**การทำนายสัมประสิทธิ์การดูดซับสารเคมีของดินด้วยเทคนิคการสร้างแบบจำลองเชิงโมเลกุล**

**วศิน พิทักษ์ตระกูล และ สมศักดิ์ เพียรวณิช**

**ภาควิชาเคมี คณะวิทยาศาสตร์ จุฬาลงกรณ์มหาวิทยาลัย**

**Email: 6172058023@student.chula.ac.th**

**บทคัดย่อ**

มลพิษทางดิน เป็นหนึ่งในมลพิษที่ได้รับผลกระทบจากการกระทำของธรรมชาติและมนุษย์ ซึ่งส่วนใหญ่เกิดจากการใช้ยาฆ่าแมลง และปุ๋ยเคมี เพื่อให้พืชออกผลผลิตและเติบโตมาอย่างงดงาม แต่ก็ทำให้ดินนั้นได้ดูดซับสารเคมีเอาไว้ ในงานวิจัยนี้จึงได้ศึกษาหาสมบัติของสารเคมีที่มีผลต่อการดูดซับสารของดิน และหาแบบจำลองเพื่อทำนายค่าสัมประสิทธิ์การดูดซับสารของดิน (KOC) ด้วยเทคนิคการหาความสัมพันธ์เชิงปริมาณระหว่างโครงสร้างกับสมบัติ (QSPR) โดยเริ่มจากการนำสารทั้งหมดจำนวน 1,330 สาร มาสร้างโครงสร้าง 3 มิติ และทำการคำนวณปรับโครงสร้างด้วยวิธีเซมิ-เอ็มพิริกัล พีเอ็ม 7 จากนั้นนำมาคำนวณค่าสมบัติทางเคมีกายภาพ จำนวนรวม 1,536 สมบัติ ใช้วิธีการวิเคราะห์การถดถอยเชิงเส้นพหุคูณในการหาสมบัติของสารที่มีผลต่อการดูดซับสารเคมีของดิน และสร้างแบบจำลอง QSPR ซึ่งแบบจำลองที่ได้มีความสัมพันธ์และประสิทธิภาพที่ดีในการทำนาย โดยมีค่า r2 = 0.761 และ q2 = 0.761

**คำสำคัญ:** มลพิษทางดิน ความสัมพันธ์เชิงปริมาณระหว่างโครงสร้างกับสมบัติ ค่าสัมประสิทธิ์ของการดูดซับสารของดิน สมบัติทางเคมีกายภาพ

**Prediction of soil sorption coefficient of chemicals by molecular modeling technique**

**Wasin Pithaktrakool and Somsak Pianwanit**

**Department of Chemistry, Faculty of science, Chulalongkorn University**

**Email: 6172058023@student.chula.ac.th**

**ABSTRACT**

 Soil pollution is one of the pollutions caused by nature and human activities, which are mostly the use of insecticide and chemical fertilizer to make plants grow well and produce more products. However, these chemicals are absorbed by soil. In this research work, properties of chemical compound affecting soil sorption were analyzed, and model that can predict soil sorption coefficient of chemicals (KOC) was examined by using quantitative structure property relationship (QSPR) technique. Totally 1,330 compounds were selected, and their three-dimensional structures were constructed. Semi-empirical PM7 method was used for geometry optimization calculations. A total of 1,536 physicochemical properties were computed. Multiple linear regression was employed to analyze properties of compound affecting soil sorption and to find QSPR model. The obtained model has a good relationship and predictive ability with r2 = 0.761 and q2 = 0.761.

**Keywords:** soil pollution, quantitative structure property relationship, soil sorption coefficient, physicochemical property

**บทนำ**

 ในปัจจุบันมีสารเคมีจำนวนมากที่ถูกส่งผ่านลงไปสู่ดินผ่านกระบวนการต่าง ๆ ทั้งที่เกิดขึ้นเองตามธรรมชาติ ได้แก่ ฝนกรด แผ่นดินไหว สึนามิ และเกิดจากกิจกรรมของมนุษย์ ได้แก่ การฝังกลบขยะ การทำเหมืองแร่ การใช้ปุ๋ย การทิ้งขยะและของเสียจากโรงงานอุตสาหกรรม การใช้ยาฆ่าแมลงหรือยากำจัดวัชพืชเพื่อป้องกันพืชผลเสียหายจากแมลงและศัตรูพืช
โดยกิจกรรมเหล่านี้ทำให้มีสารเคมีตกค้างในดินเป็นจำนวนมาก จนก่อให้เกิดมลพิษทางดินอันจะส่งผลเสียต่อพืชและสิ่งมีชีวิตต่าง ๆ ที่อาศัยอยู่ในดิน และท้ายที่สุดก่อให้เกิดอันตรายต่อสุขภาพของสัตว์และมนุษย์ นักวิจัยจึงสนใจศึกษาหาว่าสารเคมีแต่ละชนิดจะถูกดูดซับไว้ในดินได้ในปริมาณมากน้อยเพียงใด เพื่อใช้เป็นข้อมูลสำหรับการวางแผนและควบคุมปริมาณสารเคมีตกค้างในดิน รวมทั้งใช้ในการประเมินค่าความเสี่ยงของผลกระทบต่อสิ่งแวดล้อมของสารเคมีแต่ละชนิด อย่างไรก็ตามการทำการทดลองเพื่อวัดปริมาณสารเคมีที่ถูกดูดซับของดินต้องมีค่าใช้จ่ายทั้งเรื่องเครื่องมือ สารเคมี และต้องใช้เวลาในการวิเคราะห์ ดังนั้นจึงได้มีการใช้เทคนิคการวิเคราะห์หาความสัมพันธ์เชิงปริมาณระหว่างโครงสร้างกับสมบัติ (Quantitative Structure Property Relationship, QSPR) มาช่วยในการทำนายค่าการดูดซับ ซึ่งจะทำให้สามารถทำนายค่าได้อย่างสะดวกและรวดเร็ว ทำให้ประหยัดเวลาและเงินทุนได้เป็นอย่างมาก

