิเตอร์คุณภาพน้ำคือ คลอไรด์ (Chloride: Cl), ค่าความเป็นกรด-ด่าง (pH) การนำไฟฟ้า (Electrical conductivity: EC) อุณหภูมิ (Temperature: Temp) ไนโตรเจนออกไซด์ (Nitrite nitrogen: $\text{NO}_2 - \text{N}$)

ไนเตรตไนโตรเจน (Nitrate nitrogen: $\text{NO}_3\text{-N}$) แอมโมเนียไนโตรเจน (Ammonia nitrogen: $\text{NH}_4\text{-N}$) และ ฟอสฟอรัสทั้งหมด (Total phosphorous: TP) ข้อมูลจากแม่น้ำไทเป ประเทศฮ่องกง ตั้งแต่ปี พ.ศ. 2534 ถึงปี พ.ศ. 2554 ข้อมูลมีทั้งหมด 252 ค่า แบ่งเป็นข้อมูลชุดการเรียนรู้จำนวน 214 ค่า และข้อมูลชุดตรวจสอบจำนวน 38 ค่า และเปรียบเทียบประสิทธิภาพของแต่ละตัวแบบโดยพิจารณาค่าสัมประสิทธิ์สหสัมพันธ์ (Correlation coefficient: CC) รากที่สองของค่าคลาดเคลื่อนกำลังสองเฉลี่ย (Root mean square error: RMSE) และสัมประสิทธิ์ประสิทธิภาพของ Nash–Sutcliffe (Nash–Sutcliffe efficiency coefficient: E) ผลการวิจัยแบ่งตามวิธีวิเคราะห์ดังนี้

ข้อมูลชุดการเรียนรู้

- ตัวแบบ MLR: CC= 0.550, RMSE= 0.961, E = 0.303
- ตัวแบบ ANFIS: CC= 0.791, RMSE= 0.707, E = 0.632
- ตัวแบบ ANN: CC= 0.796, RMSE= 0.699, E = 0.631

ข้อมูลชุดตรวจสอบ

- ตัวแบบ MLR: CC= 0.681, RMSE= 0.657, E = 0.438
- ตัวแบบ ANFIS: CC= 0.645, RMSE= 0.973, E = -0.233
- ตัวแบบ ANN: CC= 0.798, RMSE= 0.529, E = 0.636

บทที่ 3

วิธีดำเนินการศึกษา

โครงการสถิตินี้มีวัตถุประสงค์เพื่อศึกษาความสัมพันธ์ระหว่างปริมาณออกซิเจนละลายน้ำ (Dissolved Oxygen: DO) กับปัจจัยมลพิษทางน้ำ ได้แก่ ค่าความเป็นกรด-ด่าง (pH) ฟอสฟอรัสทั้งหมด (Total phosphorus: TP) ปริมาณความสกปรกในรูปสารอินทรีย์ (Biochemical oxygen demand: BOD) และ อุณหภูมิน้ำ (Water temperature: WT) ของแม่น้ำบางปะกงตั้งแต่เดือนกุมภาพันธ์ พ.ศ. 2551 ถึงเดือน สิงหาคม พ.ศ. 2561 ข้อมูลทั้งหมดเก็บรวบรวมจากสำนักงานจัดการคุณภาพน้ำ กรมควบคุมมลพิษ การดำเนินการศึกษาแบ่งออกเป็น 3 ตอน ดังนี้

ตอนที่ 1 การหาความสัมพันธ์ระหว่างปริมาณออกซิเจนละลายน้ำกับปัจจัยมลพิษทางน้ำ

ตอนที่ 2 การสร้างสมการถดถอยเพื่อพยากรณ์ปริมาณออกซิเจนละลายน้ำ

ตอนที่ 3 การตรวจสอบประสิทธิผลของสมการถดถอย

ตอนที่ 1 การหาความสัมพันธ์ระหว่างปริมาณออกซิเจนละลายน้ำกับปัจจัยมลพิษทางน้ำ

สำหรับตอนที่ 1 เป็นการดำเนินการตามวัตถุประสงค์ข้อที่ 1 คือ ศึกษาความสัมพันธ์ระหว่าง DO กับ ปัจจัยมลพิษทางน้ำของแม่น้ำบางปะกง ได้แก่ ค่าความเป็นกรด-ด่าง อุณหภูมิน้ำ ฟอสฟอรัสทั้งหมด และ ปริมาณความสกปรกในรูปสารอินทรีย์ โดยคำนวณหาค่าสัมประสิทธิ์สหสัมพันธ์

ตอนที่ 2 การสร้างสมการถดถอยเพื่อพยากรณ์ปริมาณออกซิเจนละลายน้ำ

สำหรับตอนที่ 2 เป็นการดำเนินการตามวัตถุประสงค์ข้อที่ 2 คือ การสร้างสมการถดถอยสำหรับ พยากรณ์ปริมาณออกซิเจนละลายน้ำของแม่น้ำบางปะกง ดำเนินตามขั้นตอนดังนี้

1. วิเคราะห์ความสัมพันธ์ระหว่างปริมาณออกซิเจนละลายน้ำและปัจจัยมลพิษทางน้ำของแม่น้ำบางปะกงตั้งแต่เดือนกุมภาพันธ์ พ.ศ. 2551 ถึงเดือนสิงหาคม พ.ศ. 2561

2. สร้างสมการถดถอยเพื่อพยากรณ์ปริมาณออกซิเจนละลายน้ำของแม่น้ำบางปะกงตั้งแต่เดือนกุมภาพันธ์ พ.ศ. 2551 ถึงเดือนสิงหาคม พ.ศ. 2561

สำหรับการคัดเลือกสมการถดถอยจะใช้วิธีเซตย่อยที่ดีที่สุดที่สร้างจากปัจจัยมลพิษทางน้ำทั้ง 4 ตัว จาก ได้สมการถดถอยทั้งหมด 16 สมการ (2^4) สำหรับตัวแบบการถดถอยที่มีปัจจัยมลพิษทางน้ำทั้งหมดแสดงดังนี้

$$DO = \beta_0 + \beta_1 pH + \beta_2 TP + \beta_3 BOD + \beta_4 WT + \epsilon_{DO} \quad (8)$$

เมื่อ DO คือ ค่า DO เฉลี่ยรายวัน

pH คือ ค่า pH เฉลี่ยรายวัน

TP คือ ค่า TP เฉลี่ยรายวัน

BOD คือ ค่า BOD เฉลี่ยรายวัน

WT คือ ค่า WT เฉลี่ยรายวัน

β_i

ε_{DO} คือ ค่าคลาดเคลื่อนของตัวแบบการถดถอยสำหรับพยากรณ์ DO

3. ตรวจสอบข้อสมมุติของการวิเคราะห์ถดถอยเชิงเส้นพหุคูณดังนี้

- 3.1 ตรวจสอบการแจกแจงปกติของค่าคลาดเคลื่อน
- 3.2 ตรวจสอบความเป็นอิสระกันของค่าคลาดเคลื่อน
- 3.3 ตรวจสอบความคงที่ของความแปรปรวนของค่าคลาดเคลื่อน
- 3.4 ตรวจสอบความเป็นอิสระกันของตัวแปรอิสระ

ตอนที่ 3 การตรวจสอบประสิทธิภาพของสมการถดถอย

สำหรับตอนที่ 3 เป็นการดำเนินการตามวัตถุประสงค์ข้อที่ 3 คือ ตรวจสอบประสิทธิภาพของสมการถดถอยที่ใช้พยากรณ์ปริมาณออกซิเจนละลายน้ำของแม่น้ำบางปะกง ดังนี้

1. ตรวจสอบประสิทธิภาพของสมการถดถอยด้วยค่าคลาดเคลื่อนสัมบูรณ์เฉลี่ย (Mean absolute error: MAE) สามารถคำนวณได้ดังนี้

$$MAE = \frac{\sum_{i=1}^n |O_i - P_i|}{n} \quad (9)$$

2. ตรวจสอบประสิทธิภาพของสมการถดถอยด้วยค่ารากของค่าคลาดเคลื่อนกำลังสองเฉลี่ย (Root mean square error: RMSE) สามารถคำนวณได้ดังนี้

$$RMSE = \sqrt{\frac{\sum_{i=1}^n (O_i - P_i)^2}{n}} \quad (10)$$

เมื่อ O_i คือ ค่าจริงปริมาณออกซิเจนละลายน้ำของแม่น้ำบางปะกงจากค่าสังเกตที่ i

P_i คือ ค่าพยากรณ์ปริมาณออกซิเจนละลายน้ำของแม่น้ำบางปะกงจากค่าสังเกตที่ i

n คือ จำนวนข้อมูลทั้งหมด

บทที่ 4

ผลการศึกษา

ผลการศึกษาโครงการสถิตินี้มี 3 ตอน ได้แก่

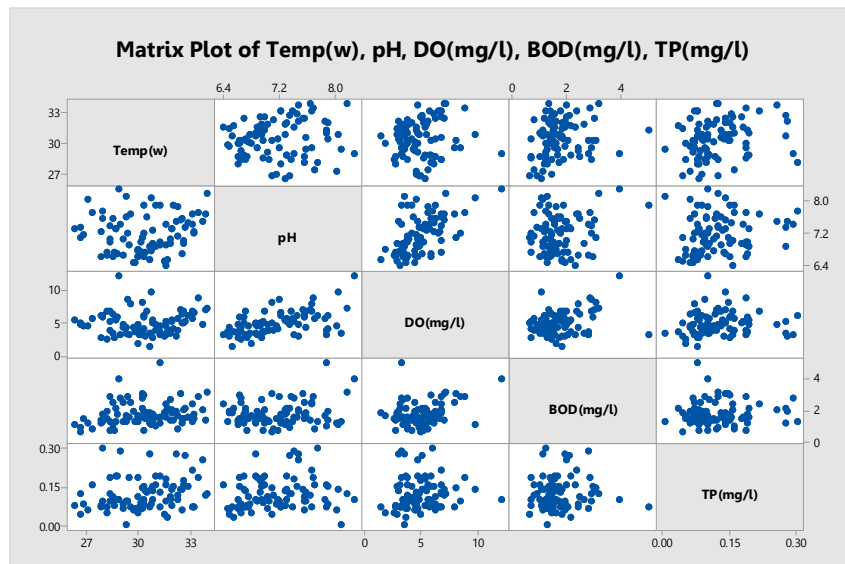
ตอนที่ 1 ผลการหาความสัมพันธ์ระหว่างปริมาณออกซิเจนละลายน้ำกับปัจจัยมลพิษทางน้ำ

ตอนที่ 2 ผลการสร้างสมการถดถอยเพื่อพยากรณ์ปริมาณออกซิเจนละลายน้ำ

ตอนที่ 3 ผลการตรวจสอบประสิทธิภาพของสมการถดถอย

ตอนที่ 1 ผลการหาความสัมพันธ์ระหว่างปริมาณออกซิเจนละลายน้ำกับปัจจัยมลพิษทางน้ำ

การศึกษาความสัมพันธ์ระหว่าง DO กับปัจจัยมลพิษทางน้ำ ได้แก่ pH, TP, BOD และ WT ด้วยแผนภาพการกระจายให้ผลดังภาพที่ 2



