

## วิธีการหาจุดตัดจากการวิเคราะห์ข้อมูลการจราจรขนาดใหญ่ในประเทศไทย

นาย นฤเดช คล้ายแก้ว รหัสประจำตัวนิสิต 6430198121 ภาควิชา วิศวกรรมโยธา

นาย ธนภพ วงศ์น้อย รหัสประจำตัวนิสิต 6430153921 ภาควิชา วิศวกรรมโยธา

นาย วรณัฐ บำบัดเนตร รหัสประจำตัวนิสิต 6431139321 ภาควิชา วิศวกรรมสำรวจ

ปริญญานิพนธ์นี้เป็นส่วนหนึ่งของการศึกษาวิชาโครงการทางวิศวกรรม (ข้ามสาขาวิชา)

ตามหลักสูตรวิศวกรรมศาสตรบัณฑิต

คณะวิศวกรรมศาสตร์ จุฬาลงกรณ์มหาวิทยาลัย

ปีการศึกษา 2567

## Study on Bottleneck Identification methodology from Traffic Big Data in Thailand

Mr. Naruedech Klaikaew Student ID 6430198121 Department of civil engineering

Mr. Tanapob Wongnoi Student ID 6430153921 Department of civil engineering

Mr. Wannatuch Bangbamned Student ID 6431139321 Department of Survey  
Engineering

A Senior Project Submitted in Partial Fulfillment of the Requirements

For the Degree of Bachelor of Engineering Program

Faculty of Engineering Chulalongkorn University

Academic Year 2024

หัวข้อโครงการ	วิธีการหาจุดติดขัดจากการวิเคราะห์ข้อมูลการจราจรขนาดใหญ่ในประเทศไทย
โดย	นาย นฤเดช คล้ายแก้ว นาย ธนภพ วงศ์น้อย นาย วรณรัช บำบำเนตร
อาจารย์ที่ปรึกษา	รศ.ดร.สรวิศ นฤปิติ ผศ.ดร.กรวิก ตันภษรานนท์

---

### บทคัดย่อ

งานวิจัยนี้มีวัตถุประสงค์เพื่อพัฒนาแนวทางการวิเคราะห์สภาพการจราจรโดยอาศัยข้อมูล GPS ของยานพาหนะ แล้วแปลงให้อยู่ในรูปของดัชนี Traffic Performance Index (TPI) ซึ่งเป็นตัวชี้วัดเชิงปริมาณที่สามารถสะท้อนระดับความหนาแน่นของจราจรในแต่ละช่วงเวลาได้อย่างเป็นระบบ

โดยข้อมูล GPS ที่ใช้ในการศึกษาได้จากการบันทึกตำแหน่งและความเร็วของรถยนต์ที่วิ่งผ่านถนนสายหลักภายในพื้นที่ จากนั้นนำมาประมวลผลด้วยวิธีการจัดกลุ่มข้อมูลเชิงเวลา เพื่อคำนวณค่าความเร็วเฉลี่ยในแต่ละช่วงเวลา และแปลงเป็นค่าดัชนี TPI ให้อยู่ในช่วง 0 ถึง 1 เพื่อให้สามารถเปรียบเทียบระดับความแออัดระหว่างวันและช่วงเวลาได้อย่างเหมาะสม

ต่อจากนั้น ได้จัดทำตารางค่า TPI รายวันของถนนเป้าหมาย และนำข้อมูลดังกล่าวไปใช้ในการสร้างแบบจำลองเพื่อพยากรณ์ค่าดัชนี TPI สำหรับช่วงเวลาถัดไป โดยใช้ข้อมูลย้อนหลังเป็นปัจจัยอิสระ เพื่อประเมินแนวโน้มของสภาพการจราจรล่วงหน้า

ผลการศึกษาแสดงให้เห็นว่าแนวทางที่พัฒนาขึ้นสามารถแปลงข้อมูล GPS ให้เป็นดัชนีที่สะท้อนสภาพการจราจรได้อย่างมีประสิทธิภาพ และสามารถนำไปต่อยอดในการพัฒนาระบบช่วยตัดสินใจด้านการบริหารจัดการจราจรต่อไป

คำสำคัญ: ข้อมูล GPS, ดัชนี TPI, การจราจรในเขตเมือง, การพยากรณ์การจราจร, ถนนพระราม 4

**Title** Study on Bottleneck Identification methodology from Traffic Big Data in Thailand

**Students** Mr. Naruedech Klaikaew  
Mr. Tanapob Wongnoi  
Mr. Wannatuch Bangbamned

**Advisor** Associate Professor Sorawit Narupiti Ph.D.  
Assistant Professor Garavig Tanaksaranond Ph.D.

---

### **Abstract**

This study aims to develop a methodology for analyzing urban traffic conditions using GPS data collected from vehicles, which are then converted into the Traffic Performance Index (TPI). TPI is a quantitative indicator designed to systematically reflect the level of traffic congestion at different times.

The GPS data used in this study were obtained from the recorded positions and speeds of vehicles traveling on main roads within the study area. The data were processed through a time-series grouping approach to calculate the average vehicle speed at each time interval. Subsequently, the TPI values were normalized to a range between 0 and 1, allowing for meaningful comparison of congestion levels across different days and time periods.

Daily TPI tables were then constructed for the target road, and these data were used to develop a forecasting model for predicting TPI values in the next time interval. The model utilized historical data as independent variables to estimate the upcoming traffic conditions.

The results demonstrate that the proposed approach effectively transforms raw GPS data into an interpretable index that reflects real-world traffic conditions. Furthermore, the method shows potential for application in decision-support systems for traffic management.

Keywords: GPS data, Traffic Performance Index (TPI), urban traffic, traffic forecasting,  
Rama IV Road

## กิตติกรรมประกาศ

ข้าพเจ้าขอกราบขอบพระคุณ รศ.ดร.สรวิศ นฤปิติ และ ผศ.ดร.กรวิก ตันเกษรนนท์ อาจารย์ที่ปรึกษา ซึ่งได้กรุณาให้คำแนะนำ ตรวจสอบ ตลอดจนให้ข้อเสนอแนะที่เป็นประโยชน์อย่างยิ่งตลอดระยะเวลาการดำเนินงานวิจัยฉบับนี้ ด้วยความเมตตา อดทน และใส่ใจในรายละเอียด ซึ่งเป็นกำลังสำคัญที่ทำให้งานวิจัยฉบับนี้สำเร็จลงได้ด้วยดี

ขอขอบคุณ บริษัท Grab (ประเทศไทย) จำกัด ที่เอื้อเฟื้อข้อมูลและสนับสนุนเครื่องมือที่ใช้ในการดำเนินงานวิจัยฉบับนี้ ซึ่งมีบทบาทสำคัญต่อการวิเคราะห์และพัฒนาแนวทางการประเมินสภาพจราจรในเขตเมืองได้อย่างเป็นรูปธรรม

สุดท้ายนี้ ข้าพเจ้าขอขอบพระคุณทุกท่านที่มีส่วนเกี่ยวข้องในการดำเนินงานวิจัยฉบับนี้ แม้จะได้เียนามไว้ ณ ที่นี้

นฤเดช คล้ายแก้ว

ธนภาพ วงศ์น้อย

วรรณรัช บางบำเนตร

คณะผู้จัดทำ

## สารบัญ

บทคัดย่อ .....	ก
Abstract.....	ข
กิตติกรรมประกาศ .....	ค
สารบัญ.....	ง
สารบัญรูป .....	ฉ
บทที่ 1 บทนำ .....	1
1.1 ที่มาและความสำคัญ.....	1
1.2 วัตถุประสงค์ .....	2
1.3 ขอบเขตการวิจัย .....	2
1.4 ประโยชน์ที่คาดว่าจะได้รับ.....	2
บทที่ 2 วรรณกรรมที่เกี่ยวข้อง.....	3
2.1 งานวิจัยที่เกี่ยวข้อง .....	3
2.1.1 Turn-level traffic bottleneck identification using vehicle trajectory data...	3
2.1.2 Study on Traffic Congestion Patterns of Large City in China Taking Beijing as an Example .....	5
2.2 ทฤษฎีที่เกี่ยวข้อง .....	7
บทที่ 3 ระเบียบวิธีวิจัย .....	11
3.1 เตรียมข้อมูล.....	12
3.1.1 ดาวน์โหลดข้อมูลจากฐานข้อมูล.....	12
3.1.2 กรองข้อมูล.....	12
3.2 วิเคราะห์ข้อมูลเบื้องต้น.....	13
3.2.1 วิเคราะห์ข้อมูล.....	13
3.3 ปรับปรุงแบบจำลอง.....	15
3.3.1 การทำนายค่าด้วยวิธีการ Linear Regression.....	15

3.3.2 การทำนายค่าด้วยวิธีการ Bayesian ( $H_1/H_2$ ).....	16
3.3.3 การทำนายค่าด้วยวิธีการ Bayesian Continuous Distribution KDE.....	17
บทที่ 4 ผลการดำเนินงานวิจัย .....	19
4.1 ค่า TPI ในแต่ละชั่วโมงของทุกวัน.....	19
4.2 แบบจำลองที่ใช้ในงานวิจัย.....	21
4.2.1 วิธีการ Linear Regression.....	21
4.2.2 วิธีการ Bayesian ( $H_1/H_2$ ).....	22
4.2.3 วิธีการ Bayesian Continuous Distribution KDE.....	24
4.3 ผลการเปรียบเทียบทั้ง 3 แบบจำลอง.....	25
บทที่ 5 สรุปผลการวิจัย อภิปรายผล และข้อเสนอแนะ .....	27
5.1 สรุปผลงานวิจัย.....	27
5.2 อภิปรายผลงานวิจัย.....	28
5.3 ข้อเสนอแนะ.....	29
5.3.1 ข้อเสนอแนะในการนำผลวิจัยไปใช้.....	29
5.3.2 ข้อเสนอแนะในการทำวิจัยครั้งต่อไป.....	29
เอกสารอ้างอิง .....	31

## สารบัญรูป

รูปที่ 2.1 ช่วงค่าดัชนี TPI และการจำแนกระดับสภาพการจราจร	3
รูปที่ 2.2 กราฟค่าดัชนี TPI รายเวลาในช่วง 06:00–13:00 น.	4
รูปที่ 2.3 ภาพแสดงความสัมพันธ์ของการติดขัดระหว่างจุดตรวจสอบการจราจร	4
รูปที่ 2.4 กราฟรูปแบบเส้นโค้งค่า TPI รายวันในวันเสาร์และวันอาทิตย์	6
รูปที่ 3.1 ตัวอย่างข้อมูลที่คำนวณโหลตมาจากฐานข้อมูล โดยเลือกเฉพาะหลักที่จำเป็นในการวิเคราะห์ข้อมูล คือ driverid, wayids, date, time, speed ,projectedlat และ projectedlng	13
รูปที่ 3.2 กราฟค่า TPI ของแต่ละชั่วโมงของทุกวัน ในรูปแบบ Boxplot บริเวณประมาณ 100 เมตร ก่อนถึงแยกซอยบ้านกล้วยใต้ ถนนฝั่งมุ่งหน้าสู่แยกพระโขนง	14
รูปที่ 3.3 ตัวอย่างค่า TPI แต่ละชั่วโมงของทุกวัน บริเวณประมาณ 100 เมตร ก่อนถึงแยกซอยบ้านกล้วยใต้ ถนนฝั่งมุ่งหน้าสู่แยกพระโขนง	14
รูปที่ 3.4 กราฟการเปรียบเทียบระหว่างค่า True TPI และ ค่าที่ทำนายจากแบบจำลองด้วยวิธีการ Linear Regression	15
รูปที่ 3.5 กราฟการเปรียบเทียบระหว่างค่า True TPI และ ค่าที่ทำนายจากแบบจำลองด้วยวิธีการ Bayesian ( $H_1/H_2$ )	16
รูปที่ 3.6 กราฟการแจกแจงค่า TPI ของ Prior ให้เป็น Continuous Distribution โดยใช้วิธี Kernel Density Estimation	17
รูปที่ 3.7 กราฟการเปรียบเทียบระหว่างค่า True TPI และ ค่าที่ทำนายจากแบบจำลองด้วยวิธีการ Bayesian Continuous Distribution KDE	18
รูปที่ 4.1 กราฟการ Box plot เปรียบเทียบระหว่างค่า TPI ในแต่ละชั่วโมง ของทุกวัน 1 มิถุนายน – 31 สิงหาคม พ.ศ. 2562 บริเวณประมาณ 100 เมตร ก่อนถึงแยกซอยบ้านกล้วยใต้	19
รูปที่ 4.2 กราฟการ Box plot เปรียบเทียบระหว่างค่า TPI ในแต่ละชั่วโมง ของทุกวัน 1 มิถุนายน – 31 สิงหาคม พ.ศ. 2562 ระหว่างแยกบ้านกล้วยใต้ ถึง แยกกล้วยน้ำไท	20
รูปที่ 4.3 กราฟการเปรียบเทียบระหว่างค่า TPI จริง และค่าที่ทำนายจากแบบจำลองด้วยวิธีการ Linear Regression ทั้งหมด 27 วัน ของถนนประมาณ 100 เมตร ก่อนถึงแยกซอยบ้านกล้วยใต้	21