 ความสัมพันธ์เชิงปริมาณระหว่างโครงสร้างกับสมบัติ เป็นเทคนิคทางสารสนเทศทางเคมีที่เกี่ยวกับการสร้างแบบจำลองทางคณิตศาสตร์ที่แสดงความสัมพันธ์เชิงปริมาณระหว่างโครงสร้างของสารกับสมบัติของสาร ซึ่งวิธีนี้จะเป็นการแสดงออกของโมเลกุลผ่านค่าตัวเลขต่าง ๆ เรียกค่าตัวเลขที่แสดงออกว่า “ตัวบ่งชี้ (descriptor)” หลังจากนั้นจะใช้ตัวบ่งชี้เพื่อคำนวณสมบัติทางเคมีกายภาพของโมเลกุล โดยแบบจำลอง QSPR จะให้ค่าที่มีประสิทธิภาพในการประมาณค่าสมบัติทางเคมีกายภาพ เช่น ค่าคงที่การแตกตัว ความสามารถในการละลาย ไดโพลโมเมนต์ เป็นต้น

 สารอินทรีย์ในดินมีความสามารถในการดูดซับสารเคมีที่ชอบละลายในไขมัน ซึ่งการดูดซับสารเคมีของดินนิยมวัดค่าในรูปของค่าสัมประสิทธิ์การดูดซับสารของดิน (organic carbon-water partition coefficient, Koc) จากอัตราส่วนระหว่างความเข้มข้นของสารเคมีในดินที่ละลายอยู่ในชั้นสารอินทรีย์ (CSOC) ต่อความเข้มข้นของสารเคมีที่ละลายในน้ำ (CW) ดังสมการ

$$K\_{OC} = \frac{C\_{soc}}{C\_{w}}$$

จากสมการข้างต้น จะพบว่าค่าสัมประสิทธิ์การดูดซับสารของดิน (Koc) น่าจะมีความสัมพันธ์กับค่าสัมประสิทธิ์การกระจายของสารในชั้นน้ำมันกับน้ำ (octanol-water partition coefficient, log P) ซึ่งงานวิจัยบางส่วนใช้สมบัติ log P เป็นพารามิเตอร์ในการสร้างแบบจำลอง แต่ก็มีงานวิจัยอื่น ๆ ที่ไม่ได้ใช้ log P เป็นพารามิเตอร์ ทำให้แบบจำลอง QSPR มีความแตกต่างที่หลากหลาย โดยงานวิจัยที่ได้ศึกษาหาแบบจำลอง QSPR ในการทำนายค่าสัมประสิทธิ์การดูดซับสารของดินมีอยู่หลายงาน มีการใช้พารามิเตอร์ที่หลากหลายและแตกต่าง เช่น งานของ Eduardo และคณะในปี ค.ศ. 2003 ได้สร้างแบบจำลอง QSPR ของสารจำนวน 82 สาร โดยใช้พารามิเตอร์จำนวน 5 ตัว ได้แก่ จำนวนวงเบนซีน มวลโมเลกุล จำนวนอะตอมของไนโตรเจน ออกซิเจน และซัลเฟอร์ แต่งานของ Carlos และคณะในปี ค.ศ. 2019 ได้สร้างแบบจำลอง QSPR ของสารจำนวน 639 สาร โดยใช้พารามิเตอร์แค่เพียงตัวเดียว คือ AlogPs ดังนั้นผู้วิจัยจึงต้องการศึกษาหาว่าสมบัติใดของสารเคมีที่มีผลต่อค่าการดูดซับ และสร้างแบบจำลอง QSPR สำหรับกลุ่มสารที่มีจำนวนมากยิ่งขึ้น

 ดังนั้นคณะผู้วิจัยมีความสนใจที่จะศึกษาหาแบบจำลองทางคณิตศาสตร์ที่สามารถทำนายค่าสัมประสิทธิ์ของการดูดซับของการดูดสารเคมีของดิน โดยการศึกษาตัวแปรต่าง ๆ ของสารทั้งหมด ด้วยเทคนิคความสัมพันธ์เชิงปริมาณระหว่างโครงสร้างกับสมบัติ เพื่อทราบตัวแปรที่เป็นปัจจัยต่อการดูดซับสารเคมีของดิน