ภาพที่ 2 แผนภาพการกระจายระหว่าง DO กับปัจจัยมลพิษทางน้ำ

จากภาพที่ 2 แสดงให้เห็นว่าการกระจายของจุดมีลักษณะคล้ายเส้นตรงหลายคู่ เช่น

ระหว่าง DO และ pH

ระหว่าง BOD และ WT

ระหว่าง TP และ WT

ระหว่าง DO และ BOD เป็นต้น

เมื่อศึกษาความสัมพันธ์ระหว่าง DO กับปัจจัยมลพิษทางน้ำ ได้แก่ pH, TP, BOD และ WT ด้วยค่าสัมประสิทธิ์สหสัมพันธ์ให้ผลดังตารางที่ 2

ตารางที่ 2 ค่าสัมประสิทธิ์สหสัมพันธ์ระหว่าง DO กับปัจจัยมลพิษทางน้ำ

	WT	pH	DO	BOD
pH	0.108 (0.322)			
DO	0.108 (0.323)	0.556* (0.000)		
BOD	0.228* (0.034)	0.203 (0.061)	0.323* (0.002)	
TP	0.216* (0.046)	0.160 (0.141)	0.085 (0.439)	0.051 (0.638)

*มีนัยสำคัญทางสถิติที่ระดับนัยสำคัญ 0.05

จากตารางที่ 2 พบว่า มีตัวแปร 4 คู่ ที่มีความสัมพันธ์เชิงเส้นอย่างมีนัยสำคัญ โดยมีตัวแปร 2 คู่ แสดงความสัมพันธ์เชิงเส้นระหว่างตัวแปรตามและตัวแปรอิสระ ได้แก่

DO และ BOD โดยให้ค่าสัมประสิทธิ์สหสัมพันธ์เท่ากับ 0.323 (ค่าพี เท่ากับ 0.002)

DO และ pH โดยให้ค่าสัมประสิทธิ์สหสัมพันธ์เท่ากับ 0.556 (ค่าพี เท่ากับ 0.000)

และอีก 2 คู่ แสดงความสัมพันธ์เชิงเส้นระหว่างตัวแปรอิสระด้วยกัน ได้แก่

TP และ WT โดยให้ค่าสัมประสิทธิ์สหสัมพันธ์เท่ากับ 0.216 (ค่าพี เท่ากับ 0.046)

BOD และ WT โดยให้ค่าสัมประสิทธิ์สหสัมพันธ์เท่ากับ 0.228 (ค่าพี เท่ากับ 0.034)

ตอนที่ 2 ผลการสร้างสมการถดถอยเพื่อพยากรณ์ปริมาณออกซิเจนละลายน้ำ

สำหรับการสร้างสมการถดถอยด้วยวิธีเซตย่อยที่ดีที่สุดจากโปรแกรมสำเร็จรูป Minitab 18 แสดงดังนี้ ตารางที่ 3 สมการถดถอยที่เป็นไปได้ทั้งหมดสำหรับพยากรณ์ DO

Variable	R ²	R _{adj} ²	Mellow's Cp	S	WT	pH	BOD	TP
1	30.9	30.1	4.8	1.5502		x		
1	10.4	9.4	30.6	1.7107			x	
2	35.6	34.0	1.0	1.4598		x	x	
2	31.2	29.5	6.5	1.5087	x	x		
3	35.6	33.2	3.0	1.4686		x	x	x
3	35.6	33.2	3.0	1.4687	x	x	x	
4	35.6	32.4	3.0	1.4776	x	x	x	x

จากตารางที่ 3 เมื่อพิจารณาจากค่าซีพีของมอลลิโรว์และค่าคลาดเคลื่อนมาตรฐานของการถดถอย (S) ที่มีค่าน้อย จะได้สมการถดถอยที่ประกอบไปด้วยตัวแปรอิสระ 2 ตัว ได้แก่ pH และ BOD ซึ่งมีค่าซีพีของมอลลิโรว์น้อยที่สุดเท่ากับ 1.0 และค่า S เท่ากับ 1.4598 และค่า R_{adj}^2 เท่ากับ 34.0 ดังนั้นสมการถดถอยเชิงเส้นพหุคูณเป็นดังนี้

$$\widehat{DO} = -10.02 + 1.960pH + 0.523BOD \quad (11)$$

จากนั้นทำการตรวจสอบความเหมาะสมของสมการที่ 11 ด้วยการวิเคราะห์ความแปรปรวน (Analysis of variance: ANOVA) ให้ผลดังนี้

ตารางที่ 4 การวิเคราะห์ความแปรปรวนของสมการถดถอยสมการที่ 11

Source	DF	SS	MS	F	P-value
Regression	2	97.64	48.819	22.91	0.000
Error	83	176.88	2.131		
Total	85	274.52			

จากตารางที่ 4 พิจารณาค่าสถิติทดสอบ F มีค่าเท่ากับ 22.91 เมื่อเทียบกับค่าวิกฤตมีค่าเท่ากับ 3.10 ซึ่งมีค่าที่เท่ากับ 0.000 พบว่ามีนัยสำคัญทางสถิติ สรุปได้ว่าสมการถดถอยที่เลือกมีความเหมาะสม

ตารางที่ 5 สัมประสิทธิ์การถดถอยและการทดสอบสมการถดถอยสมการที่ 11

Predictor	Coef	SE Coef	T	P-value
Constant	-10.02	2.42	-4.13	0.000
PH	1.960	0.345	5.96	0.000
BOD	0.523	0.214	2.44	0.017

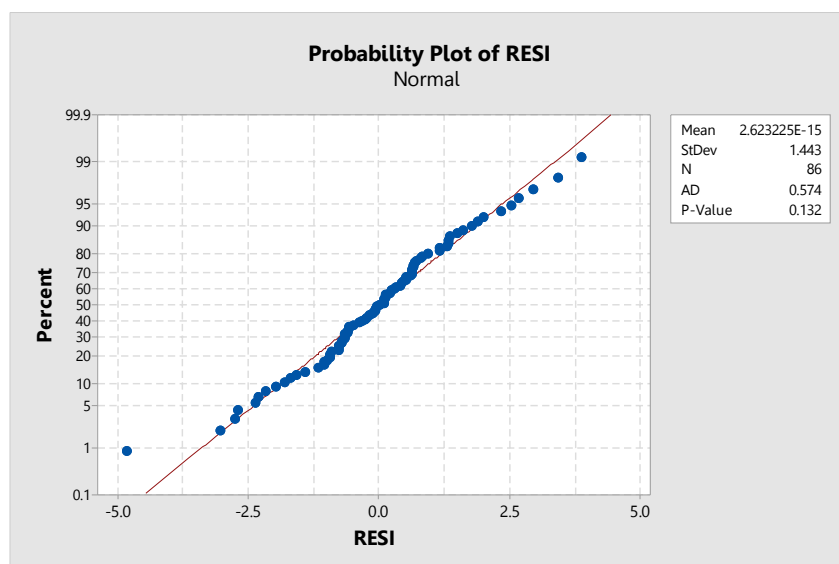
จากตารางที่ 5 พิจารณาค่าสถิติทดสอบ T สำหรับการทดสอบสัมประสิทธิ์การถดถอยแต่ละตัวดังนี้

- ค่าสถิติทดสอบ T สำหรับทดสอบตัวแปร PH มีค่าเท่ากับ 5.96 เมื่อเทียบกับค่าวิกฤตมีค่าเท่ากับ 1.988 หรือค่าพีมีค่าเท่ากับ 0.000 พบว่าตัวแปรอิสระ pH มีนัยสำคัญทางสถิติ
- ค่าสถิติทดสอบ T สำหรับทดสอบตัวแปร BOD มีค่าเท่ากับ 2.44 เมื่อเทียบกับค่าวิกฤตมีค่าเท่ากับ 1.988 หรือค่าพีมีค่าเท่ากับ 0.017 พบว่าตัวแปรอิสระ BOD มีนัยสำคัญทางสถิติ

ดังนั้นสรุปได้ว่า สมการถดถอยสมการที่ 11 มีความเหมาะสม จากนั้นทำการตรวจสอบข้อสมมุติของการวิเคราะห์การถดถอยดังนี้

1. การตรวจสอบการแจกแจงปกติของค่าคลาดเคลื่อน

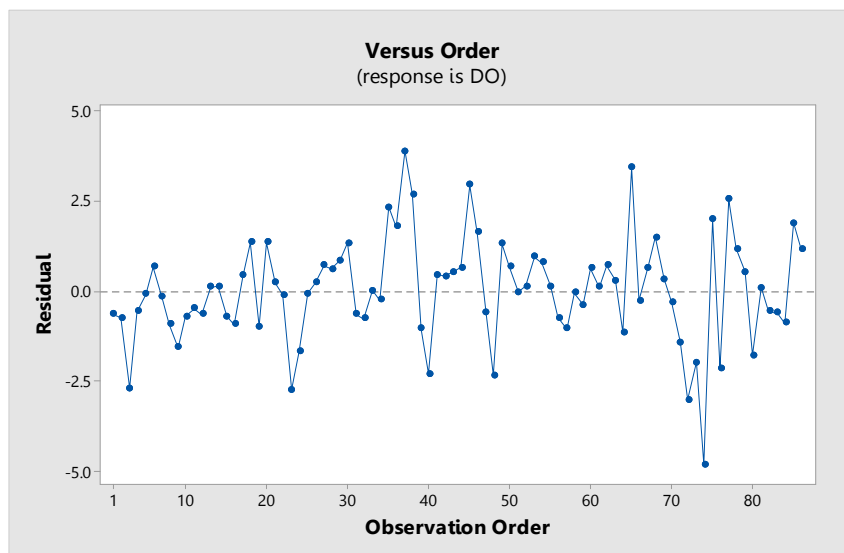
การตรวจสอบข้อสมมุติเกี่ยวกับการแจกแจงปกติของค่าคลาดเคลื่อนโดยใช้แผนภาพของส่วนเหลือ (Residuals) โดยเรียกแผนภาพนี้ว่า แผนภาพความน่าจะเป็นปกติ (Normal probability plot) ให้ผลดังภาพที่ 3



ภาพที่ 3 แผนภาพความน่าจะเป็นปกติของส่วนเหลือจากสมการถดถอยสมการที่ 11 เมื่อตรวจสอบการแจกแจงของค่าคลาดเคลื่อนของสมการถดถอยที่ 11 พบว่า AD มีค่าเท่ากับ 0.574 เมื่อเทียบกับค่าวิกฤตจากตารางที่ 1ก (ภาคผนวก) มีค่าเท่ากับ 0.752 หรือค่าพีเท่ากับ 0.132 พบว่า ไม่มีนัยสำคัญทางสถิติ สรุปได้ว่าค่าคลาดเคลื่อนมีการแจกแจงปกติ

2. การตรวจสอบความเป็นอิสระกันของค่าคลาดเคลื่อน

การตรวจสอบความเป็นอิสระกันของค่าคลาดเคลื่อนสามารถทำได้โดยใช้แผนภาพอนุกรมเวลาของส่วนเหลือให้ผลดังภาพที่ 4