รูปที่ 4.4 กราฟการเปรียบเทียบระหว่างค่า TPI จริง และค่าที่ทำนายจากแบบจำลอง ด้วยวิธีการ Bayesian ( $H_1/H_2$ ) ทั้งหมด 27 วัน ของถนนประมาณ 100 เมตร ก่อนถึง แยกซอยบ้านกล้วยใต้	22
รูปที่ 4.5 กราฟการเปรียบเทียบระหว่างค่า TPI จริง และค่าที่ทำนายจากแบบจำลอง ด้วยวิธีการ Bayesian Continuous Distribution KDE ทั้งหมด 27 วัน ของถนน ประมาณ 100 เมตร ก่อนถึงแยกซอยบ้านกล้วยใต้	24
รูปที่ 4.6 กราฟการเปรียบเทียบระหว่างค่า TPI จริง และค่าที่ทำนายจากแบบจำลอง ด้วยวิธีการ Linear Regression, Bayesian ( $H_1/H_2$ ), Bayesian Continuous Distribution KDE ทั้งหมด 27 วัน ของถนนประมาณ 100 เมตร ก่อนถึงแยกซอยบ้าน กล้วยใต้	25

## บทที่ 1

### บทนำ

#### 1.1 ที่มาและความสำคัญ

ปัญหาการจราจรติดขัดในเขตเมือง โดยเฉพาะในกรุงเทพมหานคร ยังคงเป็นปัญหาหลักที่ส่งผลกระทบต่อประสิทธิภาพการเดินทาง การใช้พลังงาน และคุณภาพชีวิตของประชาชนอย่างต่อเนื่อง การจัดการจราจรที่มีประสิทธิภาพจำเป็นต้องอาศัยข้อมูลที่สามารถสะท้อนสถานการณ์จริงได้อย่างแม่นยำและทันเวลา เพื่อใช้ประกอบการวางแผนหรือปรับเปลี่ยนมาตรการควบคุมการจราจรได้ตรงจุด

การทราบสภาพความติดขัดของการจราจรที่แม่นยำในแต่ละช่วงเวลา ช่วยให้หน่วยงานสามารถประเมินความรุนแรงของปัญหาในพื้นที่ต่าง ๆ ได้เป็นระบบ และสามารถใช้อ้างอิงข้อมูลดังกล่าวในการประเมินผลหรือกำหนดนโยบายด้านการจราจร เช่น การปรับสัญญาณไฟจราจร การจัดเจ้าหน้าที่ในพื้นที่แออัด หรือการประชาสัมพันธ์เส้นทางเลี่ยง โดยเฉพาะในพื้นที่ที่มีข้อจำกัดด้านโครงสร้างพื้นฐานหรือมีความต้องการเดินทางสูงในช่วงเวลาเร่งด่วน

ยิ่งไปกว่านั้น ความสามารถในการพยากรณ์แนวโน้มความติดขัดล่วงหน้าอย่างแม่นยำ จะช่วยให้สามารถเตรียมการเชิงรุก (proactive traffic management) ได้อย่างมีประสิทธิภาพ เช่น การแจ้งเตือนผู้ใช้ถนนล่วงหน้า การจัดการเส้นทางก่อนเกิดจุดวิกฤต หรือการสื่อสารกับหน่วยงานภาคสนามได้ทันท่วงที ซึ่งจะช่วยลดผลกระทบจากความแออัดที่เกิดขึ้นกะทันหัน และเพิ่มความยืดหยุ่นของระบบจราจรโดยรวม

ในปัจจุบัน เทคโนโลยี GPS ที่ติดตั้งในยานพาหนะ เช่น Grab Taxi หรือระบบขนส่งสาธารณะ ได้กลายเป็นแหล่งข้อมูลที่มีความครอบคลุมและทันกาลสูงที่สุด เนื่องจากสามารถสะท้อนพฤติกรรมการเดินทางจริงของผู้ใช้ถนนในช่วงเวลาต่าง ๆ ได้ต่อเนื่องและหลากหลาย จึงมีศักยภาพในการนำมาประมวลผลเพื่อประเมินและพยากรณ์สภาพการจราจรได้อย่างครอบคลุมกว่าวิธีดั้งเดิม เช่น กล้องวงจรปิดหรือเครื่องนับรถ

จากแนวคิดดังกล่าว งานวิจัยนี้จึงมุ่งเน้นการใช้ข้อมูล GPS จากรถแท็กซี่มาประมวลผลเป็นดัชนี TPI (Traffic Performance Index) เพื่อสะท้อนระดับความหนาแน่นของการจราจรในแต่ละช่วงเวลา และพัฒนาแบบจำลองสำหรับการพยากรณ์ค่าดัชนี TPI ในช่วงเวลาถัดไป โดยเปรียบเทียบแบบจำลองหลากหลายรูปแบบ เพื่อหาแนวทางที่เหมาะสมที่สุดในการนำไปประยุกต์ใช้กับการจัดการจราจรเชิงรุกในอนาคต

## 1.2 วัตถุประสงค์

ศึกษาและวิเคราะห์การติดขัดของถนนโดยใช้ข้อมูล Probe Data เช่น ความเร็วเฉลี่ย ตำแหน่ง ระยะเวลาที่รถหยุดนิ่งเพื่อกำหนดเกณฑ์ที่เหมาะสมสำหรับการระบุตำแหน่งและสามารถคาดการณ์การติดขัดของถนนในช่วงเวลาต่อไปได้

## 1.3 ขอบเขตการวิจัย

1.3.1 ศึกษาข้อมูลที่มีซึ่งข้อมูลที่มีจะเป็นข้อมูลติดตามพิกัดจาก Grab Taxi Thailand โดยเลือกข้อมูล 1 มิถุนายน ถึง 31 สิงหาคม พ.ศ. 2562 มาดำเนินงาน โดยศึกษารูปแบบของข้อมูล และสิ่งที่มีอยู่ในข้อมูล

1.3.2 ศึกษาหลักการและวิธีการระบุจุดติดขัด

1.3.3 ทำการระบุขอบเขตพื้นที่ โดยพื้นที่ทำโครงการนี้คือ ถนนพระราม4 บริเวณประมาณ 100 เมตร ก่อนถึงแยกซอยบ้านกล้วยใต้ และบริเวณระหว่างแยกบ้านกล้วยใต้ ถึง แยกกล้วยน้ำไท

1.3.4 ทดลองใช้ค่าดัชนีประสิทธิภาพการจราจร (Traffic Performance Index : TPI) ในการหาจุดติดขัดบนถนนพระราม 4 และนำมาทำตารางการติดขัดในแต่ละช่วงเวลา

1.3.5 นำค่าจากตารางการติดขัด มาเข้าแบบจำลองที่พัฒนาเพื่อ คาดการณ์ค่า TPI ล่วงหน้า

## 1.4 ประโยชน์ที่คาดว่าจะได้รับ

1.4.1 สามารถพัฒนาวิธีการวิเคราะห์สภาพจราจรเชิงปริมาณ โดยใช้ดัชนี Traffic Performance Index (TPI) ซึ่งแปลงจากข้อมูล GPS ของยานพาหนะ ช่วยให้สามารถเปรียบเทียบระดับความติดขัดของถนนในช่วงเวลาต่าง ๆ ได้อย่างเป็นระบบ

1.4.2 สามารถใช้แบบจำลองที่พัฒนาเพื่อ คาดการณ์ค่าดัชนี TPI ล่วงหน้า ช่วยสนับสนุนการตัดสินใจในการเลือกเส้นทางของผู้ขับขี่ โดยเฉพาะในช่วงเวลาที่มีแนวโน้มเกิดความแออัด

1.4.3 ช่วยให้สามารถระบุรูปแบบความติดขัดที่เกิดขึ้นในช่วงเวลาต่าง ๆ ของวัน หรือในแต่ละวันของสัปดาห์ ซึ่งเป็นประโยชน์ต่อการวางแผนการเดินทาง

## บทที่ 2

### วรรณกรรมที่เกี่ยวข้อง

#### 2.1 งานวิจัยที่เกี่ยวข้อง

##### 2.1.1 Turn-level network traffic bottleneck identification using vehicle trajectory data<sup>[1]</sup>

(Wei et al., 2022)

งานวิจัยนี้มีการวิเคราะห์ลักษณะของการจราจรในพื้นที่เมือง โดยเฉพาะการระบุช่วงเวลาหรือพื้นที่ที่มีความติดขัดสูง จำเป็นต้องอาศัยตัวชี้วัดที่สามารถเปรียบเทียบสภาพการจราจรได้ในเชิงปริมาณ งานวิจัยของ Wei และคณะ (2022) ได้เสนอวิธีการใช้ ข้อมูล GPS จาก trajectory ของยานพาหนะ เพื่อคำนวณค่าดัชนี Traffic Performance Index (TPI) เพื่อวัดระดับความติดขัดของการจราจร

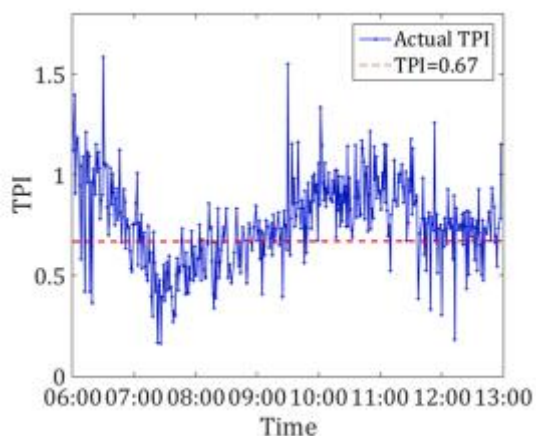
TPI เป็นตัวชี้วัดที่สะท้อนถึงประสิทธิภาพของการเคลื่อนที่ของยานพาหนะ โดยเปรียบเทียบความเร็วจริงของรถ (actual speed) กับ ความเร็วในภาวะการจราจรไหลลื่น (free-flow speed) ซึ่งหมายถึงช่วงเวลาที่ไม่มีสิ่งรบกวนจากสัญญาณไฟหรือสภาพแวดล้อมอื่น ๆ ค่า TPI ที่ได้จะมีค่าระหว่าง 0 ถึง 1 โดยค่าที่ต่ำกว่าบ่งชี้ถึงความติดขัดที่รุนแรงยิ่งขึ้น โดยในการศึกษาได้มีการแบ่งระดับของสภาพการจราจรตามค่าของ TPI

Traffic condition categories.