**วัตถุประสงค์การวิจัย**

ในการศึกษาครั้งนี้มีวัตถุประสงค์การวิจัย ประกอบด้วย

1. เพื่อศึกษาหาสมบัติของสารเคมีที่มีผลต่อการดูดซับสารของดิน

2. เพื่อศึกษาหาแบบจำลองในการทำนายค่าสัมประสิทธิ์การดูดซับสารของดิน

**วิธีดำเนินการวิจัย**

**1. เครื่องมือที่ใช้การทำวิจัย**

1.1 เครื่องคอมพิวเตอร์

 1.2 โปรแกรม HyperChem 8.0

 1.3 โปรแกรม MOPAC 2016

 1.4 โปรแกรม Materials studio

 1.5 โปรแกรม Mordred

**2. การเก็บรวบรวมข้อมูล และการวิเคราะห์ข้อมูล**

 2.1 ค้นคว้า ศึกษาเอกสาร และรวบรวมข้อมูลต่าง ๆ ที่เป็นผลจากการทดลอง ได้แก่ ชื่อสารเคมี โครงสร้างของสารเคมี ค่าสัมประสิทธิ์การดูดซับสารของดิน (log Koc) จากบทความในวารสารทางวิชาการตั้งแต่ปี ค.ศ. 1994 – 2019 จำนวนรวม 1,330 สาร

 2.2 สร้างโครงสร้าง 3 มิติของสารทั้ง 1,330 สารด้วยโปรแกรม HyperChem ทำการคำนวณปรับโครงสร้างเสถียรโดยใช้วิธีเซมิ-เอมพิริกัล พีเอ็ม 7 ด้วยโปรแกรม MOPAC 2016 และคำนวณสมบัติทางเคมีกายภาพ (physicochemical properties) ของสารแต่ละตัวโดยใช้โปรแกรม Materials studio และ Mordred จำนวนรวม 1,536 สมบัติ

 2.3 ทำการแบ่งสารออกเป็น 2 กลุ่ม ได้แก่ กลุ่มทดลอง (training set) จำนวน 930 สาร และกลุ่มทดสอบ (test set) จำนวน 400 สาร นำสารในกลุ่มทดลองมาทำการวิเคราะห์เพื่อหาว่าสมบัติของสารที่คำนวณได้จากข้อ 2.2 มีสมบัติใดที่มีผลต่อการดูดซับสารเคมีของดิน จากนั้นทำการสร้างแบบจำลอง QSPR โดยใช้สมบัตินี้ โดยใช้วิธีการวิเคราะห์การถดถอยเชิงเส้นพหุคูณ (Multiple linear regression, MLR) ด้วยโปรแกรม Materials studio

 2.4 นำแบบจำลองที่สร้างได้มาตรวจสอบประสิทธิภาพ โดยนำแบบจำลองนั้นไปทำนายค่าสัมประสิทธิ์การดูดซับสารของดินของสารในกลุ่มทดสอบ แล้วเปรียบเทียบกับค่าจริงจากการทดลอง

 2.5 เมื่อได้แบบจำลองเชิงคณิตศาสตร์แล้ว ทำการปรับแบบจำลองที่มีประสิทธิภาพในการทำนายที่ดีที่สุด

**ผลการวิจัย**

**1. การแบ่งสารออกเป็นกลุ่มทดลองและกลุ่มทดสอบ**

 หลังจากที่คำนวณสมบัติทางเคมีกายภาพของโมเลกุล การเลือกสารที่จะมาใช้เป็นกลุ่มทดลอง สามารถทำได้โดยสร้างกราฟระหว่างจำนวนสารกับช่วงค่าสัมประสิทธิ์ของการดูดซับของดิน เพื่อดูการกระจายตัวของสาร แล้วทำการเลือกสารแบบสุ่มในแต่ละช่วง เพื่อให้สารในกลุ่มทดสอบมีการกระจายตัวแบบปกติ (Normal distribution) ซึ่งในงานวิจัยนี้ประกอบด้วยสารในกลุ่มแอลเคน แอลคีน แอลไคน์ แอลกอฮอล์ อีเทอร์ สารประกอบคาร์บอนิล เอมีน เบนซีนและอนุพันธ์ สารประกอบเฮเทอโรไซคลิก สารประกอบโพลีไซคลิกอะโรมาติกไฮโดรคาร์บอน และกลุ่มโลหะออร์กาโนต่าง ๆ มีโครงสร้างทั้งหมด 1,330 สาร ซึ่งแบ่งออกเป็นกลุ่มทดลองจำนวน 930 สาร และกลุ่มทดสอบจำนวน 400 สาร

**2. การสร้างความสัมพันธ์เชิงปริมาณระหว่างโครงสร้างกับสมบัติของสาร**

 ในการสร้างความสัมพันธ์เชิงปริมาณระหว่างโครงสร้างกับสมบัติของสาร จะใช้สมบัติทางเคมีกายภาพที่คำนวณได้จากโปรแกรม Materials studio จำนวน 24 ตัวแปร และโปรแกรม Mordred 1,519 ตัวแปร ในการสร้างความสัมพันธ์