ภาพที่ 4 แผนภาพอนุกรมเวลาของส่วนเหลือจากสมการถดถอยสมการที่ 11

พิจารณาจากการทดสอบเดอ์บิน-วัตสันดังสมการที่ 11 ได้ค่าเดอ์บิน-วัตสันหรือค่า d เท่ากับ 1.537 และเมื่อเทียบกับค่าวิกฤตขอบเขตล่าง (d_L) และค่าวิกฤตขอบเขตบน (d_U) มีค่าเท่ากับ 1.600 และ 1.696 ตามลำดับ จากตารางที่ 2ก (ภาคผนวก) พบว่า ค่า d มีค่าน้อยกว่าค่าวิกฤตขอบเขตล่าง ดังนั้นค่าคลาดเคลื่อนไม่เป็นอิสระกัน เมื่อพบว่าค่าคลาดเคลื่อนไม่เป็นอิสระกันจึงแก้ปัญหาดังกล่าวด้วยการตรวจสอบค่าสังเกตที่มีอิทธิพลต่อการวิเคราะห์การถดถอย โดยพิจารณาจากค่า DFFITS ดังสมการที่ 11 และคำนวณเกณฑ์การพิจารณาดังนี้

$$2\sqrt{\frac{k+1}{n}} = 2\sqrt{\frac{3}{86}} = 0.3735$$

หากค่า $|DFFITS_i|$ มีค่ามากกว่า 0.3735 จะสรุปได้ว่าค่าสังเกตที่ i เป็นค่าที่มีอิทธิพลต่อการวิเคราะห์การถดถอย

ค่า $|DFFITS_i|$ แสดงดังตารางที่ 6 เมื่อพิจารณาพบว่า มีค่าสังเกต 8 ค่าคือ ค่าสังเกตลำดับที่ 23 37 45 65 72 74 76 และ 77 มีค่า $|DFFITS_i|$ มากกว่า 0.3735 แสดงว่าค่าสังเกตเหล่านี้มีอิทธิพลต่อการวิเคราะห์การถดถอย ดังนั้นจึงทำการตัดค่าสังเกตดังกล่าวออก และทำการวิเคราะห์สมการถดถอยใหม่อีกครั้ง

ตารางที่ 6 ค่า $|DFFITS_i|$ ของสมการถดถอยสมการที่ 11

ลำดับ	$ DFFITS_i $	ลำดับ	$ DFFITS_i $	ลำดับ	$ DFFITS_i $	ลำดับ	$ DFFITS_i $
1	0.058855	23	0.374318	45	0.428228	67	0.055036
2	0.110257	24	0.141088	46	0.202686	68	0.111134
3	0.245616	25	0.006575	47	0.106131	69	0.025207
4	0.069514	26	0.019578	48	0.320107	70	0.041024
5	0.011888	27	0.065380	49	0.249361	71	0.149729
6	0.093239	28	0.061286	50	0.097034	72	0.606286
7	0.014897	29	0.117942	51	0.003028	73	0.317019
8	0.114593	30	0.126602	52	0.009113	74	2.355271
9	0.212918	31	0.055609	53	0.094392	75	0.173267
10	0.071308	32	0.107919	54	0.077510	76	0.382903
11	0.113141	33	0.000728	55	0.019821	77	0.421093
12	0.093545	34	0.035340	56	0.133524	78	0.145280
13	0.009256	35	0.355943	57	0.093414	79	0.056598
14	0.014581	36	0.149651	58	0.008791	80	0.348867
15	0.060163	37	1.267320	59	0.033321	81	0.010782
16	0.093868	38	0.301947	60	0.099534	82	0.060313
17	0.037915	39	0.123334	61	0.018168	83	0.061250
18	0.159176	40	0.229000	62	0.067739	84	0.096029
19	0.090082	41	0.046132	63	0.025874	85	0.20811
20	0.165137	42	0.056608	64	0.133470	86	0.156517
21	0.031867	43	0.062071	65	0.713597		
22	0.010253	44	0.068563	66	0.038233		

ผลการวิเคราะห์สมการถดถอยใหม่ได้ตั้งสมการที่ (12) โดยมีค่า S เท่ากับ 1.0965 และค่า R_{adj}^2 เท่ากับ 44.30

$$\widehat{DO} = -10.69 + 2.031pH + 0.628BOD \quad (12)$$

เมื่อได้สมการถดถอยเรียบร้อยแล้วจากนั้นทำการตรวจสอบความเหมาะสมของสมการถดถอยด้วยการวิเคราะห์ความแปรปรวนได้ผลการวิเคราะห์ดังตารางที่ 7

ตารางที่ 7 การวิเคราะห์ความแปรปรวนของสมการถดถอยสมการที่ 12

Source	DF	SS	MS	F	P-Value
Regression	2	76.02	38.012	31.61	0.000
Error	75	90.18	1.202		
Total	77	166.20			

จากตารางที่ 7 พิจารณาค่าสถิติทดสอบ F มีค่าเท่ากับ 31.61 เมื่อเทียบกับค่าวิกฤตมีค่าเท่ากับ 3.11 ซึ่งมีค่าที่เท่ากับ 0.000 พบว่ามีนัยสำคัญทางสถิติ สรุปได้ว่าสมการถดถอยที่เลือกมีความเหมาะสม

ตารางที่ 8 สัมประสิทธิ์การถดถอยและการทดสอบสมการถดถอยสมการที่ 12

Predictor	Coef	SE Coef	T	P-Value
Constant	-10.69	2.05	-5.21	0.000
pH	2.031	0.290	7.00	0.000
BOD	0.628	0.215	2.93	0.005

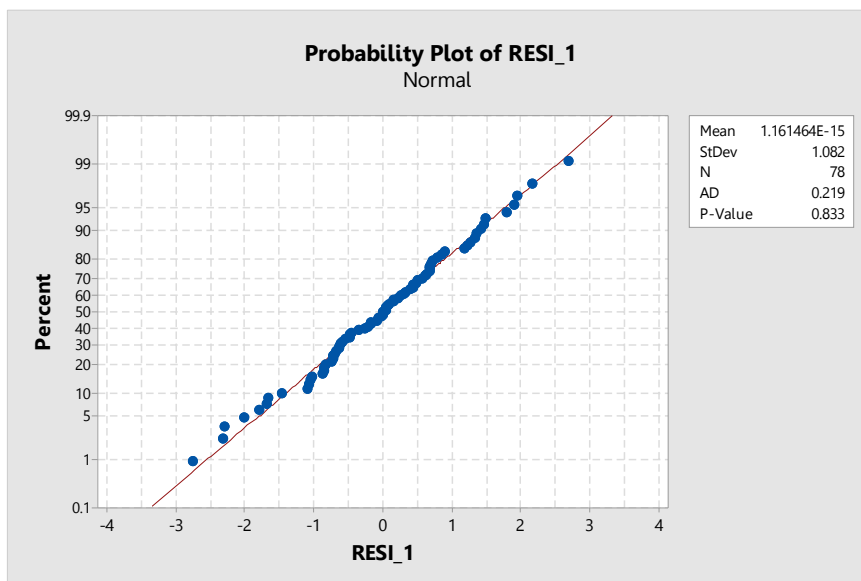
จากตารางที่ 8 พบค่าสถิติทดสอบ T สำหรับการทดสอบสัมประสิทธิ์การถดถอยแต่ละตัวดังนี้

- ค่าสถิติทดสอบ T สำหรับทดสอบตัวแปร PH มีค่าเท่ากับ 7.00 เมื่อเทียบกับค่าวิกฤต มีค่าเท่ากับ 1.992 หรือค่าที่มีค่าเท่ากับ 0.000 พบว่าตัวแปรอิสระ pH มีนัยสำคัญทางสถิติ
- ค่าสถิติทดสอบ T สำหรับทดสอบตัวแปร BOD มีค่าเท่ากับ 2.93 เมื่อเทียบกับค่าวิกฤต มีค่าเท่ากับ 1.992 หรือค่าที่มีค่าเท่ากับ 0.005 พบว่าตัวแปรอิสระ BOD มีนัยสำคัญทางสถิติ

ดังนั้นสรุปได้ว่า สมการถดถอยสมการที่ 12 มีความเหมาะสม จากนั้นทำการตรวจสอบข้อสมมุติของการวิเคราะห์การถดถอยดังนี้

1. การตรวจสอบการแจกแจงปกติของค่าคลาดเคลื่อน

การตรวจสอบข้อสมมุติเกี่ยวกับการแจกแจงปกติของค่าคลาดเคลื่อนโดยใช้แผนภาพความน่าจะเป็นปกติของส่วนเหลือให้ผลดังภาพที่ 5

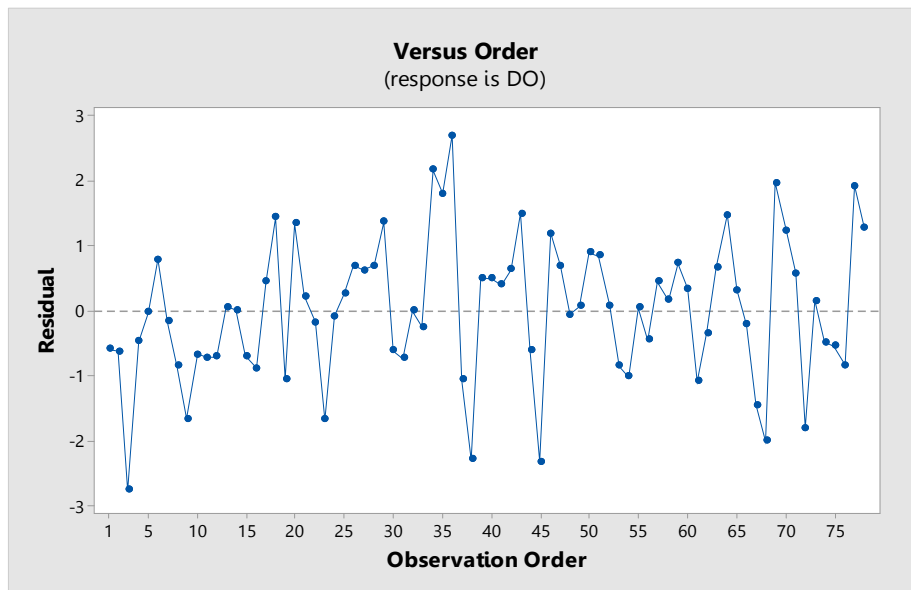


ภาพที่ 5 แผนภาพความน่าจะเป็นปกติของส่วนเหลือจากสมการถดถอยสมการที่ 12

เมื่อตรวจสอบการแจกแจงของค่าคลาดเคลื่อนของสมการถดถอยที่ 12 พบว่า AD มีค่าเท่ากับ 0.219 เมื่อเทียบกับค่าวิกฤตจากตารางที่ 1ก (ภาคผนวก) มีค่าเท่ากับ 0.752 หรือค่าพีเท่ากับ 0.833 พบว่า ไม่มีนัยสำคัญทางสถิติ สรุปได้ว่าค่าคลาดเคลื่อนมีการแจกแจงปกติ