TPI	less than 0.48	0.48-0.56	0.56-0.67	0.67-0.83	greater than 0.83
Traffic conditions	Serious congestion	Moderate congestion	Mild congestion	Basically smooth	Smooth

รูปที่ 2.1 ช่วงค่าดัชนี TPI และการจำแนกระดับสภาพการจราจร

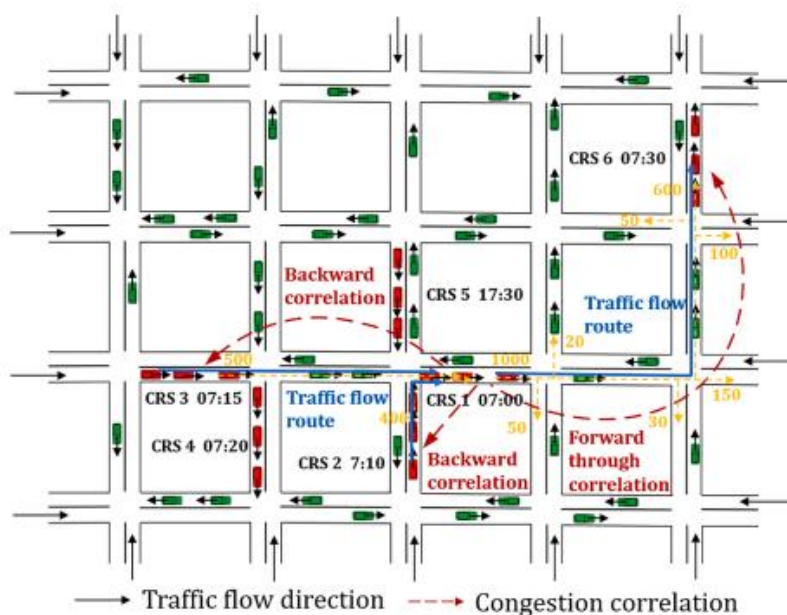
ที่มา: Turn-level network traffic bottleneck identification using vehicle trajectory data



รูปที่ 2.2 กราฟค่าดัชนี TPI รายเวลาในช่วง 06:00–13:00 น.

ที่มา: Turn-level network traffic bottleneck identification using vehicle trajectory data

จากการคำนวณค่า TPI ดังกล่าว งานวิจัยได้ใช้เกณฑ์  $TPI < 0.67$  เป็นเกณฑ์ในการพิจารณาว่าถนนนั้นอยู่ในภาวะ “ติดขัด” หากมีค่าต่ำกว่าค่าดังกล่าวอย่างต่อเนื่องในช่วงเวลาหนึ่ง ซึ่งสอดคล้องกับแนวคิดในงานนี้ที่ใช้ ค่า TPI เฉลี่ยแบบ normalized แทนค่าความเร็วตรง เพื่อให้สามารถเปรียบเทียบสภาพจราจรระหว่างวัน ช่วงเวลา หรือพื้นที่ต่าง ๆ ได้อย่างมีมาตรฐาน



รูปที่ 2.3 ภาพแสดงความสัมพันธ์ของการติดขัดระหว่างจุดตรวจสอบการจราจร

ที่มา: Turn-level network traffic bottleneck identification using vehicle trajectory data

จากนั้นได้นำแนวคิดของความน่าจะเป็นแบบเบย์ (Bayesian Inference) มาใช้ในการประเมินความน่าจะเป็นของการแพร่กระจายของการจราจรติดขัดระหว่างคู่ถนนต่าง ๆ โดยใช้การแจกแจงแบบเบตา (Beta distribution) เพื่ออัปเดตความเชื่อมั่น (posterior probability) ของความสัมพันธ์นั้นเมื่อมีข้อมูลการสังเกตเพิ่มเติม ทำให้สามารถประเมิน "ต้นทุนของความติดขัด" (congestion cost) ในแต่ละทิศทางได้อย่างแม่นยำ โดยคำนวณจากเวลาในการติดขัด (congestion time), ความล่าช้าในการเดินทาง (delay) และความยาวของคิวรถ (queue length)

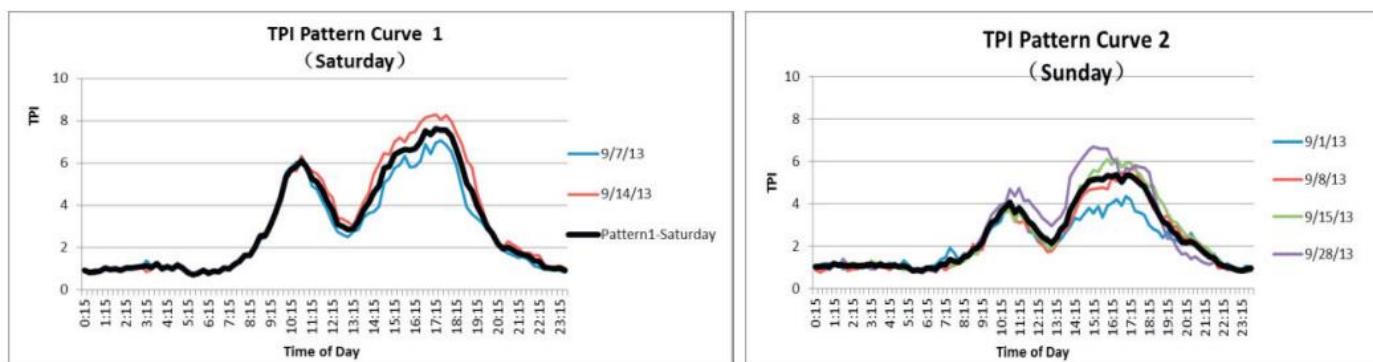
### 2.1.2 Study on Traffic Congestion Patterns of Large City in China Taking Beijing as an Example<sup>[2]</sup>

(Huimin Wen, Jianping Sun, Xi Zhang, 2014)

งานวิจัยนี้ได้ทำการศึกษา “รูปแบบความติดขัดของการจราจร” (Traffic Congestion Patterns) ในเขตเมืองขนาดใหญ่ โดยใช้กรุงปักกิ่ง ประเทศจีนเป็นพื้นที่ศึกษา งานวิจัยนี้ถือเป็นตัวอย่างสำคัญที่นำค่าดัชนี Traffic Performance Index (TPI) มาใช้วิเคราะห์และจำแนกสถานะจราจรในระดับเมือง โดยใช้ข้อมูลจากระบบ GPS ของรถแท็กซี่กว่า 40,000 คัน ซึ่งส่งข้อมูลตำแหน่งและความเร็วทุกนาทีตลอด 24 ชั่วโมง

ข้อมูล TPI ถูกประมวลผลทุก ๆ 15 นาที ครอบคลุมพื้นที่ใจกลางเมือง (ภายในวงแหวนที่ห้า) และมีการรวบรวมข้อมูลตลอดระยะเวลา 6 ปี โดยในงานวิจัยได้นำข้อมูล TPI รายวันมาทำ Clustering Analysis เพื่อจัดกลุ่มลักษณะของวันต่าง ๆ ตามรูปแบบความแออัด โดยอิงจากค่าดัชนีต่าง ๆ เช่น ค่าเฉลี่ย TPI ช่วงเวลาเช้าและเย็น, ค่าสูงสุด, ความแปรปรวนของค่า และจำนวนช่วงเวลาที่ TPI สูงเกินค่าที่กำหนด ผลการจัดกลุ่มสามารถแบ่งรูปแบบวันออกได้ถึง 6 กลุ่มหลัก คือ

- วันธรรมดาทั่วไป
- วันธรรมดาที่มีความติดขัดรุนแรง
- วันก่อนวันหยุดซึ่งมีความต้องการเดินทางสูง
- วันเสาร์
- วันอาทิตย์
- วันหยุดยาว เช่น เทศกาลไหว้พระจันทร์และวันชาติจีน



รูปที่ 2.4 กราฟรูปแบบเส้นโค้งค่า TPI รายวันในวันเสาร์และวันอาทิตย์

ที่มา: Study on Traffic Congestion Patterns of Large City in China Taking Beijing as an Example

นอกจากนี้ งานวิจัยยังเสนอว่า รูปแบบของกราฟ (TPI Curve) ของแต่ละประเภทวันมีลักษณะเฉพาะที่สามารถนำไปใช้ในการคาดการณ์ความแออัดในอนาคตได้ เช่น ลักษณะช่วงเวลาเร่งด่วน (peak time), ความสูงของยอดความแออัด (congestion crest) และความยาวของช่วงเวลาที่ติดขัด (duration)

## 2.2 ทฤษฎีที่เกี่ยวข้อง

### ทฤษฎีการเปลี่ยนค่า Latitude Longitude เป็น พิกัดระบบ UTM<sup>[3]</sup>

(Manchuk, 2009)

เนื่องจากในการหาค่าระยะห่างระหว่างจุดไม่สามารถหาได้จากค่า Latitude และ Longitude ได้โดยตรงเพราะเป็นค่าพิกัดแบบมุม จึงต้องแปลงเป็นระบบพิกัด UTM เป็นระบบพิกัดของโลกที่ถูกทำการ Projection แล้ว เพื่อจะได้ทำการหาความห่างระหว่างจุดได้แม่นยำมากขึ้น โดยที่จะใช้สมการดังนี้ในการแปลงพิกัด

$$x = k_0 N \left[ A + (1 - T + C) \frac{A^3}{6} + (5 - 18T + T^2 + 72C - 58e'^2) \frac{A^5}{120} \right] \quad (2.1)$$

$$y = k_0 \left[ M - M_0 + N \tan \phi \left[ \frac{A^2}{2} + (5 - T + 9C + 4C^2) \frac{A^4}{24} + (61 - 58T + T^2 + 600C - 300e'^2) \frac{A^6}{720} \right] \right] \quad (2.2)$$

โดยที่

$$k_0 = 0.9996$$

$$e'^2 = \frac{e^2}{1 - e^2}$$

$$N = \frac{a}{\sqrt{(1 - e^2 \sin^2 \phi)}}$$

$$T = \tan^2 \phi$$

$$C = e'^2 \cos^2 \phi$$

$$A = \lambda - \lambda_0 \cos \phi$$

$$M = a \left[ \left( 1 - \frac{e^2}{4} - \frac{3e^4}{64} - \frac{5e^6}{256} - \dots \right) \phi - \left( \frac{3e^2}{8} + \frac{3e^4}{32} + \frac{45e^6}{1024} + \dots \right) \sin^2 \phi + \left( \frac{15e^4}{256} + \frac{45e^6}{1024} + \dots \right) \sin^4 \phi - \left( \frac{35e^6}{3072} + \dots \right) \sin^6 \phi + \dots \right] \quad (2.3)$$

เนื่องจากการคำนวณหาค่าพิกัด UTM มีความยุ่งยาก ดังนั้นในการทำงานจริงจะใช้โปรแกรม คำนวณ



### ดัชนีประสิทธิภาพการจราจร (Traffic Performance Index – TPI)<sup>[1]</sup>

TPI เป็นตัวชี้วัดที่ใช้ในการประเมินระดับความหนาแน่นของการจราจรในพื้นที่เมือง โดยคำนวณจากอัตราส่วน ระหว่างความเร็วเฉลี่ยของยานพาหนะ กับความเร็วในสภาวะการจราจรที่ไหลลื่น (Free-Flow Speed)

$$TPI = \frac{v_a}{v_f} \quad (2.4)$$

โดยที่

$v_a$  คือ ความเร็วเฉลี่ยของกลุ่มยานพาหนะในช่วงเวลาที่กำหนด

$v_f$  คือ ความเร็วสถานะที่การจราจรที่ไหลลื่น (Free Flow Speed) ซึ่งสามารถหาได้จากความเร็วเฉลี่ยของกลุ่มยานพาหนะในช่วงที่มีเวลาเบาบาง หรือเปอร์เซ็นต์ไทม์ที่ 85