 **2.1** **การสร้างความสัมพันธ์โดยใช้สมบัติทางเคมีกายภาพของโมเลกุลที่คำนวณด้วยโปรแกรม Materials studio**

 **2.1.1 สร้างความสัมพันธ์โดยใช้สมบัติทางเคมีกายภาพของโมเลกุลจำนวน 1 สมบัติ**

 การสร้างความสัมพันธ์เชิงปริมาณระหว่างโครงสร้างกับสมบัติของสาร โดยใช้สมบัติทางเคมีกายภาพของโมเลกุลจำนวน 1 สมบัติ ได้แก่ AlogP98 มีค่า r2 = 0.757 และ q2 = 0.756

**แบบจำลองที่ 1** log KOC = 0.579 \* AlogP98 + 1.231

**A**

**B**

**ภาพที่ 1** การเปรียบเทียบค่า log KOC จากการทดลองและการทำนายของ (A) กลุ่มทดลอง และ (B) กลุ่มทดสอบ

 **2.1.2 สร้างความสัมพันธ์โดยใช้สมบัติทางเคมีกายภาพของโมเลกุลจำนวน 17 สมบัติ**

 การสร้างความสัมพันธ์เชิงปริมาณระหว่างโครงสร้างกับสมบัติของสาร โดยใช้สมบัติทางเคมีกายภาพของโมเลกุลจำนวน 17 สมบัติ พบว่ามีสมบัติที่เกี่ยวข้องกับความสัมพันธ์ 4 สมบัติ ได้แก่ AlogP98, Hydrogen bond acceptor, Molecular flexibility และ Zagreb index มีค่า r2 = 0.800 และ q2 = 0.783

**แบบจำลองที่ 2** log KOC = 0.493 \* AlogP98 - 0.0981 \* Hydrogen bond acceptor +
0.0252 \* Molecular flexibility + 0.00569 \* Zagreb index + 1.366

**A**

**B**

**ภาพที่ 2** การเปรียบเทียบค่า log KOC จากการทดลองและการทำนายของ (A) กลุ่มทดลอง และ (B) กลุ่มทดสอบ

 **2.1.3 สร้างความสัมพันธ์โดยใช้สมบัติทางเคมีกายภาพของโมเลกุลจำนวน 24 สมบัติ**

 การสร้างความสัมพันธ์เชิงปริมาณระหว่างโครงสร้างกับสมบัติของสาร โดยใช้สมบัติทางเคมีกายภาพของโมเลกุลจำนวน 24 สมบัติ พบว่ามีสมบัติที่เกี่ยวข้องกับความสัมพันธ์ 6 สมบัติ ได้แก่ AlogP98, Hydrogen bond acceptor, Rotatable bond, Total dipole, Quadrupole zz และ Wiener index มีค่า r2 = 0.802 และ q2 = 0.794

**แบบจำลองที่ 3** log KOC = 0.564 \* AlogP98 – 0.0669 \* Hydrogen bond acceptor –
0.0225 \* Rotatable bonds – 0.0441 \* Total dipole +
0.00777 \* Quadrupole zz+ 0.000336 \* Wiener index + 1.538

**A**

**B**

**ภาพที่ 3** การเปรียบเทียบค่า log KOC จากการทดลองและการทำนายของ (A) กลุ่มทดลอง และ (B) กลุ่มทดสอบ

 **2.2** **การสร้างความสัมพันธ์โดยใช้สมบัติทางเคมีกายภาพของโมเลกุลที่คำนวณด้วยโปรแกรม Mordred**

 **2.2.1 สร้างความสัมพันธ์โดยใช้สมบัติทางเคมีกายภาพของโมเลกุลจำนวน 1 สมบัติ**

 การสร้างความสัมพันธ์เชิงปริมาณระหว่างโครงสร้างกับสมบัติของสาร โดยใช้สมบัติทางเคมีกายภาพของโมเลกุลจำนวน 1 สมบัติ พบว่ามีสมบัติที่เกี่ยวข้องกับความสัมพันธ์ 1 สมบัติ ได้แก่ SlogP มีค่า r2 = 0.761 และ q2 = 0.761

**แบบจำลองที่ 4** log KOC = 0.636 \* SlogP + 1.152

**A**

**B**

**ภาพที่ 4** การเปรียบเทียบค่า log KOC จากการทดลองและการทำนายของ (A) กลุ่มทดลอง และ (B) กลุ่มทดสอบ

 **2.2.2 สร้างความสัมพันธ์โดยใช้สมบัติทางเคมีกายภาพของโมเลกุลจำนวน 1,519 สมบัติ**

 การสร้างความสัมพันธ์เชิงปริมาณระหว่างโครงสร้างกับสมบัติของสาร โดยใช้สมบัติทางเคมีกายภาพของโมเลกุลจำนวน 1,519 สมบัติ พบว่ามีสมบัติที่เกี่ยวข้องกับความสัมพันธ์ 6 สมบัติ ได้แก่ MDEO-11, IC0, nHRing, SlogP, nG12FAHRing และ C3SP3 มีค่า r2 = 0.803 และ q2 = 0.791

**แบบจำลองที่ 5** log KOC = 0.326 \* MDEO-11 – 0.931 \* IC0 + 0.150 \* nHRing + 0.629 \* SlogP +
1.362 \* nG12FAHRing – 0.140 \* C3SP3 + 2.434