2. การตรวจสอบความเป็นอิสระกันของค่าคลาดเคลื่อน

การตรวจสอบความเป็นอิสระกันของค่าคลาดเคลื่อนสามารถทำได้โดยใช้แผนภาพอนุกรมเวลาของส่วนเหลือ ให้ผลดังภาพที่ 6 จะเห็นได้ว่าค่าคลาดเคลื่อนมีลักษณะกระจายอย่างสุ่มรอบศูนย์ แต่เมื่อทำการพิจารณาจากการทดสอบเดอ์บิน-วัตสันตั้งสมการที่ 12 ได้ค่าเดอ์บิน-วัตสันหรือค่า d เท่ากับ 1.545 และเมื่อเทียบกับค่าวิกฤตขอบเขตล่าง (d_L) และค่าวิกฤตขอบเขตบน (d_U) มีค่าเท่ากับ 1.586 และ 1.688 ตามลำดับจากตารางที่ 2ก (ภาคผนวก) พบว่า ค่า d มีค่าน้อยกว่าค่าวิกฤตขอบเขตล่าง ดังนั้นค่าคลาดเคลื่อนเป็นไม่เป็นอิสระกัน



ภาพที่ 6 แผนภาพอนุกรมเวลาของส่วนเหลือจากสมการถดถอยสมการที่ 12

เมื่อพบว่าค่าคลาดเคลื่อนไม่เป็นอิสระกันจึงแก้ปัญหาดังกล่าวด้วยการตรวจสอบค่าสังเกตที่มีอิทธิพลต่อการวิเคราะห์การถดถอย โดยพิจารณาจากค่า DFFITS ดังสมการที่ 12 และคำนวณเกณฑ์การพิจารณาดังนี้

$$2\sqrt{\frac{k+1}{n}} = 2\sqrt{\frac{3}{78}} = 0.3922$$

หากค่า $|DFFITS_i|$ มีค่ามากกว่า 0.3922 จะสรุปได้ว่าค่าสังเกตที่ i เป็นค่าที่มีอิทธิพลต่อการวิเคราะห์การถดถอย ค่า $|DFFITS_i|$ แสดงดังตารางที่ 9 เมื่อพิจารณาพบว่า มีค่าสังเกต 6 ค่าคือ ค่าสังเกตลำดับที่ 3 34 36 45 68 และ 72 มีค่า $|DFFITS_i|$ มากกว่า 0.3922 แสดงว่าค่าสังเกตเหล่านี้มีอิทธิพลต่อการวิเคราะห์การถดถอย ดังนั้นจึงทำการตัดค่าสังเกตดังกล่าวออก และทำการวิเคราะห์สมการถดถอยใหม่อีกครั้ง

ผลการวิเคราะห์สมการถดถอยใหม่ได้ดังสมการที่ (13) โดยมีค่า S เท่ากับ 0.9045 และค่า R_{adj}^2 เท่ากับ 55.26

$$\widehat{DO} = -12.22 + 2.294pH + 0.457BOD \quad (13)$$

เมื่อได้สมการถดถอยเรียบร้อยแล้วจากนั้นทำการตรวจสอบความเหมาะสมของสมการถดถอยด้วยการวิเคราะห์ความแปรปรวนได้ผลการวิเคราะห์ดังตารางที่ 10

ตารางที่ 9 ค่า $|DFFITS_i|$ ของสมการถดถอยสมการที่ 12

ลำดับ	$ DFFITS_i $	ลำดับ	$ DFFITS_i $	ลำดับ	$ DFFITS_i $	ลำดับ	$ DFFITS_i $
1	0.077301	21	0.048943	41	0.084607	61	0.175784
2	0.144387	22	0.026122	42	0.101045	62	0.079988
3	0.410316	23	0.192521	43	0.334559	63	0.078276
4	0.091414	24	0.010978	44	0.164674	64	0.159624
5	0.005624	25	0.029902	45	0.449963	65	0.034949
6	0.175500	26	0.095350	46	0.385949	66	0.044995
7	0.018518	27	0.094397	47	0.141248	67	0.235014
8	0.148582	28	0.182042	48	0.008086	68	0.497045
9	0.388188	29	0.192442	49	0.008293	69	0.255595
10	0.091492	30	0.072594	50	0.139790	70	0.223520
11	0.320491	31	0.146268	51	0.117288	71	0.093996
12	0.155158	32	0.000548	52	0.018981	72	0.546043
13	0.008125	33	0.053907	53	0.246437	73	0.025161
14	0.000378	34	0.628421	54	0.122741	74	0.075764
15	0.078305	35	0.216228	55	0.007070	75	0.079488
16	0.122167	36	0.432215	56	0.064373	76	0.124286
17	0.054319	37	0.181900	57	0.136773	77	0.292273
18	0.244074	38	0.311170	58	0.040645	78	0.276181
19	0.156560	39	0.069553	59	0.103923		
20	0.249274	40	0.098979	60	0.043071		

ตารางที่ 10 การวิเคราะห์ความแปรปรวนของสมการถดถอยสมการที่ 13

Source	DF	SS	MS	F	P-Value
Regression	2	76.02	36.6914	44.84	0.000
Error	69	90.18	0.8182		
Total	71	166.20			

จากตารางที่ 10 พิจารณาค่าสถิติทดสอบ F มีค่าเท่ากับ 44.84 เมื่อเทียบกับค่าวิกฤตมีค่าเท่ากับ 3.12 ซึ่งมีค่าพีเท่ากับ 0.000 พบว่ามีนัยสำคัญทางสถิติ สรุปได้ว่าสมการถดถอยที่เลือกมีความเหมาะสม

ตารางที่ 11 สัมประสิทธิ์การถดถอยและการทดสอบสมการถดถอยสมการที่ 13

Predictor	Coef	SE Coef	T	P-Value
Constant	-12.22	1.85	-6.61	0.000
pH	2.294	0.263	8.71	0.000
BOD	0.457	0.185	2.47	0.016

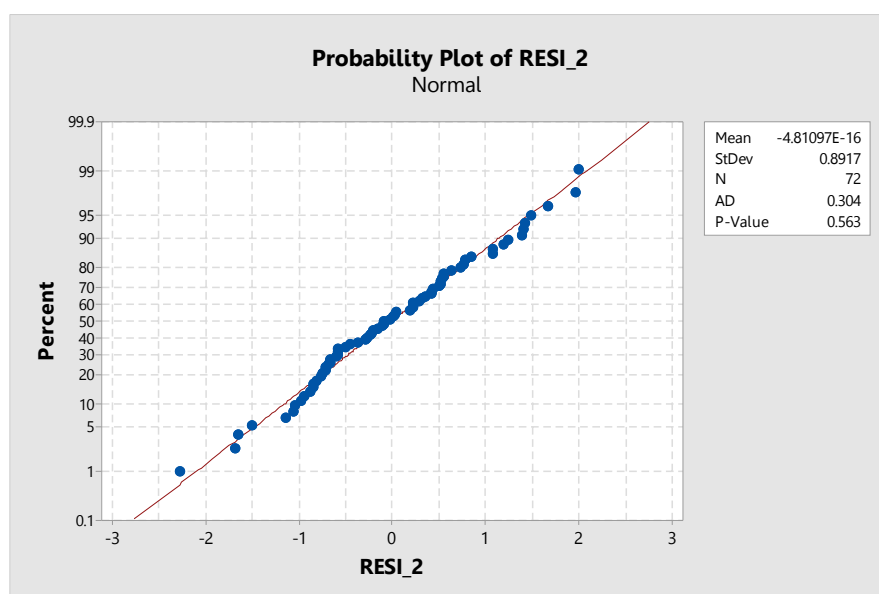
จากตารางที่ 11 พบค่าสถิติทดสอบ T สำหรับการทดสอบสัมประสิทธิ์การถดถอยแต่ละตัวดังนี้

- ค่าสถิติทดสอบ T สำหรับทดสอบตัวแปร PH มีค่าเท่ากับ 8.71 เมื่อเทียบกับค่าวิกฤต มีค่าเท่ากับ 1.994 หรือค่าที่มีค่าเท่ากับ 0.000 พบว่าตัวแปรอิสระ pH มีนัยสำคัญทางสถิติ
- ค่าสถิติทดสอบ T สำหรับทดสอบตัวแปร BOD มีค่าเท่ากับ 2.47 เมื่อเทียบกับค่าวิกฤต มีค่าเท่ากับ 1.994 หรือค่าที่มีค่าเท่ากับ 0.016 พบว่าตัวแปรอิสระ BOD มีนัยสำคัญทางสถิติ

ดังนั้นสรุปได้ว่า สมการถดถอยสมการที่ 13 มีความเหมาะสม จากนั้นทำการตรวจสอบข้อสมมุติของการวิเคราะห์การถดถอยดังนี้

1. การตรวจสอบการแจกแจงปกติของค่าคลาดเคลื่อน

การตรวจสอบข้อสมมุติเกี่ยวกับการแจกแจงปกติของค่าคลาดเคลื่อนโดยใช้แผนภาพความน่าจะเป็นปกติของส่วนเหลือให้ผลดังภาพที่ 7

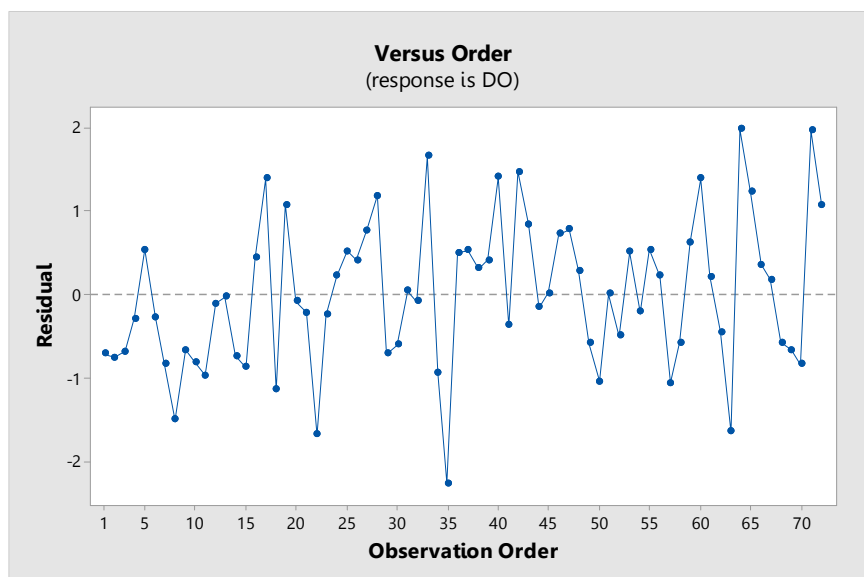


ภาพที่ 7 แผนภาพความน่าจะเป็นปกติของส่วนเหลือจากสมการถดถอยสมการที่ 13

เมื่อตรวจสอบการแจกแจงของค่าคลาดเคลื่อนของสมการถดถอยที่ 13 พบว่า AD มีค่าเท่ากับ 0.304 เมื่อเทียบกับค่าวิกฤตจากตารางที่ 1ก (ภาคผนวก) มีค่าเท่ากับ 0.752 หรือค่าพีเท่ากับ 0.563 พบว่า ไม่มีนัยสำคัญทางสถิติ สรุปได้ว่าค่าคลาดเคลื่อนมีการแจกแจงปกติ

2. การตรวจสอบความเป็นอิสระกันของค่าคลาดเคลื่อน

การตรวจสอบความเป็นอิสระกันของค่าคลาดเคลื่อนสามารถทำได้โดยใช้แผนภาพอนุกรมเวลาของส่วนเหลือ ให้ผลดังภาพที่ 8