### การวิเคราะห์การถดถอยเชิงเส้น (Linear Regression Analysis)

การถดถอยเชิงเส้น (Linear Regression) เป็นวิธีทางสถิติที่ใช้สร้างความสัมพันธ์ระหว่างตัวแปรต้น  $x$  และตัวแปรตาม  $y$  โดยสมมติว่า  $y$  สามารถประมาณค่าจาก  $x$  ด้วยสมการเชิงเส้น

$$y = ax + b \quad (2.5)$$

โดยที่

$a$  คือ ความชันของเส้นตรง

$b$  คือ ค่าตัดแกน  $y$  (y-intercept)

ซึ่งสามารถหาค่าได้จากสูตรดังนี้

$$a = \frac{\sum_{i=1}^n (x_i - \bar{x})(y_i - \bar{y})}{\sum_{i=1}^n (x_i - \bar{x})^2} \quad (2.6)$$

$$b = \bar{y} - a\bar{x} \quad (2.7)$$

โดยที่

$x_i, y_i$  คือ ค่าของข้อมูลแต่ละจุด

$\bar{x}, \bar{y}$  คือ ค่าเฉลี่ยของ  $x$  และ  $y$  ตามลำดับ

$n$  คือ จำนวนข้อมูลทั้งหมด

### ทฤษฎีบทของเบย์ (Bayes' Theorem)<sup>[4]</sup>

(Hayes, 2024)

ทฤษฎีบทของเบย์ (Bayes' Theorem) เป็นกรอบแนวคิดพื้นฐานทางสถิติที่ใช้สำหรับปรับปรุงค่าความน่าจะเป็นของสมมติฐานเมื่อมีข้อมูลใหม่เข้ามา โดยอาศัยการคำนวณแบบมีเงื่อนไข ซึ่งเหมาะสมอย่างยิ่งสำหรับการวิเคราะห์ข้อมูลที่มีการเปลี่ยนแปลงต่อเนื่องตามเวลา เช่น ข้อมูลจราจร

$$P(H|D) = \frac{P(D|H) \times P(H)}{P(D)} \quad (2.7)$$

โดยที่

$H$  คือ สมมติฐาน เช่น “TPI ณ เวลาที่สนใจมีค่าสูง”

$D$  คือ ข้อมูลที่สังเกตได้ เช่น ค่า TPI เวลาที่ผ่านมา

$P(H)$  คือ ค่าความน่าจะเป็นเดิม (Prior)

$P(D|H)$  คือ ความน่าจะเป็นที่จะได้ข้อมูล  $D$  เมื่อทราบว่า  $H$  เป็นจริง (Likelihood)

$P(D)$  คือ ความน่าจะเป็นของข้อมูล  $D$  ทั้งหมด (Evidence)

$P(H|D)$  คือ ค่าความน่าจะเป็นใหม่เมื่อมีข้อมูล  $D$  แล้ว (Posterior)

เมื่อต้องการประเมินค่าคาดหวัง (Expected Value) ของตัวแปรที่สนใจ เช่น TPI ในช่วงเวลาที่สนใจสามารถคำนวณได้จากสมการ

$$E[H|D] = \sum_{i=1}^n H_i \times P(H_i|D) \quad (2.8)$$

### การอนุมานแบบเบย์เชิงต่อเนื่อง (Continuous Bayesian Inference)<sup>[6]</sup>

ทฤษฎีของเบย์สามารถขยายจากกรณีแบบไม่ต่อเนื่องไปสู่กรณีที่ค่าความน่าจะเป็นของสมมติฐานหรือค่าที่ต้องการทำนายเป็นตัวแปรต่อเนื่องได้ โดยในกรณีนี้ การคำนวณค่าความน่าจะเป็นจะใช้ ฟังก์ชันความหนาแน่น (probability density function) และการอินทิเกรตแทนการบวกค่าความน่าจะเป็นแบบจำกัด

$$P(v|x_{obs}) = \frac{P(x_{obs}|v) \times P(v)}{\int P(x_{obs}|v) \times P(v) dv} \quad (2.9)$$

โดยที่

$v$  คือ ค่าที่ต้องการพยากรณ์ (เช่น TPI เวลา 17:00)

$x_{obs}$  คือ ข้อมูลที่สังเกตได้ เช่น TPI เวลา 16:00

$P(v)$  คือ prior distribution ของ  $v$

$P(x_{obs}|v)$  คือ likelihood

$\int P(x_{obs}|v) \times P(v) dv$  คือ การอินทิเกรตเพื่อหาค่าความน่าจะเป็นรวม (evidence)

$P(v|x_{obs})$  คือ posterior distribution

เมื่อได้ posterior distribution แล้ว ค่าสรุปที่สนใจสามารถคำนวณได้ เช่น ค่าคาดหวัง (Expected Value)

$$E[v|x_{obs}] = \int v \times P(v|x_{obs}) dv \quad (2.10)$$

ในงานวิจัยนี้ ได้ประยุกต์ใช้ทฤษฎีเบย์ในกรณีต่อเนื่อง เพื่อสร้างแบบจำลองการพยากรณ์ค่าดัชนี TPI ในช่วงเวลาถัดไป โดยอาศัยข้อมูลที่สังเกตได้จากช่วงเวลาก่อนหน้า พร้อมปรับค่าความเชื่อเดิม (prior) ตามข้อมูลจากอดีต ซึ่งทำให้สามารถพยากรณ์ผลลัพธ์ในลักษณะเป็นช่วงความน่าจะเป็นได้อย่างแม่นยำและยืดหยุ่นมากขึ้น

### การประมาณความหนาแน่นด้วย Kernel (Kernel Density Estimation – KDE)<sup>[6]</sup>

Kernel Density Estimation (KDE) เป็นเทคนิคทางสถิติเชิงไม่พารามิเตอร์ที่ใช้ในการประมาณฟังก์ชันความหนาแน่นของข้อมูลต่อเนื่อง โดยไม่ต้องสมมติรูปร่างของการแจกแจงล่วงหน้า ซึ่งเหมาะสมอย่างยิ่งสำหรับใช้สร้างฟังก์ชัน Prior ในแบบจำลอง Bayesian จากชุดข้อมูลจริง

$$P(v) = \frac{1}{nh} \sum_{i=1}^n K(u) \left( \frac{v - y_i}{h} \right) \quad (2.11)$$

โดยที่

$P(v)$  คือ โอกาสที่จะได้ค่า TPI เวลา 17:00 เป็น  $v$

$y_i$  คือ ข้อมูลจริง เช่น ค่า TPI เวลา 17:00 จากวันต่าง ๆ

$n$  คือ จำนวนข้อมูลทั้งหมด เช่น จำนวนวัน

$h$  คือ ค่าความกว้างแถบ (bandwidth) ซึ่งควบคุมระดับความเรียบของกราฟความหนาแน่น

$K(u)$  คือ Kernel function ที่ใช้ในการประมาณค่า เช่น Gaussian kernel

$v$  คือ ค่าที่ต้องการประเมินความน่าจะเป็น

สำหรับ Gaussian kernel ที่ใช้ในงานนี้ จะมีรูปสมการเป็น

$$K(u) = \frac{1}{\sqrt{2\pi}} e^{-\frac{1}{2}u^2} \quad (2.12)$$

โดยที่

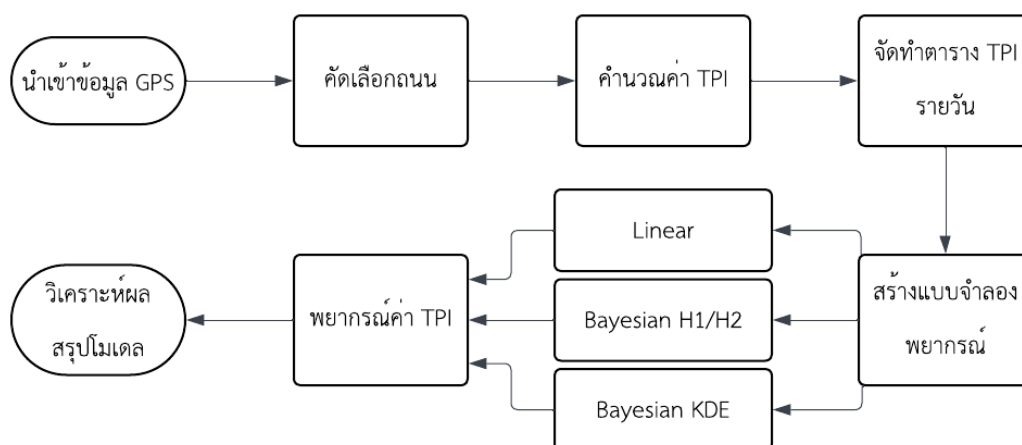
$u = \frac{v - y_i}{h}$  คือ เป็นระยะเชิงมาตรฐานระหว่างค่าที่สนใจ  $v$  และข้อมูล  $y_i$

ในงานวิจัยนี้ ได้ใช้ KDE สำหรับประมาณ Prior distribution ของค่า TPI เวลา 17:00 โดยอิงจากข้อมูลจริงที่บันทึกไว้ในอดีต เพื่อใช้ในกระบวนการคำนวณแบบจำลองเบย์เชิงต่อเนื่อง (ตามหัวข้อก่อนหน้า)

### บทที่ 3

#### ระเบียบวิธีวิจัย

เพื่อให้เห็นภาพรวมของกระบวนการดำเนินงานวิจัย แผนภาพที่แสดงด้านล่างแสดงขั้นตอนหลักในการดำเนินการ ตั้งแต่การเตรียมข้อมูลจนถึงการพัฒนาแบบจำลองพยากรณ์และวิเคราะห์ผล



#### 3.1 เตรียมข้อมูล

##### 3.1.1 ดาวน์โหลดข้อมูลจากฐานข้อมูล

ขอข้อมูลการติดตาม GPS จาก บริษัท Grab Thailand ปีพ.ศ.2562 จุดทั้งหมด 1,346,029,571 จุด เป็นจุด GPS ที่กระจายตัวอยู่ทั่วกรุงเทพมหานคร

##### 3.1.2 คัดเลือกถนน

ทำการกรองจุดข้อมูลให้เหลือเฉพาะพื้นที่ศึกษา คือ ถนนพระราม4 บริเวณหน้าสถานีรถไฟกรุงเทพฯ จนถึงจุดตัดบริเวณแยกพระโขนง โดยการใช้การเขียนโปรแกรมผ่านแพลตฟอร์ม Jupyter นำข้อมูลเข้าแพลตฟอร์ม โดยใช้ Library "Pyspark"

- 1) ทำการกรองข้อมูลให้อยู่ในช่วงวันที่ต้องการทำการศึกษา คือ 1 มิถุนายน – 31 สิงหาคม
- 2) ทำการเลือก wayid ที่อยู่ในถนนพระราม4 โดยวิธีการติกรอบถนนที่ละน้อยๆ ใน Google map และนำค่า Latitude และ Longitude จุดเริ่มและจุดจบของกรอบนั้นมาทำการเขียนโปรแกรมเพื่อกรองให้เหลือ wayid ที่ต้องการ และนำ wayid ที่กรองได้ทั้งหมดมาต่อกันจนเป็นถนนทั้งเส้นพระราม 4
- 3) ทำการเลือกข้อมูลคอลัมน์ที่จำเป็นต้องใช้ คือ driverid, wayids, date, time, speed ,projectedlat และ projectedlng