**A**

**B**

**ภาพที่ 5** การเปรียบเทียบค่า log KOC จากการทดลองและการทำนายของ (A) กลุ่มทดลอง และ (B) กลุ่มทดสอบ

**ตารางที่ 1** การเปรียบเทียบตัวแปร ค่า r2 และค่า q2 ของสมการทั้ง 5 สมการ

|  |  |  |  |
| --- | --- | --- | --- |
| **แบบจำลอง** | **ตัวแปร** | **r2** | **q2** |
| 1 | AlogP98 | 0.757 | 0.756 |
| 2 | AlogP98, Hydrogen bond acceptor, Molecular flexibility, Zagreb index | 0.800 | 0.783 |
| 3 | AlogP98, Hydrogen bond acceptor, Rotatable bonds, Total dipole, Quadrupole zz, Wiener index | 0.802 | 0.792 |
| 4 | SlogP | 0.761 | 0.761 |
| 5 | SlogP, MDEO-11, IC0, nHRing, nG12FAHRing, C3SP3 | 0.803 | 0.794 |

 ตารางที่ 1 แสดงให้เห็นแบบจำลองทั้ง 5 แบบ สมบัติทางเคมีกายภาพของโมเลกุลในแบบจำลองที่ 1-3 คำนวณจากโปรแกรม Materials studio และในแบบจำลองที่ 4-5 คำนวณจากโปรแกรม Mordred เมื่อเปรียบเทียบแบบจำลองทั้ง 5 แบบ พบว่าแบบจำลองที่มีค่าทางสถิติสูงที่สุดคือ แบบจำลองที่ 5 อย่างไรก็ตามแบบจำลองนี้ใช้ตัวแปร 6 ตัวแปร และเมื่อเทียบกับแบบจำลองที่ 4 ซึ่งใช้ตัวแปรเพียง 1 ตัว และมีค่าทางสถิติที่ไม่แตกต่างอย่างมีนัยสำคัญ จึงทำให้เลือกแบบจำลองที่ 4 ซี่งใช้ตัวแปรเพียง 1 ตัว คือ SlogP นอกจากนั้นสมบัติทางเคมีกายภาพนี้ยังสามารถคำนวณได้จากโปรแกรม Mordred ซึ่งเป็นโปรแกรมฟรี จึงทำให้นักวิจัยสามารถใช้งานได้อย่างสะดวก ไม่มีข้อจำกัดเรื่องค่าใช้จ่ายของโปรแกรม

**3. การเปรียบเทียบค่า r2 ของสารแต่ละกลุ่มโดยใช้แบบจำลองที่ 4 และ 5**

 เนื่องจากแบบจำลองที่ 4 สร้างขึ้นสำหรับสารที่มีโครงสร้างหลากหลายกลุ่ม ผู้วิจัยจึงสนใจที่จะศึกษาว่าหากทำการแยกสารออกเป็นแต่ละกลุ่มโครงสร้าง การใช้สมบัติ SlogP เพียงสมบัติเดียวจะยังคงให้แบบจำลองที่มีประสิทธิภาพดี หรือการเพิ่มสมบัติอื่น ๆ เข้าไปจะทำให้ได้แบบจำลองที่มีประสิทธิภาพสูงขึ้น ดังนั้นจึงได้นำตัวแปรของแบบจำลองที่ 5 ได้แก่ MDEO-11, IC0, nHRing, SlogP, nG12FAHRing และ C3SP3 มาสร้างความสัมพันธ์กับสารในแต่ละกลุ่มโครงสร้าง ซึ่งมีทั้งหมด 13 กลุ่มโครงสร้าง และมีจำนวนสารในแต่ละกลุ่มที่แตกต่างกัน ดังแสดงในตารางที่ 2 ผลการคำนวณพบว่า สารกลุ่มที่ 1, 2 และ 3 ยังคงใช้สมบัติ SlogP เพียงสมบัติเดียว แต่ค่าทางสถิติไม่สูงนัก บ่งชี้ว่าสารกลุ่มโครงสร้างเหล่านี้อาจจะมีปัจจัยอื่นที่ส่งผลต่อการดูดซับของดินหรือมีกลไกการดูดซับที่ซับซ้อนและแตกต่างออกไป สารกลุ่มที่ 4 ซึ่งมีโครงสร้างหลักเป็นแอลกอฮอล์ เมื่อเพิ่มสมบัติ MDEO-11 เข้าไปในแบบจำลอง คือใช้ 2 ตัวแปร (SlogP และ MDEO-11) ทำให้แบบจำลองมีประสิทธิภาพที่ดีมาก โดยมีค่า r2 สูงถึง 0.960 และค่า q2 สูงถึง 0.895 สารกลุ่มที่ 5 ใช้ตัวแปร SlogP และ MDEO-11 เหมือนสารกลุ่มที่ 4 แต่มีค่าทางสถิติที่ต่ำกว่า สารกลุ่มที่ 6, 7 และ 8 ใช้ตัวแปร 2 ตัว เช่นกัน แต่เปลี่ยนเป็น SlogP และ IC0 สารกลุ่มที่ 9 ใช้ตัวแปรเพิ่มขึ้นเป็น 3 ตัวแปร คือ SlogP, MDEO-11, IC0 ส่วนสารกลุ่มที่ 10 และ 11 ใช้ตัวแปร 4 ตัว คือ SlogP, MDEO-11, IC0 และ C3SP3 สารกลุ่มที่ 12 ใช้ตัวแปร 4 ตัว คือ SlogP, MDEO-11, nG12FAHRing และ C3SP3 ผลที่ได้บ่งชี้ว่ากลไกการดูดซับสารของดินน่าจะมีความซับซ้อนและในแต่ละขั้นตอนมีความเกี่ยวข้องกับสมบัติของสารที่หลากหลายและแตกต่างกัน จึงทำให้ได้แบบจำลองที่มีความแตกต่างกันในแต่ละกลุ่มโครงสร้าง