ภาพที่ 8 แผนภาพอนุกรมเวลาของส่วนเหลือจากสมการถดถอยสมการที่ 13

จากภาพที่ 8 จะเห็นได้ว่าค่าคลาดเคลื่อนมีลักษณะกระจายอย่างสุ่มรอบศูนย์ และเมื่อทำการพิจารณาจากการทดสอบเดอ์บิน-วัตสันตั้งสมการที่ 13 ได้ค่าเดอ์บิน-วัตสันหรือค่า d เท่ากับ 1.653 แต่เมื่อเทียบกับค่าวิกฤตขอบเขตล่าง (d_L) และค่าวิกฤตขอบเขตบน (d_U) มีค่าเท่ากับ 1.554 และ 1.672 ตามลำดับ จากตารางที่ 2ก (ภาคผนวก) พบว่า ค่า d มีอยู่ระหว่างค่าวิกฤตขอบเขตบนและขอบเขตล่าง ดังนั้นค่าคลาดเคลื่อนเป็นไม่เป็นอิสระกัน

เมื่อพบว่าค่าคลาดเคลื่อนไม่เป็นอิสระกันจึงแก้ปัญหาดังกล่าวด้วยการตรวจสอบค่าสังเกตที่มีอิทธิพลต่อการวิเคราะห์การถดถอย โดยพิจารณาจากค่า DFFITS ตั้งสมการที่ 13 และคำนวณเกณฑ์การพิจารณาดังนี้

$$2\sqrt{\frac{k+1}{n}} = 2\sqrt{\frac{3}{72}} = 0.4082$$

ตารางที่ 12 ค่า $|DFFITS_i|$ ของสมการถดถอยสมการที่ 13

ลำดับ	$ DFFITS_i $	ลำดับ	$ DFFITS_i $	ลำดับ	$ DFFITS_i $	ลำดับ	$ DFFITS_i $
1	0.114709	21	0.039388	41	0.129667	61	0.031046
2	0.217343	22	0.244425	42	0.634693	62	0.120234
3	0.164606	23	0.042633	43	0.226702	63	0.362062
4	0.072114	24	0.032573	44	0.023394	64	0.334663
5	0.153963	25	0.095840	45	0.002939	65	0.282246
6	0.038702	26	0.082539	46	0.153467	66	0.075827
7	0.183806	27	0.263824	47	0.135079	67	0.038023
8	0.451536	28	0.215075	48	0.086333	68	0.110265
9	0.113741	29	0.102968	49	0.221673	69	0.120505
10	0.479186	30	0.154801	50	0.156052	70	0.155443
11	0.291641	31	0.007179	51	0.001004	71	0.387905
12	0.022081	32	0.022448	52	0.092114	72	0.293211
13	0.008210	33	0.258024	53	0.204322		
14	0.104024	34	0.208192	54	0.072668		
15	0.151807	35	0.394413	55	0.100137		
16	0.066854	36	0.087515	56	0.036042		
17	0.299250	37	0.138536	57	0.216087		
18	0.226984	38	0.089208	58	0.178468		
19	0.271803	39	0.089714	59	0.091261		
20	0.024576	40	0.421246	60	0.196766		

หากค่า $|DFFITS_i|$ มีค่ามากกว่า 0.4082 จะสรุปได้ว่าค่าสังเกตที่ i เป็นค่าที่มีอิทธิพลต่อการวิเคราะห์การถดถอย ค่า $|DFFITS_i|$ แสดงดังตารางที่ 12 เมื่อพิจารณาพบว่า มีค่าสังเกต 4 ค่าคือ ค่าสังเกตลำดับที่ 8 10 40 และ 42 มีค่า $|DFFITS_i|$ มากกว่า 0.4082 แสดงว่าค่าสังเกตเหล่านี้มีอิทธิพลต่อการวิเคราะห์การถดถอย ดังนั้นจึงทำการตัดค่าสังเกตดังกล่าวออก และทำการวิเคราะห์สมการถดถอยใหม่อีกครั้ง

ผลการวิเคราะห์สมการถดถอยใหม่ได้ตั้งสมการที่ (14) โดยมีค่า S เท่ากับ 0.8710 และค่า R_{adj}^2 เท่ากับ 54.15

$$\widehat{DO} = -12.57 + 2.347\text{pH} + 0.437\text{BOD} \quad (14)$$

เมื่อได้สมการถดถอยเรียบร้อยแล้วจากนั้นทำการตรวจสอบความเหมาะสมของสมการถดถอยด้วยการวิเคราะห์ความแปรปรวนได้ผลการวิเคราะห์ดังตารางที่ 13

ตารางที่ 13 การวิเคราะห์ความแปรปรวนของสมการถดถอยสมการที่ 14

Source	DF	SS	MS	F	P-Value
Regression	2	61.540	30.7701	40.56	0.000
Error	65	49.312	0.7586		
Total	67	110.852			

จากตารางที่ 13 พิจารณาค่าสถิติทดสอบ F มีค่าเท่ากับ 40.56 เมื่อเทียบกับค่าวิกฤตมีค่าเท่ากับ 3.13 ซึ่งมีค่า P เท่ากับ 0.000 พบว่ามีนัยสำคัญทางสถิติ สรุปได้ว่าสมการถดถอยที่เลือกมีความเหมาะสม

ตารางที่ 14 สัมประสิทธิ์การถดถอยและการทดสอบสมการถดถอยสมการที่ 14

Predictor	Coef	SE Coef	T	P-Value
Constant	-12.57	1.94	-6.49	0.000
pH	2.347	0.273	8.58	0.000
BOD	0.437	0.204	2.15	0.036

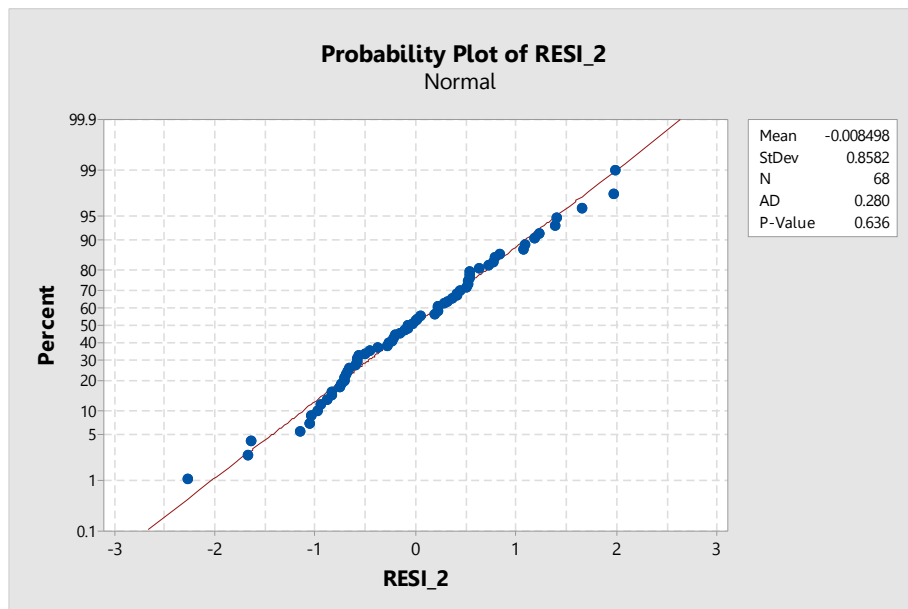
จากตารางที่ 14 พบค่าสถิติทดสอบ T สำหรับการทดสอบสัมประสิทธิ์การถดถอยแต่ละตัวดังนี้

- ค่าสถิติทดสอบ T สำหรับทดสอบตัวแปร PH มีค่าเท่ากับ 8.58 เมื่อเทียบกับค่าวิกฤตมีค่าเท่ากับ 1.997 หรือค่า P มีค่าเท่ากับ 0.000 พบว่าตัวแปรอิสระ pH มีนัยสำคัญทางสถิติ
- ค่าสถิติทดสอบ T สำหรับทดสอบตัวแปร BOD มีค่าเท่ากับ 2.15 เมื่อเทียบกับค่าวิกฤตมีค่าเท่ากับ 1.997 หรือค่า P มีค่าเท่ากับ 0.036 พบว่าตัวแปรอิสระ BOD มีนัยสำคัญทางสถิติ

ดังนั้นสรุปได้ว่า สมการถดถอยสมการที่ 14 มีความเหมาะสม จากนั้นทำการตรวจสอบข้อสมมุติของการวิเคราะห์การถดถอยดังนี้

1. การตรวจสอบการแจกแจงปกติของค่าคลาดเคลื่อน

การตรวจสอบข้อสมมติเกี่ยวกับการแจกแจงปกติของค่าคลาดเคลื่อนโดยใช้แผนภาพความน่าจะเป็นปกติของส่วนเหลือให้ผลดังภาพที่ 9

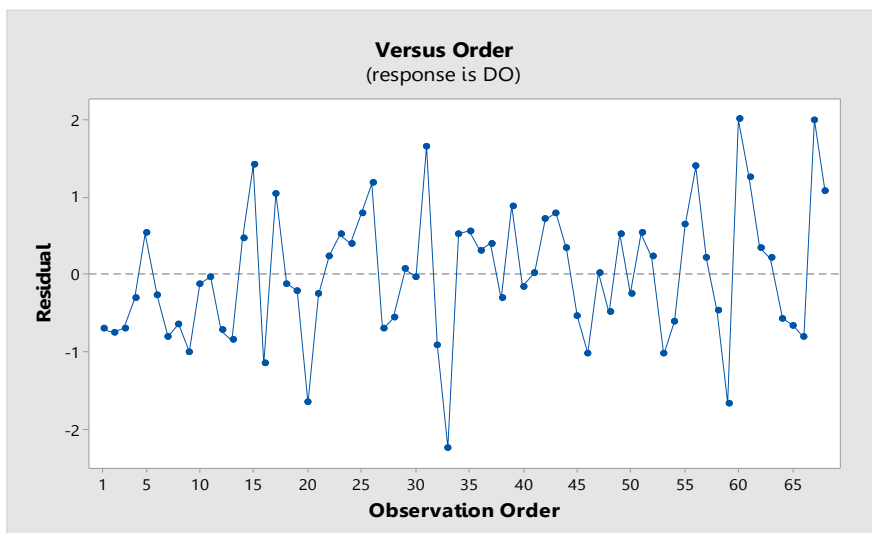


ภาพที่ 9 แผนภาพความน่าจะเป็นปกติของส่วนเหลือสมการถดถอยสมการที่ 14

เมื่อตรวจสอบการแจกแจงของค่าคลาดเคลื่อนของสมการถดถอยที่ 14 พบว่า AD มีค่าเท่ากับ 0.280 เมื่อเทียบกับค่าวิกฤตจากตารางที่ 1ก (ภาคผนวก) มีค่าเท่ากับ 0.752 หรือค่าพีเท่ากับ 0.636 พบว่า ไม่มีนัยสำคัญทางสถิติ สรุปได้ว่าค่าคลาดเคลื่อนมีการแจกแจงปกติ