- 4) แปลงข้อมูลที่เลือกมาจาก Library "Pyspark" ไป Library "Pandas" เพื่อการคำนวณได้ง่ายยิ่งขึ้น

	driverid	wayids	date	time	speed	projectedlat	projectedlng
0	826c00d50cf541f35445b3eb178b641662c89a0982ce6b...	459492912	2019-06-02	05:09:49	10.17	13.721782	100.555492
1	04da1c82aae938e9c9c69ed25f99ea642dab1b000b7fdb...	459492912	2019-06-02	05:27:27	24.35	13.721287	100.556687
2	04da1c82aae938e9c9c69ed25f99ea642dab1b000b7fdb...	459492912	2019-06-02	05:27:31	18.10	13.720963	100.557425
3	2ab60462ca330a9fd443061718cc46618fdf19138dda2e...	459492912	2019-06-02	05:42:13	11.99	13.721786	100.555481
4	2ab60462ca330a9fd443061718cc46618fdf19138dda2e...	459492912	2019-06-02	05:42:17	8.22	13.721646	100.555815
...	...	...	...	...	...	...	...
3827	2c7d2cc5df76d66153d43a814e879bd95222aec8a483c6...	459492912	2019-06-12	05:58:14	14.25	13.721651	100.555803
3828	2c7d2cc5df76d66153d43a814e879bd95222aec8a483c6...	459492912	2019-06-12	05:58:18	14.45	13.721444	100.556305
3829	2c7d2cc5df76d66153d43a814e879bd95222aec8a483c6...	459492912	2019-06-12	05:58:22	13.20	13.721254	100.556767
3830	405aaf5dce2edd067ee3ef48b793dfe845cb86633dfdf0...	459492912	2019-06-27	05:31:22	15.50	13.721796	100.555459
3831	405aaf5dce2edd067ee3ef48b793dfe845cb86633dfdf0...	459492912	2019-06-27	05:31:26	16.83	13.721565	100.556009

รูปที่ 3.1 ตัวอย่างข้อมูลที่ดาวน์โหลดมาจากรฐานข้อมูล โดยเลือกเฉพาะหลักที่จำเป็นในการวิเคราะห์ข้อมูล คือ driverid, wayids, date, time, speed ,projectedlat และ projectedlng

### 3.2 วิเคราะห์ข้อมูลเบื้องต้น

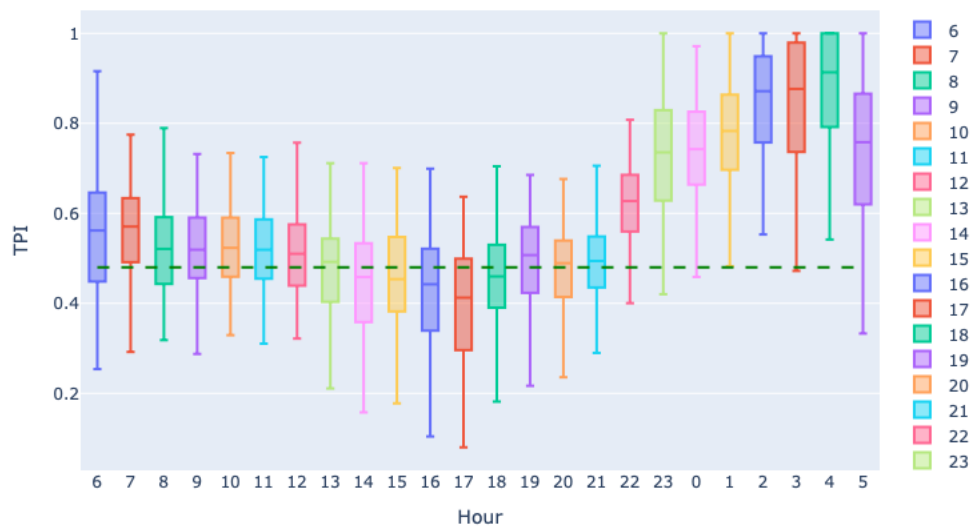
#### 3.2.1 วิเคราะห์ข้อมูลเพื่อหาค่า TPI

แยกการดูข้อมูลบริเวณที่เป็นพื้นที่การศึกษาเป็นเวลาทุกๆ 1 ชั่วโมง ในทุกๆวัน เพื่อดูการเกิดรถติดจากการหาค่า TPI ดังนี้

- 1) หาค่า  $V_a$  คือค่าความเร็วเฉลี่ยของรถในช่วงเวลาที่เราส่งใจ หาได้จาก การสุ่มกลุ่มตัวอย่างจากข้อมูลมากกว่า 30 ตัวอย่าง แล้วหาค่าเฉลี่ยความเร็วของรถโดยการแปลงพิกัด Latitude และ Longitude ให้เป็นระบบพิกัด UTM ตามในหัวข้อทฤษฎีที่เกี่ยวข้อง<sup>[3]</sup> และหาความห่างจุดปลาย-จุดต้น เทียบกับเวลา
- 2) ค่า  $V_f$  คือค่าความเร็วของรถในช่วงเวลาที่รถสามารถเคลื่อนตัวได้คล่องตัว<sup>[5]</sup> (Forbes et al., 2012) ในที่นี้จะยกกลุ่มตัวอย่าง ช่วงเวลา 1:00 - 2:00 น. จากข้อมูล มากกว่า 30 ตัวอย่าง แล้วหาค่าความเร็วของรถแต่ละคัน โดยการแปลงพิกัด Latitude และ Longitude ให้เป็นระบบพิกัด UTM ตามในหัวข้อทฤษฎีที่เกี่ยวข้อง และหาความเร็วจุดปลาย-จุดต้น เทียบกับเวลา แล้วมาทำการเรียงลำดับจากมากไปน้อยเพื่อหาข้อมูลเปอร์เซ็นไทล์ที่ 85 มาเป็นค่า  $V_f$ <sup>[7]</sup> (Finkelstein, 2016)

- 3) หาความสัมพันธ์ของข้อมูลโดยการพลอตกราฟทุกๆ 1 ชั่วโมง ในทุกๆวัน แบ่งออกเป็น 24 ชั่วโมง เพื่อดูแนวโน้มค่า TPI ของแต่ละชั่วโมง และทำการตัดข้อมูลที่เป็นค่าโดด

Box Plot of TPI for Each Hour 3M



รูปที่ 3.2 กราฟค่า TPI ของแต่ละชั่วโมงของทุกวัน ในรูปแบบ Boxplot บริเวณประมาณ 100 เมตร ก่อนถึงแยกซอยบ้านกล้วยใต้ ถนนฝั่งมุ่งหน้าสู่แยกพระโขนง

- 4) สร้างตาราง TPI ในทุกชั่วโมงในทุกวัน และนำไปใช้งานต่อ

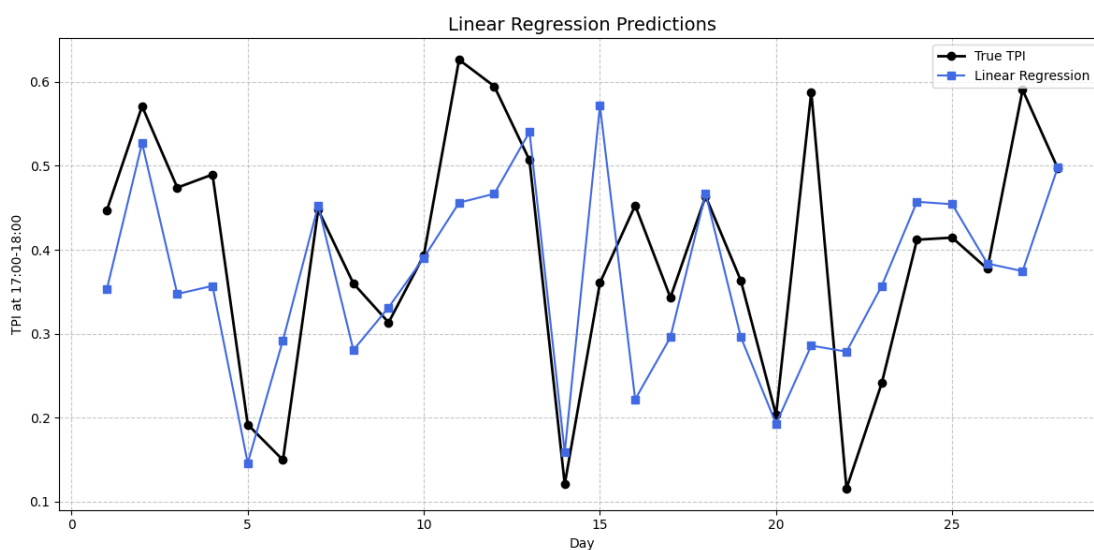
hour	0	1	2	3	4	5	6	7	8	9	...	14	15	16	17	18	19	20
date																		
2019-06-01	NaN	NaN	NaN	NaN	NaN	NaN	0.643059	0.676740	0.733566	0.705351	...	0.711398	0.647175	0.583293	0.536565	0.497574	0.637894	0.664751
2019-06-02	0.838857	0.805945	1.000000	0.948304	0.907798	0.620463	0.563938	0.646087	0.586426	0.443025	...	0.517674	0.553976	0.517437	0.476854	0.637440	0.603415	0.535578
2019-06-03	0.708940	0.753129	0.608335	0.815792	1.000000	0.577042	0.472742	0.498413	0.415910	0.446991	...	0.492745	0.429344	0.474278	0.500685	0.452180	0.463869	0.411601
2019-06-04	0.727689	0.809786	1.000000	0.623470	0.935100	0.824306	0.465484	0.604144	0.442282	0.394583	...	0.590306	0.538050	0.534511	0.463748	0.447126	0.552142	0.464541
2019-06-05	0.517825	0.482381	0.581289	1.000000	0.541808	0.332906	0.370148	0.397321	0.385117	0.342084	...	0.356251	0.271714	0.316044	0.363302	0.361825	0.321166	0.329947
...	...	...	...	...	...	...	...	...	...	...	...	...	...	...	...	...	...	...
2019-08-27	0.458603	0.683205	0.578124	1.000000	0.651731	0.530787	0.253843	NaN	0.339623	0.313284	...	0.333766	0.360501	0.347153	0.216733	NaN	NaN	NaN
2019-08-28	0.823666	0.818217	0.910116	0.876474	1.000000	0.850955	0.682458	0.618197	0.657947	0.514029	...	0.235651	0.560646	0.602109	0.256001	0.329834	0.531033	0.676274
2019-08-29	0.741402	0.828867	1.000000	0.824966	0.803989	0.927787	0.530721	0.549004	0.623724	0.605890	...	0.157733	0.177715	0.305762	0.243310	0.430736	0.549001	0.505945
2019-08-30	0.549291	0.535847	NaN	0.484127	1.000000	0.411512	0.327523	0.292328	0.355156	0.291192	...	0.202148	0.261852	0.315611	0.343005	0.397755	0.356381	0.356370
2019-08-31	0.910803	0.646250	1.000000	0.998661	0.840977	0.767785	0.648750	0.681229	0.650572	0.636134	...	0.638680	0.307204	0.235939	0.267421	0.394597	0.652233	0.605313

รูปที่ 3.3 ตัวอย่างค่า TPI แต่ละชั่วโมงของทุกวัน บริเวณประมาณ 100 เมตร ก่อนถึงแยกซอยบ้านกล้วยใต้ ถนนฝั่งมุ่งหน้าสู่แยกพระโขนง

### 3.3 ปรับปรุงแบบจำลอง

#### 3.3.1 การทำนายค่า TPI ด้วยวิธีการ Linear Regression

- 1) นำข้อมูลที่ได้จากตารางค่า TPI มาวิเคราะห์ โดยเลือกเวลาที่ข้อมูล TPI มีแนวโน้มค่าน้อยที่สุด คือ ในช่วงเวลา 16.00 – 17.00 น. มาทำนายค่า TPI ในเวลา 17.00 – 18.00 น.
- 2) รวบรวมข้อมูลค่า TPI ที่เกิดขึ้นจริงในช่วงเวลาดังกล่าว และทำการแบ่งออกเป็น 2 ชุดข้อมูล คือ ชุดข้อมูลสำหรับการฝึกแบบจำลอง (training set) 70% และ ชุดข้อมูลสำหรับการทดสอบ (test set) 30%
- 3) นำข้อมูล training set มาหาความสัมพันธ์เชิงเส้นของค่า TPI ในช่วงเวลา 16.00 – 17.00 น. กับเวลาที่จะเกิดขึ้น 16.00 – 17.00 น.
- 4) เมื่อได้แบบจำลองแล้ว นำแบบจำลองนั้นมาทดสอบกับชุดข้อมูลทดสอบในช่วงเวลา 17.00 – 18.00 น. ค่าที่ได้จากการทำนายจะถูกนำมาเปรียบเทียบกับค่าจริงเพื่อประเมินค่าคลาดเคลื่อนของแบบจำลองนี้