**ตารางที่ 2** เปรียบเทียบความสัมพันธ์ของแบบจำลองสารแต่ละกลุ่มเมื่อใช้ตัวแปร SlogP กับใช้ตัวแปรเพิ่มขึ้น

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| **กลุ่ม** | **กลุ่มโครงสร้างสาร** | **จำนวนสาร** | **แบบจำลองที่ใช้****ตัวแปร SlogP** | **แบบจำลองที่ใช้****ตัวแปรเพิ่มขึ้น** |
| **r2** | **q2** | **r2** | **q2** |
| 1 | Polycyclic aromatic hydrocarbon | 58 | 0.344 | 0.310 | 0.344 | 0.310 |
| 2 | Ether | 20 | 0.494 | 0.339 | 0.494 | 0.339 |
| 3 | Organochloride | 11 | 0.410 | -0.070 | 0.410 | -0.070 |
| 4 | Alcohol | 55 | 0.922 | 0.877 | 0.960 | 0.895 |
| 5 | Organophosphorus | 60 | 0.278 | 0.225 | 0.334 | 0.188 |
| 6 | Amine | 35 | 0.517 | 0.403 | 0.723 | 0.693 |
| 7 | Nitrile | 18 | 0.738 | 0.723 | 0.796 | 0.742 |
| 8 | Organosulfur | 19 | 0.370 | -0.401 | 0.546 | -0.407 |
| 9 | Alkane, Alkene, Alkyne | 149 | 0.555 | 0.550 | 0.647 | 0.607 |
| 10 | Carbonyl compound | 184 | 0.614 | 0.602 | 0.673 | 0.614 |
| 11 | Benzene & Benzene derivatives | 539 | 0.566 | 0.562 | 0.639 | 0.627 |
| 12 | Heterocyclic compound | 173 | 0.326 | 0.316 | 0.417 | 0.281 |
| 13 | Other compounds | 9 | - | - | - | - |

**สรุปผลและอภิปรายผล**

 ในงานวิจัยนี้ได้รวบรวมข้อมูลสารเคมีทั้งหมด 1,330 สาร นำมาสร้างโครงสร้าง 3 มิติ ทำการคำนวณปรับโครงสร้างเสถียรด้วยวิธี semi-empirical PM7 และคำนวณสมบัติทางเคมีกายภาพของสารโดยใช้โปรแกรม Materials studio และ Mordred จำนวนรวม 1,536 สมบัติ จากนั้นแบ่งออกเป็น 2 กลุ่ม ได้แก่ กลุ่มทดลอง 930 สาร และกลุ่มทดสอบ 400 สาร นำสารในกลุ่มทดลองมาทำการวิเคราะห์พบว่า การดูดซับสารของดินจะขึ้นกับสมบัติ log P ของสาร โดยค่าสัมประสิทธิ์การดูดซับสารของดินมีความสัมพันธ์เชิงบวกกับค่า log P ซึ่งแสดงว่าสารที่มีค่า log P สูงจะถูกดูดซับโดยดินได้มาก ผลที่ได้นี้สอดคล้องกับงานวิจัยของ Ralpho และคณะ (2013), Yonghua และคณะ (2014) และ Carlos และคณะ (2019) อีกทั้งยังเป็นการสนับสนุนทฤษฎีที่คิดว่าค่าสัมประสิทธิ์การดูดซับสารของดิน (Koc) มีความสัมพันธ์กับค่า log P