2. การตรวจสอบความเป็นอิสระกันของค่าคลาดเคลื่อน

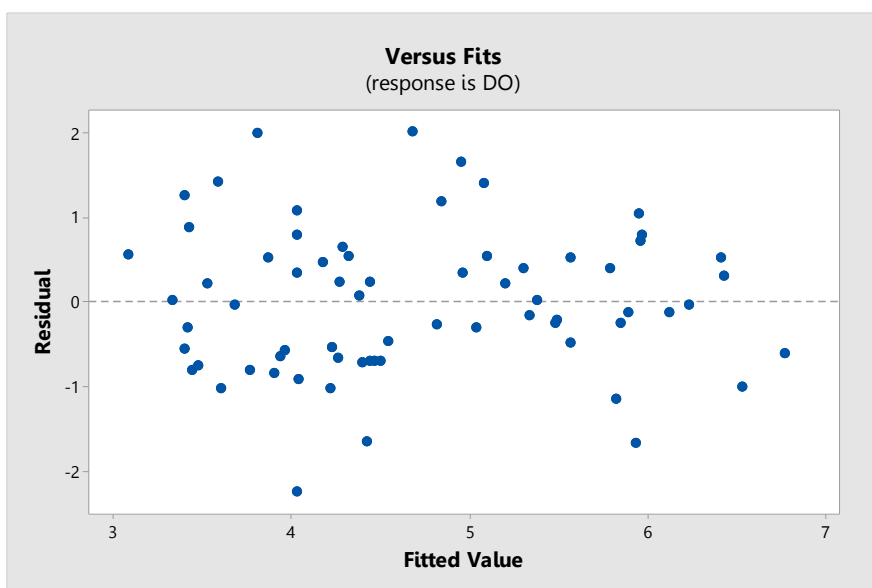
การตรวจสอบความเป็นอิสระกันของค่าคลาดเคลื่อนสามารถทำได้โดยใช้แผนภาพอนุกรมเวลาของส่วนเหลือ ให้ผลดังภาพที่ 10 จากภาพที่ 10 จะเห็นได้ว่าค่าคลาดเคลื่อนมีลักษณะกระจายอย่างสุ่มรอบศูนย์ และเมื่อทำการพิจารณาจากการทดสอบเดอร์บิน-วัตสันดังสมการที่ 14 ได้ค่าเดอร์บิน-วัตสันหรือค่า d เท่ากับ 1.752 แต่เมื่อเทียบกับค่าวิกฤตขอบเขตล่าง (d_L) และค่าวิกฤตขอบเขตบน (d_U) มีค่าเท่ากับ 1.554 และ 1.672 ตามลำดับ จากตารางที่ 2ก (ภาคผนวก) พบว่า ค่า d มีค่ามากกว่าค่าวิกฤตขอบเขตบน ดังนั้นค่าคลาดเคลื่อนเป็นอิสระกัน



ภาพที่ 10 แผนภาพอนุกรมเวลาของส่วนเหลือจากสมการถดถอยสมการที่ 14

3. การตรวจสอบความแปรปรวนคงที่ของค่าคลาดเคลื่อน

การตรวจสอบค่าคลาดเคลื่อนมีความแปรปรวนคงที่สามารถทำได้โดยใช้แผนภาพการกระจายระหว่างส่วนเหลือและค่าประมาณของตัวแปรตามให้ผลดังภาพที่ 11



ภาพที่ 11 แผนภาพการกระจายระหว่างส่วนเหลือและค่าประมาณจากสมการถดถอยสมการที่ 14

จากภาพที่ 11 จะเห็นได้ว่า ค่าคลาดเคลื่อนมีลักษณะกระจายอย่างสุ่มรอบศูนย์ เมื่อตรวจสอบโดยใช้การทดสอบบูซ-เพเกินจากสมการที่ 14 โดยค่า SSR* ดูได้จากตารางที่ 15 และค่า SSE ดูได้จากตารางที่ 9 จากนั้นคำนวณค่าสถิติทดสอบบูซ-เพเกินหรือค่า BP มีค่าเท่ากับ 1.0855 เมื่อเปรียบเทียบกับค่าวิกฤตจากตารางที่ 16 มีค่าเท่ากับ 3.841 หรือค่าพีเท่ากับ 0.297 พบว่า มีนัยสำคัญทางสถิติ ดังนั้นค่าคลาดเคลื่อนมีความแปรปรวนคงที่

ตารางที่ 15 การวิเคราะห์ความแปรปรวนของสมการถดถอยเมื่อค่าคลาดเคลื่อนกำลังสองเป็นตัวแปรตาม

Source	DF	SS	MS	F	P-Value
Regression	2	1.1417	0.57086	0.53	0.593
Error	65	70.3865	1.08287		
Total	67	71.5283			

ตารางที่ 16 ค่าวิกฤตไคกำลังสองด้วยองศาเสรีเท่ากับ 1

P(X<=x)	x
0.95	3.841

4. การตรวจสอบความเป็นอิสระกันของตัวแปรอิสระ

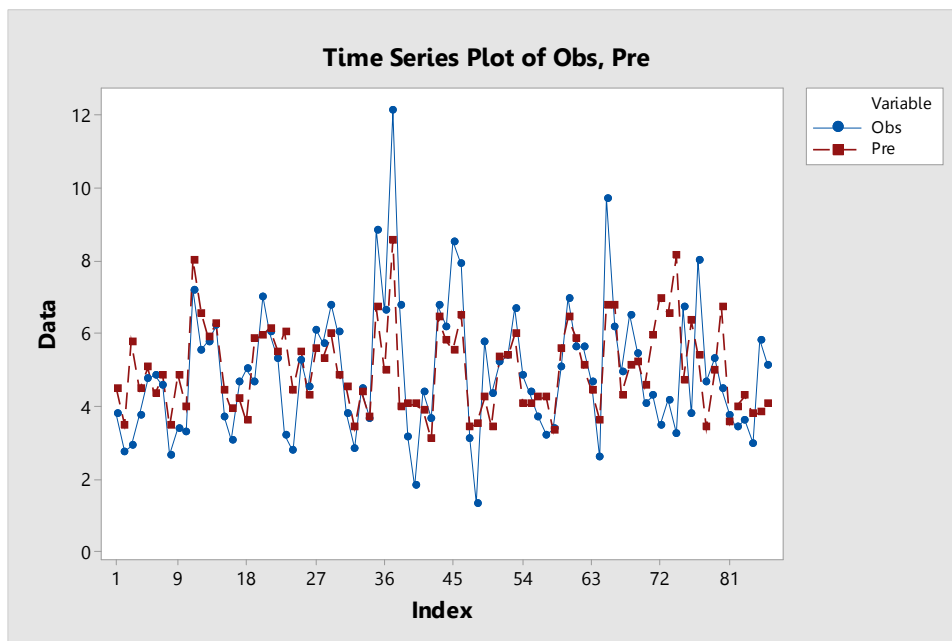
พิจารณาจากค่า VIF จากสมการถดถอยสมการที่ 14 จากตารางที่ 17 พบว่าค่า VIF มีค่าเท่ากับ 1.00 ทั้งสองตัวแปร ซึ่งมีค่าไม่เกิน 5 สรุปได้ว่าจึงเกิดพหุสัมพันธ์ระหว่างตัวแปรอิสระน้อย

ตารางที่ 17 ค่า VIF ของสมการถดถอยสมการที่ 14

Predictor	VIF
pH	1.00
BOD	1.00

ตอนที่ 3 ผลการตรวจสอบประสิทธิภาพของสมการถดถอย

เมื่อนำค่าพยากรณ์ DO ที่ได้จากสมการถดถอยสมการที่ 14 และค่าจริง DO ตั้งแต่ปีพ.ศ. 2551 ถึงปี พ.ศ. 2561 มาสร้างแผนภาพอนุกรมเวลาเพื่อทำการเปรียบเทียบแสดงดังภาพที่ 12



ภาพที่ 12 แผนภาพอนุกรมเวลาระหว่างค่าพยากรณ์และค่าจริงของปริมาณออกซิเจนละลายน้ำ

จากภาพที่ 12 พบว่าแผนภาพทั้งสองมีลักษณะใกล้เคียงกัน เมื่อตรวจสอบประสิทธิภาพด้วยค่าคลาดเคลื่อนสัมบูรณ์เฉลี่ยจากสมการที่ 9 ให้ผลดังตารางที่ 18 และค่าคลาดเคลื่อนสัมบูรณ์เฉลี่ยคำนวณได้ดังนี้

$$\begin{aligned}
 MAE &= \frac{\sum_{i=1}^n |O_i - P_i|}{n} \\
 &= \frac{|3.7567 - 4.4683| + |2.7200 - 3.4741| + |2.9000 - 5.7678| + \dots + |5.1000 - 4.0257|}{86} \\
 &= \frac{92.3829}{86} \\
 &= 1.0742
 \end{aligned}$$

ตารางที่ 18 การตรวจสอบประสิทธิภาพของสมการถดถอย

ลำดับที่ i	O_i	P_i	ลำดับที่ i	O_i	P_i	ลำดับที่ i	O_i	P_i	ลำดับที่ i	O_i	P_i
1	3.7567	4.4683	23	3.2000	6.0214	45	8.5000	5.5380	67	4.9267	4.2887
2	2.7200	3.4741	24	2.7625	4.4258	46	7.8900	6.4838	68	6.4800	5.0814
3	2.9000	5.7678	25	5.2267	5.5632	47	3.1000	3.4157	69	5.4100	5.1975
4	3.7450	4.4433	26	4.5050	4.2680	48	1.3200	3.4838	70	4.0700	4.5042
5	4.7267	5.0357	27	6.0733	5.5632	49	5.7600	4.2235	71	4.2667	5.9340
6	4.8500	4.3206	28	5.7000	5.3022	50	4.3100	3.4232	72	3.4550	6.9400
7	4.5400	4.8139	29	6.7567	5.9697	51	5.1800	5.3355	73	4.1467	6.5148
8	2.6300	3.4413	30	6.0200	4.8396	52	5.3950	5.3740	74	3.2300	8.1194
9	3.3700	4.8365	31	3.7933	4.4982	53	6.6733	4.9610	75	6.6867	4.6782
10	3.2900	3.9394	32	2.8300	3.3943	54	4.8250	4.0297	76	3.7650	6.3261
11	7.1767	8.0122	33	4.4500	4.3836	55	4.3633	4.0299	77	8.0000	5.3868
12	5.5200	6.5296	34	3.6400	3.6771	56	3.6900	4.2259	78	4.6550	3.3976
13	5.7667	5.8916	35	8.8400	6.6938	57	3.1933	4.2190	79	5.3033	4.9599
14	6.2000	6.2306	36	6.6050	4.9475	58	3.3550	3.3307	80	4.4700	6.7219
15	3.6667	4.3971	37	12.1133	8.5629	59	5.0700	5.5668	81	3.7333	3.5241
16	3.0500	3.9042	38	6.7500	3.9464	60	6.9400	6.4122	82	3.3950	3.9641
17	4.6333	4.1759	39	3.1233	4.0380	61	5.6000	5.8515	83	3.6000	4.2614
18	5.0000	3.5870	40	1.7800	4.0271	62	5.6275	5.0984	84	2.9500	3.7667
19	4.6333	5.8204	41	4.3867	3.8667	63	4.6633	4.4370	85	5.8000	3.8050
20	7.0000	5.9510	42	3.6540	3.0813	64	2.5650	3.6007	86	5.1000	4.0257
21	6.0000	6.1201	43	6.7333	6.4246	65	9.7000	6.7592			
22	5.2750	5.4858	44	6.1800	5.7856	66	6.1600	6.7663			