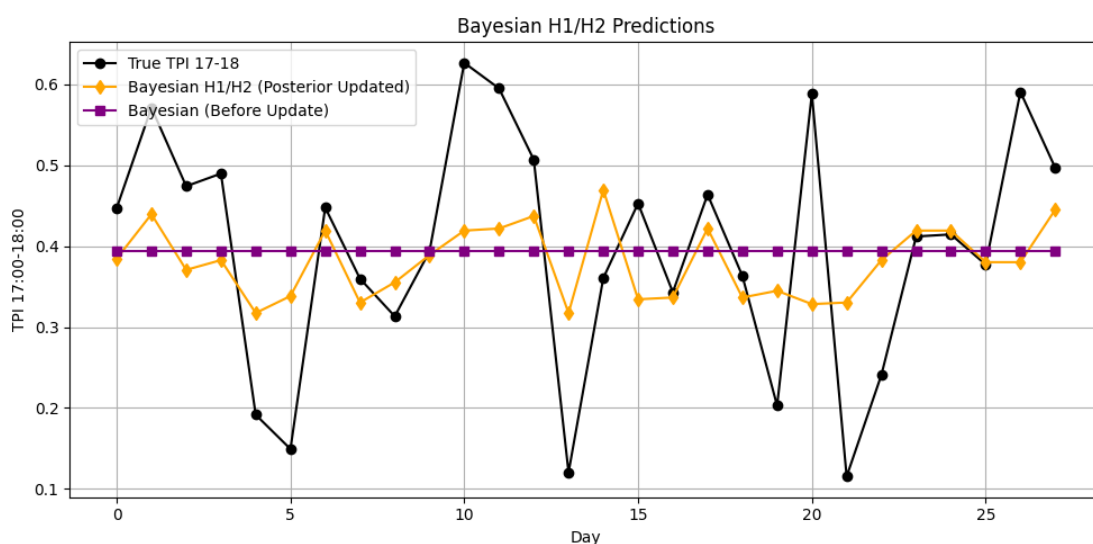
รูปที่ 3.4 กราฟการเปรียบเทียบระหว่างค่า True TPI และ ค่าที่ทำนายจากแบบจำลองด้วยวิธีการ Linear Regression



### 3.3.2 การทำนายค่า TPI ด้วยวิธีการ Bayesian ( $H_1/H_2$ )

นำข้อมูลที่ได้จากตารางค่า TPI มาวิเคราะห์ โดยเลือกเวลาที่ข้อมูล TPI มีแนวโน้มค่าน้อยที่สุด คือ ในช่วงเวลา 16.00 – 17.00 น. มาทำนายค่า TPI ในเวลา 17.00 – 18.00 น. และแบ่งค่าออกเป็น 2 กลุ่ม จากการกำหนดค่าเกณฑ์ล่วงหน้า คือ ค่า TPI = 0.48<sup>[1]</sup> โดยที่ค่า TPI มากกว่าเกณฑ์ดังกล่าวจะจัดอยู่ในค่า  $H_1$  ซึ่งหมายถึงช่วงเวลาที่แนวโน้มการจราจรหนาแน่น และ  $H_2$  คือ ค่าที่มี TPI น้อยกว่าค่าดังกล่าว ซึ่งหมายถึงช่วงเวลาที่แนวโน้มการจราจรไม่หนาแน่น

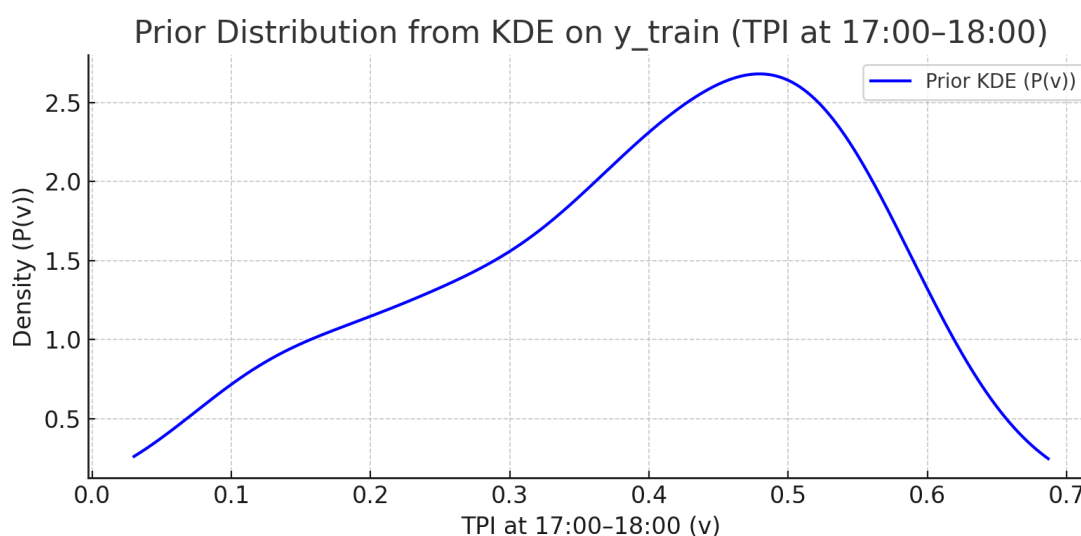
- 1) คำนวณค่าความน่าจะเป็นล่วงหน้าที่เราเรียกว่า prior probability และค่าเฉลี่ยของผลลัพธ์ในแต่ละกลุ่ม เพื่อใช้ในการทำนาย
- 2) ทำการวัดค่าความใกล้เคียงของข้อมูลใหม่ โดยใช้หลักการที่เราเรียกว่า bandwidth คือการนำข้อมูลใหม่ใส่ลงไป และข้อมูลใหม่นั้นมีค่าอยู่ในช่วงที่กำหนดอยู่ในช่วงกลุ่มใด จะถือว่าเกิดจากกลุ่มนั้นมากกว่า
- 3) ค่าความจะเป็นที่ได้จะถูกนำมาคำนวณในรูปแบบ weight average โดยใช้กฎของ Bayes เพื่อทำนายค่าความน่าจะเป็นแบบมีเงื่อนไข ที่เรียกว่า posterior probability ของแต่ละกลุ่ม และนำค่าคาดหวังของผลลัพธ์ในแต่ละกลุ่มมารวมกันเพื่อทำนายค่าสุดท้าย
- 4) ในกรณีที่ข้อมูลใหม่นั้นไม่ใกล้เคียงกับข้อมูลใดเลยจะใช้ข้อมูลที่ใกล้เคียงที่สุดในชุดข้อมูลเดิมมาทำนายแทน เพื่อป้องกันความผิดพลาด



รูปที่ 3.5 กราฟการเปรียบเทียบระหว่างค่า True TPI และ ค่าที่ทำนายจากแบบจำลองด้วยวิธีการ Bayesian ( $H_1/H_2$ )

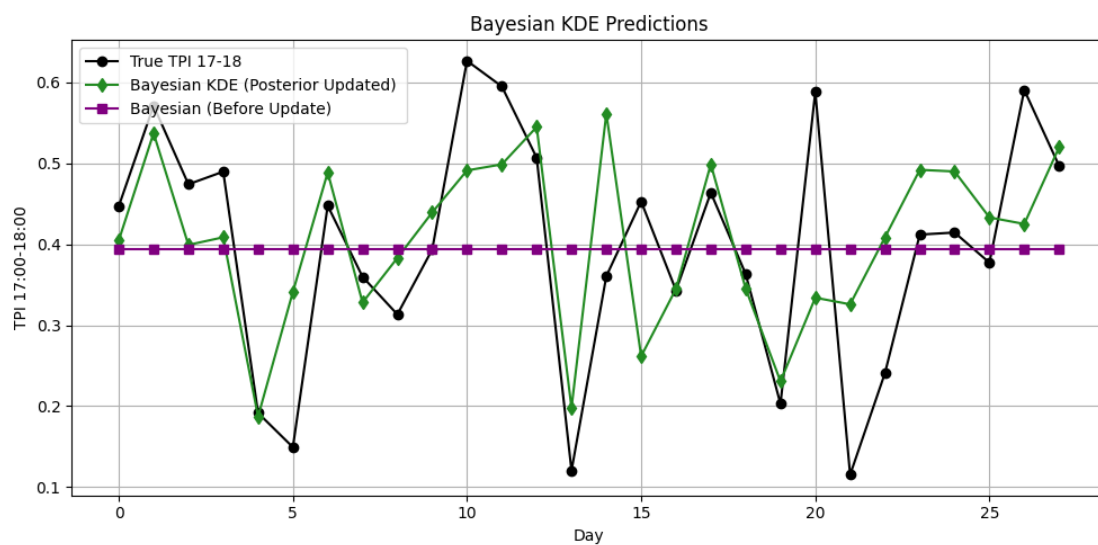
### 3.3.3 การทำนายค่า TPI ด้วยวิธีการ Bayesian Continuous Distribution KDE

- 1) ประมาณค่า Prior Distribution โดยการนำข้อมูลผลลัพธ์ค่า TPI ในช่วงเวลา 17.00 – 18.00 น. มาสร้าง Probability Continuous Distribution โดยใช้ Kernel Density Estimation<sup>[6]</sup> ซึ่งจะทำให้การกระจายตัวของข้อมูลเป็นดังรูป



รูปที่ 3.6 กราฟการแจกแจงค่า TPI ของ Prior ให้เป็น Continuous Distribution โดยใช้วิธี Kernel Density Estimation

- 2) กำหนดค่า Likelihood Function โดยพิจารณาจากค่าแนวโน้ม TPI ที่จะเกิดขึ้น โดยใช้ฟังก์ชันความหนาแน่นปกติ และคำนวณส่วนเบี่ยงเบนระหว่างข้อมูลในช่วงเวลา 16.00 – 17.00 น. กับ 17.00 – 18.00 น. ของ training set
- 3) คำนวณค่า Posterior จากการนำ Prior Distribution คูณกับ Likelihood Function แล้วทำการ normalization เพื่อให้ได้ค่าความน่าจะเป็นที่แท้จริง จากนั้นจะได้ค่าทำนาย TPI จากค่าคาดหวังที่เรียกว่า Expected value บนฟังก์ชันของ Posterior



รูปที่ 3.7 กราฟการเปรียบเทียบระหว่างค่า True TPI และ ค่าที่ทำนายจากแบบจำลองด้วยวิธีการ Bayesian Continuous Distribution KDE