จากการนำสารในกลุ่มทดลองมาทำการสร้างแบบจำลองความสัมพันธ์ระหว่างค่าสัมประสิทธิ์การดูดซับสารของดินกับสมบัติของสารด้วยเทคนิค QSPR โดยใช้วิธีการวิเคราะห์การถดถอยเชิงเส้นพหุคูณ (MLR) พบว่าได้แบบจำลอง 5 แบบ ที่แตกต่างกัน อย่างไรก็ตามทุกแบบจำลองมีตัวแปรที่เป็นพื้นฐานเหมือนกันคือ log P และความแตกต่างอยู่ที่การเพิ่มตัวแปรอื่น ๆ เข้าไปในแบบจำลอง แบบจำลองที่ 1 และ 4 ใช้ตัวแปรเพียงตัวเดียวคือ log P แต่ตัวแปรนี้คำนวณด้วยโปรแกรมที่แตกต่างกัน กล่าวคือ AlogP98 ในแบบจำลองที่ 1 คำนวณด้วยโปรแกรม Materials studio ในขณะที่ SlogP ในแบบจำลองที่ 4 คำนวณด้วยโปรแกรม Mordred จึงทำให้ค่าตัวแปรทั้งสองนี้มีความแตกต่างกันเล็กน้อย อย่างไรก็ตามค่าทางสถิติของทั้งสองแบบจำลองก็มีค่าไม่แตกต่างกัน แต่เนื่องจากโปรแกรม Mordred เป็นโปรแกรมฟรี ทำให้ทุกคนเข้าถึงและใช้งานได้ จึงได้เลือกแบบจำลองที่ 4 ส่วนแบบจำลองที่ 2, 3 และ 5 มีการเพิ่มตัวแปรอื่น ๆ เข้าไปในแบบจำลองทำให้ใช้ตัวแปร 4 ตัว (แบบจำลองที่ 2) และ 6 ตัวแปร (แบบจำลองที่ 3 และ 5) ทำให้ค่าทางสถิติดีขึ้น ค่า r2 เพิ่มขึ้นจาก 0.761 ในแบบจำลองที่ 4 เป็น 0.800, 0.802, 0.803 ซึ่งคิดเป็นการเพิ่มขึ้นเพียงร้อยละ 5.12 ถึง 5.25 เท่านั้นในขณะที่ต้องใช้ตัวแปรเพิ่มขึ้นจาก 1 ตัวแปรเป็น 4 และ 6 ตัวแปร ในทางสถิติถือว่าเป็นการเพิ่มขึ้นอย่างไม่มีนัยสำคัญ ดังนั้นจึงไม่ควรเลือกแบบจำลองที่ 2, 3 และ 5 สรุปได้ว่าแบบจำลองที่เหมาะสมที่สุดคือ แบบจำลองที่ 4 ซึ่งมีสมการเป็น log KOC = 0.63 \* SlogP + 1.152 และมีค่า r2 = 0.761 และ q2 = 0.761 และเมื่อนำสมการที่ได้นี้ไปใช้ทำนายค่าสัมประสิทธิ์การดูดซับสารของดินของสารในกลุ่มทดสอบจำนวน 400 สาร แล้วนำมาหาความสัมพันธ์กับค่าสัมประสิทธิ์การดูดซับสารของดินจากการทดลองพบว่ามีค่า r2 = 0.572 แบบจำลองที่ได้นี้มีค่าทางสถิติที่ดีและสามารถใช้งานได้กับสารที่มีความหลากหลายทางโครงสร้างมากกว่างานวิจัยก่อนหน้าที่สร้างแบบจำลองจากสารเพียงแค่ 639 สารเท่านั้น

 นอกจากนี้ยังได้ทดลองวิเคราะห์และสร้างแบบจำลองแยกตามแต่ละกลุ่มโครงสร้างของสาร โดยทดลองใช้ตัวแปรเพียงตัวเดียวคือ SlogP และทดลองใช้ตัวแปรเพิ่มขึ้น คือ MDEO-11, IC0, nHRing, nG12FAHRing และ C3SP3 พบว่าในสารกลุ่มที่ 1, 2 และ 3 การเพิ่มตัวแปรไม่ส่งผลต่อประสิทธิภาพของแบบจำลอง กล่าวคือ แบบจำลองใช้แค่ตัวแปร SlogP เพียงตัวเดียวก็พอ อย่างไรก็ตามค่าทางสถิติของแบบจำลองไม่สูงนัก บ่งชี้ว่าสารกลุ่มโครงสร้างเหล่านี้อาจจะมีปัจจัยอื่นที่ส่งผลต่อการดูดซับสารของดินหรือมีกลไกการดูดซับสารที่ซับซ้อนและแตกต่างออกไป ในสารกลุ่มที่ 4 ถึง 12 การเพิ่มตัวแปรเข้าไปในแบบจำลองจะทำให้ค่าทางสถิติของแบบจำลองมีค่าเพิ่มขึ้น ซึ่งถือเป็นเรื่องปกติในแง่สถิติ ดังนั้นจึงได้พิจารณาเฉพาะการเพิ่มขึ้นอย่างมีนัยสำคัญคือ ค่า r2 ต้องมีค่าเพิ่มขึ้นมากกว่า 10% ต่อการเพิ่มตัวแปรเข้าไป 1 ตัว พบว่ามีเพียงสารกลุ่มที่ 5, 6 และ 8 เท่านั้นที่สอดคล้องกับเกณฑ์นี้ อย่างไรก็ตามเมื่อพิจารณาค่า q2 ของแบบจำลองสารกลุ่มเหล่านี้พบว่า สารกลุ่มที่ 5 มีค่า q2 ลดลง ในขณะที่สารกลุ่มที่ 8 ค่า q2 ลดลงและติดลบ ดังนั้นการเพิ่มตัวแปรในแบบจำลองของสารกลุ่มที่ 5 และ 8 ไม่ได้ทำให้ประสิทธิภาพของแบบจำลองดีขึ้น ส่วนสารกลุ่มที่ 6 มีค่า q2 เพิ่มขึ้นอย่างมีนัยสำคัญ ดังนั้นจึงมีเพียงสารกลุ่มที่ 6 ซึ่งมีโครงสร้างหลักเป็นเอมีนเท่านั้นที่แบบจำลองมีประสิทธิภาพเพิ่มขึ้น ผลที่ได้นี้ช่วยยืนยันว่าการดูดซับสารของดินจะขึ้นกับสมบัติ log P ของสารเป็นหลัก ไม่ว่าสารจะมีโครงสร้างเป็นแบบใด นอกจากนี้ยังบ่งชี้ว่ากลไกการดูดซับสารของดินน่าจะมีความซับซ้อนและในแต่ละขั้นตอนมีความเกี่ยวข้องกับสมบัติของสารที่หลากหลายและแตกต่างกัน จึงทำให้ได้แบบจำลองที่มีความแตกต่างกันในแต่ละกลุ่มโครงสร้าง