เมื่อตรวจสอบประสิทธิภาพด้วยรากของค่าคลาดเคลื่อนกำลังสองเฉลี่ยโดยใช้ข้อมูลจากตารางที่ 18 และสมการที่ 14 คำนวณได้ดังนี้

$$\begin{aligned}
 \text{RMSE} &= \sqrt{\frac{\sum_{i=1}^n (O_i - P_i)^2}{n}} \\
 &= \sqrt{\frac{(3.7567 - 4.4683)^2 + (2.7000 - 3.4741)^2 + (2.9000 - 5.7678)^2 + \dots + (5.1000 - 4.2057)^2}{86}} \\
 &= \sqrt{\frac{180.0484}{86}} \\
 &= 1.4469
 \end{aligned}$$

บทที่ 5

สรุปและอภิปรายผลการศึกษา

5.1 สรุปผลการศึกษา

จากการสร้างสมการถดถอยเพื่อพยากรณ์ปริมาณออกซิเจนละลายน้ำของแม่น้ำบางปะกงได้สมการถดถอยเชิงเส้นพหุคูณดังนี้

$$\widehat{DO} = -10.02 + 1.960pH + 0.523BOD$$

เมื่อทำการตรวจสอบข้อสมมุติของการวิเคราะห์การถดถอยของสมการถดถอยดังกล่าวพบว่าค่าคลาดเคลื่อนไม่เป็นอิสระกันจึงทำการตรวจสอบค่าสังเกตที่มีอิทธิพลต่อสมการถดถอยโดยพิจารณาจากค่า DFFITS พบว่ามีค่าสังเกต 8 ค่าที่เป็นค่าสังเกตที่มีอิทธิพล จึงทำการตัดค่าสังเกตเหล่านี้ออกและทำการสร้างสมการถดถอยอีกครั้งได้สมการถดถอยเชิงเส้นพหุคูณดังนี้

$$\widehat{DO} = -10.69 + 2.031pH + 0.628BOD$$

หลังจากตัดค่าสังเกตที่มีอิทธิพลออกแล้วทำการตรวจสอบข้อสมมุติของการวิเคราะห์การถดถอยของสมการถดถอยดังกล่าว พบว่าค่าคลาดเคลื่อนไม่เป็นอิสระกันจึงทำการตรวจสอบค่าสังเกตที่มีอิทธิพลต่อสมการถดถอยโดยพิจารณาจาก ค่า DFFITS พบว่ามีค่าสังเกต 6 ค่าที่เป็นค่าสังเกตที่มีอิทธิพล จึงทำการตัดค่าสังเกตเหล่านี้ออกจากและทำการสร้างสมการถดถอยอีกครั้งได้สมการถดถอยเชิงเส้นพหุคูณดังนี้

$$\widehat{DO} = -12.22 + 2.294pH + 0.457BOD$$

หลังจากตัดค่าสังเกตที่มีอิทธิพลออกแล้วทำการตรวจสอบข้อสมมุติของการวิเคราะห์การถดถอยของสมการถดถอยดังกล่าว พบว่าค่าคลาดเคลื่อนไม่เป็นอิสระกันจึงทำการตรวจสอบค่าสังเกตที่มีอิทธิพลต่อสมการถดถอยโดยพิจารณาจาก ค่า DFFITS พบว่ามีค่าสังเกต 4 ค่าที่เป็นค่าสังเกตที่มีอิทธิพล จึงทำการตัดค่าสังเกตเหล่านี้ออกจากและทำการสร้างสมการถดถอยอีกครั้งได้สมการถดถอยเชิงเส้นพหุคูณดังนี้

$$\widehat{DO} = -12.57 + 2.347pH + 0.437BOD$$

สมการถดถอยดังกล่าวมีค่าคลาดเคลื่อนของการถดถอยเท่ากับ 0.8710 และมีค่าสัมประสิทธิ์การตัดสินใจที่ปรับค่าแล้วเท่ากับ 54.15 จากนั้นทำการตรวจสอบข้อสมมุติของการวิเคราะห์การถดถอยของสมการถดถอยพบว่า ค่าคลาดเคลื่อนเป็นไปตามข้อสมมุติของการวิเคราะห์การถดถอยทุกข้อนั่นคือ 1) ค่าคลาดเคลื่อนมีการแจกแจงปกติ 2) ค่าคลาดเคลื่อนเป็นอิสระกัน 3) ความแปรปรวนของค่าคลาดเคลื่อนมีค่าคงที่ และ 4) ความเป็นอิสระกันของตัวแปรอิสระ

เมื่อตรวจสอบประสิทธิภาพของการพยากรณ์ด้วยแผนภาพอนุกรมเวลาระหว่างค่าพยากรณ์และค่าจริงของปริมาณออกซิเจนละลายน้ำพบว่า ลักษณะแผนภาพมีค่าใกล้เคียงกัน จากนั้นทำการตรวจสอบประสิทธิภาพของการพยากรณ์ด้วยค่าคลาดเคลื่อนสมบูรณ์เฉลี่ยและรากของค่าคลาดเคลื่อนกำลังสองเฉลี่ยพบว่า ค่าคลาดเคลื่อนสมบูรณ์เฉลี่ยมีค่าเท่ากับ 1.0742 และรากของค่าคลาดเคลื่อนกำลังสองเฉลี่ยมีค่าเท่ากับ 1.4469

5.2 อภิปรายผลการศึกษา

จากโครงการสถิติของกนกอร แก้วดวงสี (2560) ซึ่งศึกษาการพยากรณ์ปริมาณออกซิเจนละลายน้ำในแม่น้ำบางปะกงโดยใช้ข้อมูลตั้งแต่ปี พ.ศ. 2551 ถึงปี พ.ศ. 2560 โดยใช้การวิเคราะห์การถดถอยเชิงเส้นพหุคูณ ผลการศึกษาพบว่า สมการถดถอยเชิงเส้นพหุคูณที่ใช้พยากรณ์ปริมาณออกซิเจนละลายน้ำในแม่น้ำบางปะกงมีดังนี้

$$\widehat{DO} = -3.80 + 1.099\text{pH} + 0.000046\text{EC}$$

สมการถดถอยดังกล่าวมีค่าสัมประสิทธิ์การตัดสินใจที่ปรับค่าแล้วเท่ากับ 40.16 เมื่อเปรียบเทียบกับสมการถดถอยเชิงเส้นพหุคูณที่ใช้พยากรณ์ปริมาณออกซิเจนละลายน้ำในแม่น้ำบางปะกงในการศึกษาครั้งนี้จากสมการที่ 14 พบว่า สมการถดถอยเชิงเส้นพหุคูณสมการที่ 14 มีประสิทธิภาพในการพยากรณ์ปริมาณออกซิเจนละลายน้ำในแม่น้ำบางปะกงมากกว่า

5.2 ข้อเสนอแนะ

ในการศึกษาครั้งนี้ปริมาณออกซิเจนละลายน้ำในแม่น้ำบางปะกง หากนำสมการถดถอยที่ได้ไปใช้ในการพยากรณ์ปริมาณออกซิเจนละลายน้ำในแม่น้ำอื่น ค่าพยากรณ์ที่ได้อาจมีค่าไม่ใกล้เคียงกับค่าจริง ดังนั้นควรศึกษาปัจจัยมลพิษทางน้ำเพิ่มเติม เพื่อให้มีประสิทธิภาพและความแม่นยำในการพยากรณ์มากขึ้น

บรรณานุกรม

- กนกอร แก้วดวงสี (2560). *การพยากรณ์ปริมาณออกซิเจนละลายน้ำในแม่น้ำบางปะกงโดยสมการถดถอยเชิงเส้นพหุคูณ*. คณะวิทยาศาสตร์ มหาวิทยาลัยบูรพา.
- กรมควบคุมมลพิษ, *การบริหารจัดการและพัฒนาคุณภาพน้ำบางปะกงแบบบูรณาการ*. วันที่ค้นหาข้อมูล 21 กุมภาพันธ์ 2562 เข้าถึงได้จาก, http://www.pcd.go.th/info_serv/pol_bangprakong.html
- กรมควบคุมมลพิษ, *มาตรฐานคุณภาพน้ำ*. วันที่ค้นหาข้อมูล 17 กุมภาพันธ์ 2562 เข้าถึงได้จาก, http://www.pcd.go.th/info_serv/reg_std_water05.html
- กรมควบคุมมลพิษ. (2550), *รายงานประจำปีสำนักจัดการคุณภาพน้ำ*. วันที่ค้นหาข้อมูล 6 มีนาคม 2562 เข้าถึงได้จาก, http://infofile.pcd.go.th/water/water_annual50.pdf
- กลุ่มงานกลยุทธ์ศาสตร์และข้อมูลเพื่อการพัฒนาจังหวัด. (2560), *ข้อมูลสภาพทั่วไปของจังหวัด*. วันที่ค้นหาข้อมูล 27 กุมภาพันธ์ 2562 เข้าถึงได้จาก, http://www.chachoengsao.go.th/cco/images/file_links/plan2557-2560.pdf
- จตุภัทร เมฆพ่ายพ. (2560). *การวิเคราะห์การถดถอย*. เอกสารประกอบการสอน ภาควิชาคณิตศาสตร์ คณะวิทยาศาสตร์ มหาวิทยาลัยบูรพา.
- จันทวัน เบ็ญจวรรณ, *ทรัพยากรน้ำ*. วันที่ค้นหาข้อมูล 16 กุมภาพันธ์ 2562 เข้าถึงได้จาก, http://human.tru.ac.th/elearning/Human%20Being/human-detail3_6.html
- มงคล ต๊ะอุ่นและสุทธิพงษ์ เป็รื่องคำ. (2546), *น้ำเสียและแนวทางการลดผลกระทบ*. วันที่ค้นหาข้อมูล 17 กุมภาพันธ์ 2562 เข้าถึงได้จาก, https://home.kku.ac.th/uac/journal/year_11_2_2546/07_11_2_2546.pdf
- วนิดา พงษ์ศักดิ์ชาติ. *การวิเคราะห์การถดถอย*. เอกสารประกอบการสอน ภาควิชาคณิตศาสตร์ คณะวิทยาศาสตร์ มหาวิทยาลัยบูรพา.
- วรชาดา บัวไพร. (2555), *คุณภาพน้ำเพื่อการอุปโภคบริโภคอย่างยั่งยืนในเขตพื้นที่ อำเภอบางคนที จังหวัดสมุทรสงคราม*. วันที่ค้นหาข้อมูล 9 กุมภาพันธ์ 2562 เข้าถึงได้จาก, http://www.ssruir.ssru.ac.th/bitstream/ssruir/859/1/rid_031_2555.pdf
- สำนักจัดการน้ำคุณภาพน้ำ กรมควบคุมมลพิษ. (2561), *จุดตรวจวัดคุณภาพน้ำ*. วันที่ค้นหาข้อมูล 9 กุมภาพันธ์ 2562 เข้าถึงได้จาก, http://iwis.pcd.go.th/index.php?method=water_quality&etc=1550340427183
- Anderson, T. W., & Darling, D. A. (1952). Asymptotic theory of certain "goodness of fit" criteria based on stochastic processes. *The annals of mathematical statistics*, 23(2), 193-212.