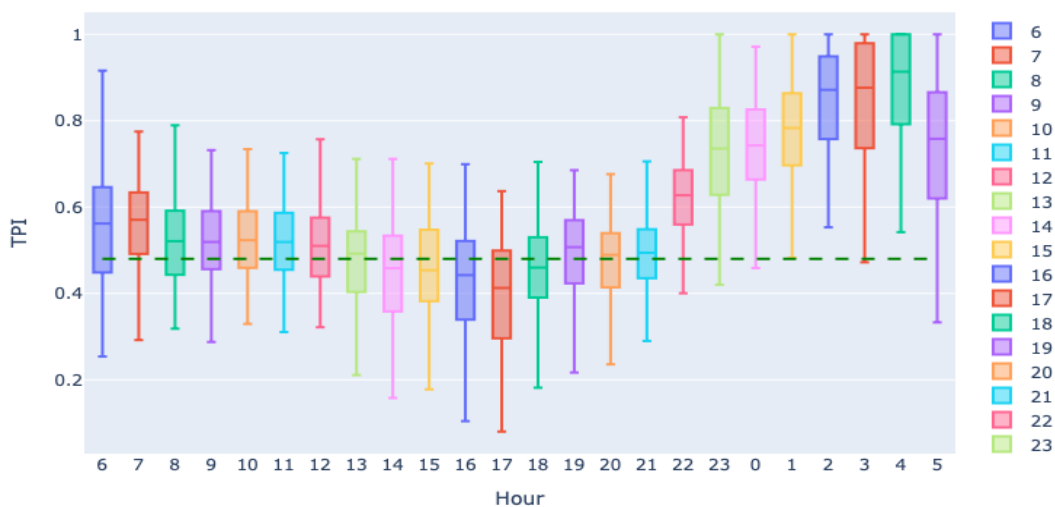
## บทที่ 4

### ผลการดำเนินงานวิจัย

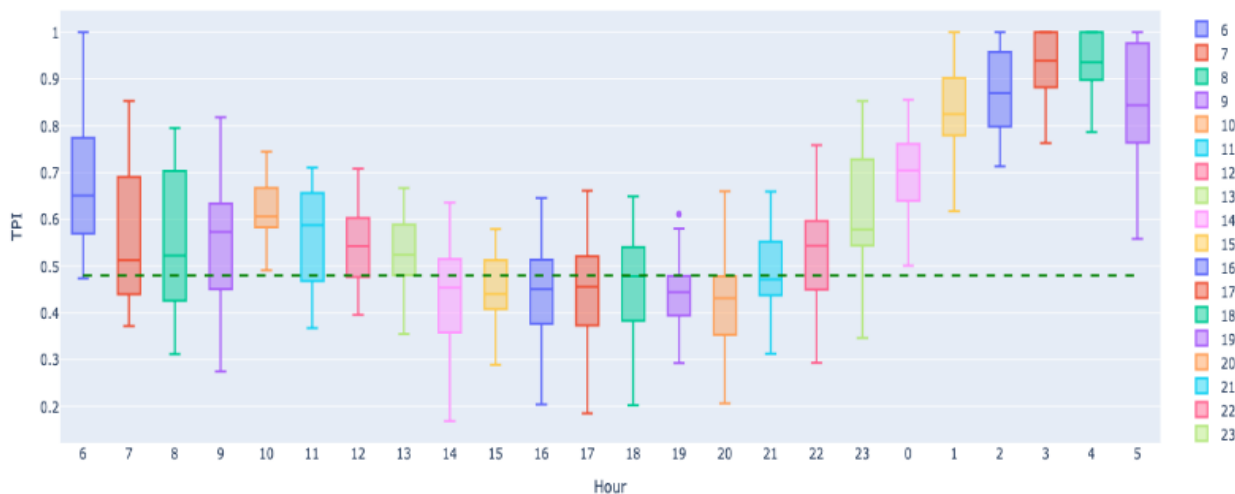
หลังจากการใช้แบบจำลองทั้ง 3 แบบจำลองในการวิเคราะห์สภาพจราจรจากการทำนายค่า TPI จากข้อมูลค่าการจราจรจากชั่วโมงก่อนหน้า เพื่อทำนายค่าการจราจรของชั่วโมงต่อมา กล่าวคือ ในงานวิจัยนี้ใช้ข้อมูลในช่วงเวลา 16.00 – 17.00 น. มาทำนาย ค่าการจราจรในช่วงเวลา 17.00 – 18.00 น. ได้ผลการวิจัยดังนี้

#### 4.1 ค่า TPI ในแต่ละชั่วโมงของทุกวัน

จากการคำนวณ ค่า TPI ของทุกชั่วโมงในแต่ละวัน เพื่อวิเคราะห์ว่า ในชั่วโมงใดบ้างที่มีค่า TPI ต่ำกว่าเกณฑ์  $TPI = 0.48$  เพื่อดูแลแนวโน้มและวิเคราะห์ ชั่วโมงที่มีสภาพจราจรติดขัดรุนแรง ได้ทำการวิเคราะห์ใน 2 ช่วงถนน คือ บริเวณประมาณ 100 เมตร ก่อนถึงแยกซอยบ้านกล้วยใต้ ถนนฝั่งมุ่งหน้าสู่แยกพระโขนง และระหว่างแยกบ้านกล้วยใต้ ถึง แยกกล้วยน้ำไท ได้ผลเป็นกราฟดังรูป



รูปที่ 4.1 กราฟการ Box plot เปรียบเทียบระหว่างค่า TPI ในแต่ละชั่วโมง ของทุกวัน 1 มิถุนายน – 31 สิงหาคม พ.ศ. 2562 บริเวณประมาณ 100 เมตร ก่อนถึงแยกซอยบ้านกล้วยใต้



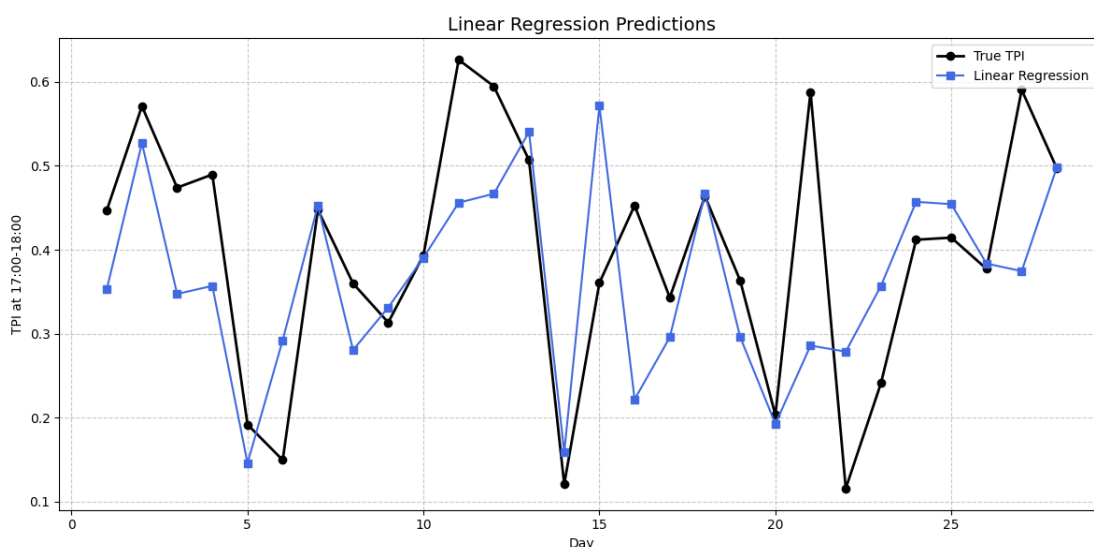
รูปที่ 4.2 กราฟการ Box plot เปรียบเทียบระหว่างค่า TPI ในแต่ละชั่วโมง ของทุกวัน 1 มิถุนายน – 31 สิงหาคม พ.ศ. 2562 ระหว่างแยกบ้านกล้วยใต้ ถึง แยกกล้วยน้ำไ้

จากรูปที่ 4.1 และรูปที่ 4.2 แสดงค่าเปรียบเทียบค่า TPI ในถนน 2 จุด เส้นประสีเขียวแทนจุดที่  $TPI = 0.48$  จะเห็นได้ว่า รูปที่ 4.1 จาก 06.00 – 07.00 น. ถึง 17.00 – 18.00 น. ค่า TPI มีแนวโน้มลดลงเรื่อยๆ จากค่ากลาง TPI ที่ทำได้ในเวลา 06.00 – 07.00 น. คือ 0.56 จนกระทั่งถึง ค่ากลาง TPI คือ 0.41 ในเวลา 17.00 – 18.00 น. และแนวโน้ม TPI มีค่าเพิ่มขึ้นเรื่อยๆอีกครั้ง หลังจากเวลา 17.00 – 18.00 น. จนกระทั่งค่ากลาง TPI มีค่าสูงสุดถึง 0.91 ในเวลา 04.00 – 05.00 น. ในขณะเดียวกัน รูปที่ 4.2 มีแนวโน้มของค่ากลาง TPI ในลักษณะที่สอดคล้องกับรูปที่ 4.1 โดยมีค่ากลาง TPI ที่ต่ำที่สุดอยู่ที่ 0.45 ในเวลา 17.00 – 18.00 น. และมีค่ากลาง TPI มากที่สุดอยู่ที่ 0.93 ในเวลา 04.00 – 05.00 น. ซึ่งค่อนข้างชัดเจนว่า ช่วงเวลาที่มีการจราจรติดขัดรุนแรงมากที่สุด จะอยู่ในช่วง 17.00 – 18.00 น. ในทางกลับกัน 04.00 – 05.00 น.จะเป็นช่วงที่มีการจราจรคล่องตัวมากที่สุดของวัน อย่างไรก็ตาม กราฟทั้ง 2 กราฟนี้ มีช่วงพิสัยที่ค่อนข้างกว้าง ที่เป็นเช่นนั้นเพราะในแต่ละวันมีเงื่อนไขที่จะทำให้การจราจรแตกต่างกันออกไป เช่น อาจจะมีการเกิดอุบัติเหตุ หรือฝนตก หรือบางวันถนนโล่งมาก จึงทำให้จราจรเคลื่อนตัวได้อย่างคล่องตัว

## 4.2 แบบจำลองที่ใช้ในงานวิจัย

จากการวิธีดำเนินงานวิจัย ในวิจัยนี้ใช้แบบจำลองในการทำนายค่า TPI ได้แก่วิธีการดังนี้

### 4.2.1 วิธีการ Linear Regression



รูปที่ 4.3 กราฟการเปรียบเทียบระหว่างค่า TPI จริง และค่าที่ทำนายจากแบบจำลองด้วยวิธีการ Linear Regression ทั้งหมด 27 วัน ของถนนประมาณ 100 เมตร ก่อนถึงแยกซอยบ้านกล้วยใต้

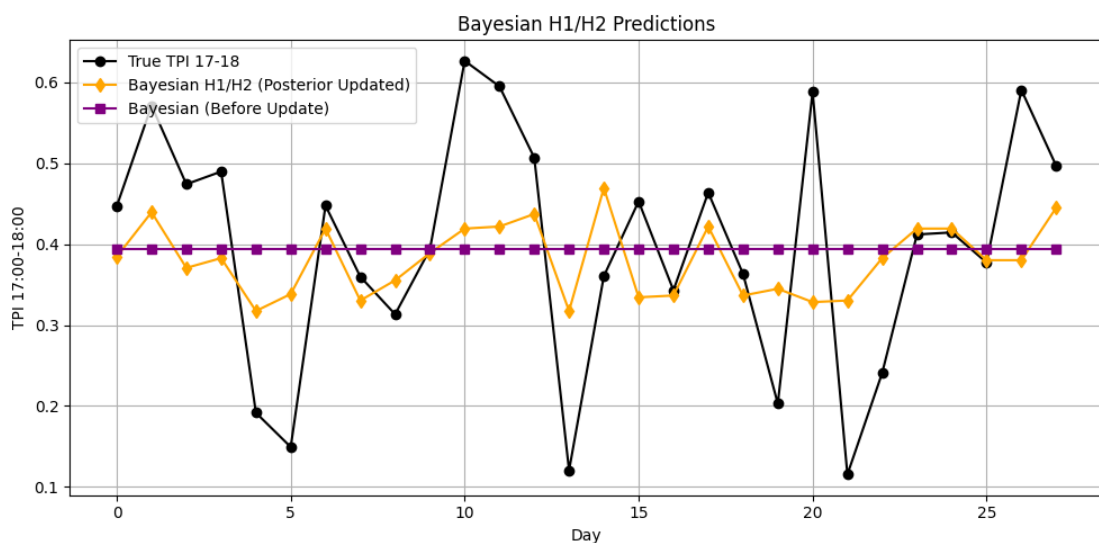
จากรูปที่ 4.3 แสดงการเปรียบเทียบค่าดัชนี TPI ที่ได้จากข้อมูลจริงกับค่าที่ได้จากการทำนายโดยใช้แบบจำลอง Linear Regression ซึ่งพิจารณาค่าดัชนี TPI ในช่วงเวลา 17:00–18:00 ของแต่ละวันตลอดช่วงเวลาที่ทำการศึกษา จะเห็นได้ว่าแบบจำลองสามารถติดตามแนวโน้มโดยรวมของข้อมูลจริงได้ในระดับหนึ่ง โดยเฉพาะในช่วงวันที่ 13–14 ซึ่งค่าดัชนี TPI ลดลงอย่างต่อเนื่อง แบบจำลองสามารถสะท้อนทิศทางการเปลี่ยนแปลงในช่วงเวลาดังกล่าวได้อย่างแม่นยำ นอกจากนี้ ยังพบว่าในบางช่วง แบบจำลองสามารถติดตามการเปลี่ยนแปลงของข้อมูลที่มีลักษณะต่อเนื่องได้ดี เช่น ช่วงวันที่ 9–11 และ 24–26 ซึ่งแสดงให้เห็นว่าโมเดลสามารถเรียนรู้แนวโน้มของข้อมูลในช่วงเวลาที่มีความเสถียรได้อย่างมีประสิทธิภาพ