**ข้อเสนอแนะ**

 งานวิจัยนี้ให้ข้อมูลเกี่ยวกับค่าสัมประสิทธิ์ของการดูดซับของดิน โดยค่าที่ได้จากการทดลองและการทำนายนั้นอาจมีความคลาดเคลื่อน เนื่องจากการทดลองจริงกับการทำนายนั้นต่างกัน การทดลองในห้องปฏิบัติการนั้นเป็นการทดลองทางโมเลกุลจริงที่เกิดอันตรกิริยาระหว่างรูพรุนของดินกับสารเคมีที่ละลายอยู่ในสารละลาย และคำนวณหาปริมาณสารที่เหลือจากการดูดซับของดิน เพื่อหาค่าสัมประสิทธิ์ของการดูดซับของดินได้ แต่การทำนายเป็นการทดลองทางคอมพิวเตอร์ ที่สามารถทำนายค่าสัมประสิทธิ์การดูดซับสารของดิน ที่คำนวณได้จากสมบัติของสาร ซึ่งเมื่อเทียบค่าสัมประสิทธิ์ของการดูดซับของดินระหว่างค่าที่ได้จากการทดลองจริงกับค่าที่ได้จากการทำนายนั้น อาจทำให้มีค่าตรงกันหรือต่างกันก็ได้ อย่างไรก็ตามการทดลองทางคอมพิวเตอร์ช่วยให้ประหยัดงบประมาณ ประหยัดเวลาในการทดลอง และเป็นการลดการใช้สารเคมี ไม่ก่อให้เกิดของเสียสู่สิ่งแวดล้อม แบบจำลองที่ได้ในงานวิจัยนี้มีประสิทธิภาพที่ดี แต่อาจจะสามารถพัฒนาปรับปรุงเพิ่มประสิทธิภาพขึ้นได้โดยการเพิ่มจำนวนสมบัติทางเคมีกายภาพให้ครอบคลุมยิ่งขึ้น เช่น สมบัติที่เกี่ยวข้องกับกลไกการดูดซับ ซึ่งต้องมีการศึกษาเพิ่มเติมโดยใช้โปรแกรมที่มีความละเอียดแม่นยำสูงและใช้เวลาการคำนวณที่นานขึ้น แล้วนำมาหาความสัมพันธ์ต่อไป

**เอกสารอ้างอิง**

กุลยา โอตากะ. (2547). *เคมีสิ่งแวดล้อม* (พิมพ์ครั้งที่ 5). กรุงเทพฯ: สำนักพิมพ์มหาวิทยาลัยรามคำแหง

H. Lohninger. (1994). Estimation of soil partition coefficients of pesticides from their chemical structure*. Chemosphere, 28*(8), 1611-1626.

E.J. Delgado, J.B. Alderete & G.A. Jaña. (2003). A simple QSPR model for predicting soil sorption coefficient of polar and nonpolar organic compound from molecular formula. *The Journal for Chemical Information and Computer scientists, 43*(6), 1928-1932.

I. Kahn, D. Fara, M. Karelson & U. Maran. (2005). QSPR treatment of the soil sorption coefficient of organic pollutants. *Journal of Chemical Information and Modeling, 45*(1), 94-105.

P.R. Duchowicz, M.P. Gonzalez, A.M. Helguera, M.N.D.S. Cordeiro & E.A. Castro. (2005). Application of the replacement method as novel variable selection in QSPR. 2. Soil sorption coefficients. *Chemometrics and intelligent laboratory system, 88*(2), 197-203.

N. Goudarzi, M. Goodzaqri, M.C.U. Araujo & R.K.H. Galvao. (2009). QSPR modeling of soil sorption coefficients (KOC) of pesticides using SPA-ANN and SPA-MLR. *Journal of agricultural and food chemistry, 57*(15), 7153-7158.

R.R. dos Reis, S.C. Sampiao & E.B. de Melo. (2013). The effect of different log P algorithms on the modeling of the soil sorption coefficient of nonionic pesticides. *Water Research, 47*(15), 5751-5759.

Y. Shao, J. Liu, M. Wang, L. Shi, X. Yao & P. Gramatica. (2014). Integrated QSPR models to predict the soil sorption coefficient for a large diverse set of compounds by using different modeling methods. *Atmospheric Environment, 88*, 212-218.

C.J.M Olguin, S.C. Sampaio, R.R dos Reis, M.B. Remor & C.F.A. Olguin. (2019). QSPR modelling of the soil sorption coefficient from training sets of different sizes. *SAR and QSAR in environmental research,* *30*(5), 299-311.