บรรณานุกรม (ต่อ)

- Antanasijević, D., Pocajt, V., Povrenović, D., Perić-Grujić, A., & Ristić, M. (2013). Modelling of dissolved oxygen content using artificial neural networks: Danube River, North Serbia, case study. *Environmental Science and Pollution Research*, 20(12), 9006-9013.
- Chen, W. B., & Liu, W. C. (2015). Water quality modeling in reservoirs using multivariate linear regression and two neural network models. *Advances in Artificial Neural Systems*, 2015, 6.
- Nemati, S., Fazelifard, M. H., Terzi, Ö., & Ghorbani, M. A. (2015). Estimation of dissolved oxygen using data-driven techniques in the Tai Po River, Hong Kong. *Environmental earth sciences*, 74(5), 4065-4073.
- Welsch, R. E., & Kuh, E. (1977). Technical Report 923-77: Linear Regression Diagnostics. Cambridge, MA: Sloan School of Management, Massachusetts Institute of Technology, 775.

ภาคผนวก

ตารางที่ 1ก ค่าวิกฤตของตัวสถิติแอนเดอร์สัน-ดาร์ลิง

α	0.1	0.05	0.025	0.01
A_{crit}^2	0.631	0.752	0.873	1.035

ที่มา : <http://www.statisticalengineering.com/goodness.htm>

ตารางที่ 2ก ค่าวิกฤตของตัวสถิติเดออร์บิน-วัตสัน ที่ระดับนัยสำคัญ 0.05

จำนวนตัวแปรทำนายในตัวแบบการถดถอย																					
n	k=1		k=2		k=3		k=4		k=5		k=6		k=7		k=8		k=9		k=10		
	d _L	d _U	d _L	d _U	d _L	d _U	d _L	d _U	d _L	d _U	n	d _L	d _U	d _L	d _U	d _L	d _U	d _L	d _U	d _L	d _U
6	0.610	1.400	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----
7	0.700	1.356	0.467	1.896	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----
8	0.763	1.332	0.559	1.777	0.367	2.287	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----
9	0.824	1.320	0.629	1.699	0.455	2.128	0.296	2.588	----	----	----	----	----	----	----	----	----	----	----	----	----
10	0.879	1.320	0.697	1.641	0.525	2.016	0.376	2.414	0.243	2.822	----	----	----	----	----	----	----	----	----	----	----
11	0.927	1.324	0.758	1.604	0.595	1.928	0.444	2.283	0.315	2.645	0.203	3.004	----	----	----	----	----	----	----	----	----
12	0.971	1.331	0.812	1.579	0.658	1.864	0.512	2.177	0.380	2.506	0.268	2.832	0.171	3.149	----	----	----	----	----	----	----
13	1.010	1.340	0.861	1.562	0.715	1.816	0.574	2.094	0.444	2.390	0.328	2.692	0.230	2.985	0.147	3.266	----	----	----	----	----
14	1.045	1.350	0.905	1.551	0.767	1.779	0.632	2.030	0.505	2.296	0.389	2.572	0.286	2.848	0.200	3.111	0.127	3.360	----	----	----
15	1.077	1.361	0.946	1.543	0.814	1.750	0.685	1.977	0.562	2.220	0.447	2.471	0.343	2.727	0.251	2.979	0.175	3.216	0.111	3.438	----
16	1.106	1.371	0.982	1.539	0.857	1.728	0.734	1.935	0.615	2.157	0.502	2.388	0.398	2.624	0.304	2.860	0.222	3.090	0.155	3.304	----
17	1.133	1.381	1.015	1.536	0.897	1.710	0.779	1.900	0.664	2.104	0.554	2.318	0.451	2.537	0.356	2.757	0.272	2.975	0.198	3.184	----
18	1.158	1.391	1.046	1.535	0.933	1.696	0.820	1.872	0.710	2.060	0.603	2.258	0.502	2.461	0.407	2.668	0.321	2.873	0.244	3.073	----
19	1.180	1.401	1.074	1.536	0.967	1.685	0.859	1.848	0.752	2.023	0.649	2.206	0.549	2.396	0.456	2.589	0.369	2.783	0.290	2.974	----
20	1.201	1.411	1.100	1.537	0.998	1.676	0.894	1.828	0.792	1.991	0.691	2.162	0.595	2.339	0.502	2.521	0.416	2.704	0.336	2.885	----
21	1.221	1.420	1.125	1.538	1.026	1.669	0.927	1.812	0.829	1.964	0.731	2.124	0.637	2.290	0.546	2.461	0.461	2.633	0.380	2.806	----
22	1.239	1.429	1.147	1.541	1.053	1.664	0.958	1.797	0.863	1.940	0.769	2.090	0.677	2.246	0.588	2.407	0.504	2.571	0.424	2.735	----
23	1.257	1.437	1.168	1.543	1.078	1.660	0.986	1.785	0.895	1.920	0.804	2.061	0.715	2.208	0.628	2.360	0.545	2.514	0.465	2.670	----
24	1.273	1.446	1.188	1.546	1.101	1.656	1.013	1.775	0.925	1.902	0.837	2.035	0.750	2.174	0.666	2.318	0.584	2.464	0.506	2.613	----
25	1.288	1.454	1.206	1.550	1.123	1.654	1.038	1.767	0.953	1.886	0.868	2.013	0.784	2.144	0.702	2.280	0.621	2.419	0.544	2.560	----
26	1.302	1.461	1.224	1.553	1.143	1.652	1.062	1.759	0.979	1.873	0.897	1.992	0.816	2.117	0.735	2.246	0.657	2.379	0.581	2.513	----
27	1.316	1.469	1.240	1.556	1.162	1.651	1.084	1.753	1.004	1.861	0.925	1.974	0.845	2.093	0.767	2.216	0.691	2.342	0.616	2.470	----
28	1.328	1.476	1.255	1.560	1.181	1.650	1.104	1.747	1.028	1.850	0.951	1.959	0.874	2.071	0.798	2.188	0.723	2.309	0.649	2.431	----
29	1.341	1.483	1.270	1.563	1.198	1.650	1.124	1.743	1.050	1.841	0.975	1.944	0.900	2.052	0.826	2.164	0.753	2.278	0.681	2.396	----
30	1.352	1.489	1.284	1.567	1.214	1.650	1.143	1.739	1.071	1.833	0.998	1.931	0.926	2.034	0.854	2.141	0.782	2.251	0.712	2.363	----
31	1.363	1.496	1.297	1.570	1.229	1.650	1.160	1.735	1.090	1.825	1.020	1.920	0.950	2.018	0.879	2.120	0.810	2.226	0.741	2.333	----
32	1.373	1.502	1.309	1.574	1.244	1.650	1.177	1.732	1.109	1.819	1.041	1.909	0.972	2.004	0.904	2.102	0.836	2.203	0.769	2.306	----
33	1.383	1.508	1.321	1.577	1.258	1.651	1.193	1.730	1.127	1.813	1.061	1.900	0.994	1.991	0.927	2.085	0.861	2.181	0.796	2.281	----
34	1.393	1.514	1.333	1.580	1.271	1.652	1.208	1.728	1.144	1.808	1.079	1.891	1.015	1.978	0.950	2.069	0.885	2.162	0.821	2.257	----
35	1.402	1.519	1.343	1.584	1.283	1.653	1.222	1.726	1.160	1.803	1.097	1.884	1.034	1.967	0.971	2.054	0.908	2.144	0.845	2.236	----
36	1.411	1.525	1.354	1.587	1.295	1.654	1.236	1.724	1.175	1.799	1.114	1.876	1.053	1.957	0.991	2.041	0.930	2.127	0.868	2.216	----
37	1.419	1.530	1.364	1.590	1.307	1.655	1.249	1.723	1.190	1.795	1.131	1.870	1.071	1.948	1.011	2.029	0.951	2.112	0.891	2.197	----
38	1.427	1.535	1.373	1.594	1.318	1.656	1.261	1.722	1.204	1.792	1.146	1.864	1.088	1.939	1.029	2.017	0.970	2.098	0.912	2.180	----
39	1.435	1.540	1.382	1.597	1.328	1.658	1.273	1.722	1.218	1.789	1.161	1.859	1.104	1.932	1.047	2.007	0.990	2.085	0.932	2.164	----
40	1.442	1.544	1.391	1.600	1.338	1.659	1.285	1.721	1.230	1.786	1.175	1.854	1.120	1.924	1.064	1.997	1.008	2.072	0.952	2.149	----
45	1.475	1.566	1.430	1.615	1.383	1.666	1.336	1.720	1.287	1.776	1.238	1.835	1.189	1.895	1.139	1.958	1.089	2.022	1.038	2.088	----
50	1.503	1.585	1.462	1.628	1.421	1.674	1.378	1.721	1.335	1.771	1.291	1.822	1.246	1.875	1.201	1.930	1.156	1.986	1.110	2.044	----
55	1.528	1.601	1.490	1.641	1.452	1.681	1.414	1.724	1.374	1.768	1.334	1.814	1.294	1.861	1.253	1.909	1.212	1.959	1.170	2.010	----
60	1.549	1.616	1.514	1.652	1.480	1.689	1.444	1.727	1.408	1.767	1.372	1.808	1.335	1.850	1.298	1.894	1.260	1.939	1.222	1.984	----
65	1.567	1.629	1.536	1.662	1.503	1.696	1.471	1.731	1.438	1.767	1.404	1.805	1.370	1.843	1.336	1.882	1.301	1.923	1.266	1.964	----
70	1.583	1.641	1.554	1.672	1.525	1.703	1.494	1.735	1.464	1.768	1.433	1.802	1.401	1.838	1.369	1.874	1.337	1.910	1.305	1.948	----
75	1.598	1.652	1.571	1.680	1.543	1.709	1.515	1.739	1.487	1.770	1.458	1.801	1.428	1.834	1.399	1.867	1.369	1.901	1.339	1.935	----
80	1.611	1.662	1.586	1.688	1.560	1.715	1.534	1.743	1.507	1.772	1.480	1.801	1.453	1.831	1.425	1.861	1.397	1.893	1.369	1.925	----
85	1.624	1.671	1.600	1.696	1.575	1.721	1.550	1.747	1.525	1.774	1.500	1.801	1.474	1.829	1.448	1.857	1.422	1.886	1.396	1.916	----
90	1.635	1.679	1.612	1.703	1.589	1.726	1.566	1.751	1.542	1.776	1.518	1.801	1.494	1.827	1.469	1.854	1.445	1.881	1.420	1.909	----
95	1.645	1.687	1.623	1.709	1.602	1.732	1.579	1.755	1.557	1.778	1.535	1.802	1.512	1.827	1.489	1.852	1.465	1.877	1.442	1.903	----
100	1.654	1.694	1.634	1.715	1.613	1.736	1.592	1.758	1.571	1.780	1.550	1.803	1.528	1.826	1.506	1.850	1.484	1.874	1.462	1.898	----
150	1.720	1.747	1.706	1.760	1.693	1.774	1.679	1.788	1.665	1.802	1.651	1.817	1.637	1.832	1.622	1.846	1.608	1.862	1.593	1.877	----
200	1.758	1.779	1.748	1.789	1.738	1.799	1.728	1.809	1.718	1.820	1.707	1.831	1.697	1.841	1.686	1.852	1.675	1.863	1.665	1.874	----

ที่มา: <http://www.dm.unibo.it>