อย่างไรก็ตาม แบบจำลอง Linear Regression ยังมีข้อจำกัดในช่วงที่ข้อมูลมีความผันผวนสูงหรือเปลี่ยนแปลงอย่างเฉียบพลัน เช่น วันที่ 4, 12, 19 และ 21 โดยค่าที่แบบจำลองทำนายได้ในช่วงเวลาดังกล่าวมักมีลักษณะเรียบหรือโน้มไปทางค่าเฉลี่ยมากกว่าค่าจริง ทำให้ไม่สามารถสะท้อนลักษณะการกระโดดของข้อมูลได้อย่างแม่นยำ ซึ่งเป็นข้อจำกัดโดยธรรมชาติของแบบจำลองเชิงเส้นที่มีสมมติฐานว่าความสัมพันธ์ระหว่างตัวแปรอิสระและตัวแปรตามเป็นแบบเชิงเส้นตลอดช่วง

แม้จะมีข้อจำกัดดังกล่าว แต่ในหลายช่วงเวลายังคงพบว่าแนวโน้มของค่าที่ทำนายมีความสอดคล้องกับแนวโน้มของค่าจริง เช่น ช่วงวันที่ 7–10 และ 25–27 โดยลักษณะของกราฟแสดงเส้นแบบจำลองที่วางตัวขนานกับเส้นค่าจริง ซึ่งสะท้อนถึงความสามารถของแบบจำลองในการเรียนรู้ทิศทางของข้อมูลพื้นฐาน แม้อาจไม่สามารถติดตามรายละเอียดความแปรปรวนได้ทั้งหมด

โดยสรุป แบบจำลอง Linear Regression มีความเหมาะสมในการใช้งานกับข้อมูลที่มีแนวโน้มต่อเนื่องและเสถียร แต่ยังไม่สามารถจัดการกับข้อมูลที่มีลักษณะไม่เป็นเชิงเส้นหรือมีความแปรปรวนสูงได้อย่างมีประสิทธิภาพ การนำแบบจำลองนี้ไปใช้งานจริงจึงควรพิจารณาร่วมกับลักษณะของข้อมูลในพื้นที่ และอาจพิจารณาแบบจำลองอื่นเพิ่มเติม หากข้อมูลมีพฤติกรรมที่ซับซ้อนหรือไม่สามารถประมาณด้วยความสัมพันธ์เชิงเส้นได้ดีพอ

#### 4.2.2 วิธีการ Bayesian ( $H_1/H_2$ )



รูปที่ 4.4 กราฟการเปรียบเทียบระหว่างค่า TPI จริง และค่าที่ทำนายจากแบบจำลองด้วยวิธีการ Bayesian ( $H_1/H_2$ ) ทั้งหมด 27 วัน ของถนนประมาณ 100 เมตร ก่อนถึงแยกซอยบ้านกล้วยใต้

จากรูปที่ 4.4 แสดงผลการทำนายค่าดัชนี TPI ในช่วงเวลา 17:00–18:00 น. โดยใช้วิธีการสร้างแบบจำลอง Bayesian ซึ่งในที่นี้ได้แบ่งข้อมูลออกเป็นสองกลุ่ม ได้แก่  $H_1$  และ  $H_2$  โดยอิงจากเกณฑ์ค่า threshold ของ TPI ที่ระดับ 0.48 เพื่อจำแนกระดับความหนาแน่นของการจราจร โดย  $H_1$  แทนค่าที่ต่ำกว่า 0.48 (แสดงถึงความแออัดสูง) และ  $H_2$  สำหรับค่าที่มากกว่าหรือเท่ากับ 0.48 (แสดงถึงสภาพการจราจรที่คล่องตัวมากขึ้น)

ในแบบจำลองนี้ ค่าที่ใช้ในการพยากรณ์ถูกคำนวณจาก expected value ที่ได้จาก posterior probability ของกลุ่ม  $H_1$  และ  $H_2$  ซึ่งเป็นผลจากการอัปเดตค่าความน่าจะเป็นเริ่มต้น (prior แบบจำแนกกลุ่ม – discrete prior) ด้วยข้อมูลจริงในแต่ละวัน โดยไม่ใช้ distribution แบบต่อเนื่อง ผลลัพธ์ที่ได้จะสะท้อนความเชื่อปรับปรุง (updated belief) ตามหลักของ Bayesian inference

ในภาพกราฟ เส้นสีดำแสดงค่าจริงของ TPI ที่ได้จากข้อมูล GPS เส้นสีส้มแสดงค่า expected value ที่ได้จาก posterior probability ที่อัปเดตแล้ว ขณะที่เส้นสีม่วงคือค่า expected value ที่ได้จาก prior probability เพียงอย่างเดียว ซึ่งไม่ปรับเปลี่ยนตามข้อมูลรายวัน แต่สะท้อนค่าเฉลี่ยจากข้อมูลในอดีต (historical data)

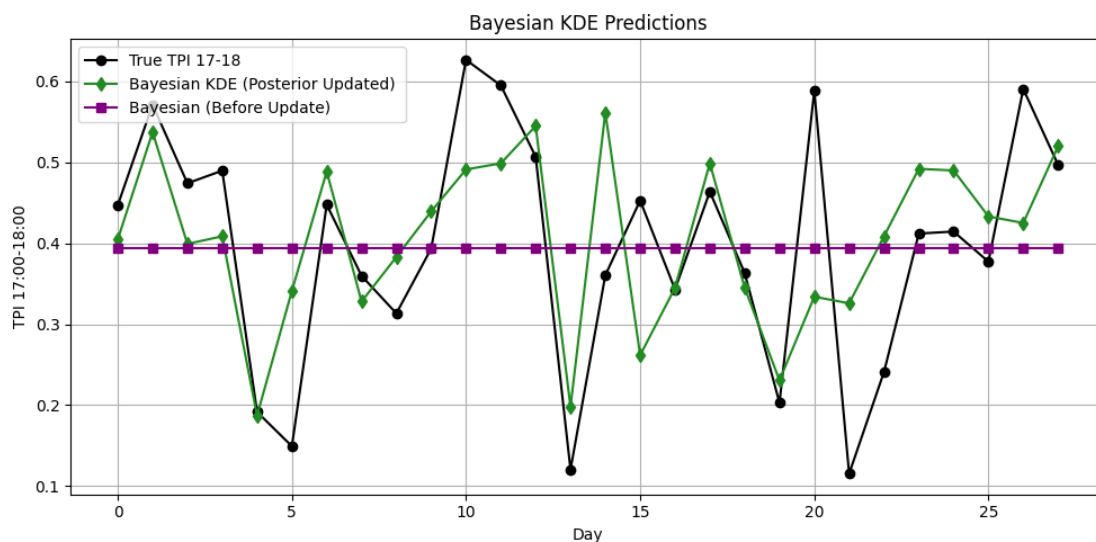
จากผลลัพธ์พบว่า แบบจำลองสามารถทำนายค่า TPI ได้ใกล้เคียงกับค่าจริงในช่วงวันที่ 7, 23, 24 และ 25 โดยค่าที่ได้จาก posterior probability มีแนวโน้มสอดคล้องกับค่าจริงในช่วงวันเหล่านี้อย่างชัดเจน โดยเฉพาะในช่วงที่ข้อมูลมีลักษณะเสถียรและค่าของ TPI อยู่ใกล้เคียงกับค่าเฉลี่ยของกลุ่ม แสดงให้เห็นว่าแบบจำลองมีความเหมาะสมในการพยากรณ์ข้อมูลที่ไม่มีความผันผวนมากนัก

ในทางกลับกัน ช่วงวันที่ 10, 19 และ 22 ซึ่งข้อมูลจริงมีการเปลี่ยนแปลงอย่างฉับพลัน แบบจำลองยังคงให้ค่าทำนายที่ใกล้กับค่าเฉลี่ยของกลุ่มเดิม ซึ่งไม่สามารถสะท้อนความเปลี่ยนแปลงในวันนั้นได้อย่างแม่นยำ แสดงถึงข้อจำกัดของแบบจำลองที่อิงกับกลุ่มแบบคงที่ และไม่สามารถตอบสนองต่อข้อมูลที่เปลี่ยนแปลงเร็วหรือมีลักษณะ time-dependent ได้อย่างเพียงพอ สำหรับเส้น expected value ที่ได้จาก prior probability (กราฟสีม่วง) จะเห็นได้ว่ามีลักษณะคงที่ตลอดเวลา เนื่องจากไม่มีการอัปเดตด้วยข้อมูลใหม่ในแต่ละวัน แม้จะสามารถใช้เป็นค่าความน่าจะเป็นเบื้องต้นจาก historical data ได้ แต่ก็ไม่สามารถสะท้อนพฤติกรรมของข้อมูลจริงในแต่ละวันได้ดีนัก โดยเฉพาะในช่วงที่ข้อมูลเปลี่ยนแปลงจากแนวโน้มเดิม

โดยสรุป แบบจำลอง Bayesian ที่แบ่งกลุ่มตาม threshold นี้มีความเหมาะสมกับข้อมูลที่มีแนวโน้มเสถียร หรืออยู่ในช่วงค่าที่ใกล้เคียงกับค่าเฉลี่ยของกลุ่ม แต่ยังมีข้อจำกัดในการพยากรณ์ข้อมูลที่มีลักษณะเปลี่ยนแปลงอย่างฉับพลัน หรือไม่สามารถจำแนกได้ด้วยการแบ่งกลุ่มเพียงอย่างเดียว การใช้งานในบริบทจริงจึงอาจต้องพิจารณาปรับปรุงโมเดล เช่น ใช้การแบ่งกลุ่มที่ยืดหยุ่นมากขึ้น หรือเพิ่มข้อมูลบริบทอื่นเข้ามาช่วยวิเคราะห์เพื่อเพิ่มความแม่นยำในการพยากรณ์ในอนาคต



#### 4.2.3 วิธีการ Bayesian Continuous Distribution KDE



รูปที่ 4.5 กราฟการเปรียบเทียบระหว่างค่า TPI จริง และค่าที่ทำนายจากแบบจำลองด้วยวิธีการ Bayesian Continuous Distribution KDE ทั้งหมด 27 วัน ของถนนประมาณ 100 เมตร ก่อนถึงแยกซอยบ้านกล้วยใต้

จากรูปที่ 4.5 แสดงผลการทำนายค่าดัชนี TPI ในช่วงเวลา 17:00–18:00 น. โดยใช้แบบจำลอง Bayesian ซึ่งในกรณีนี้ได้สร้าง prior distribution ของค่าที่ต้องการพยากรณ์จาก Kernel Density Estimation (KDE) โดยอิงจากข้อมูลย้อนหลัง (historical data) ทั้งหมด และทำการอัปเดตค่าความน่าจะเป็นใหม่ (posterior distribution) เมื่อมีข้อมูล TPI จากช่วงเวลาก่อนหน้าเข้ามาเป็นหลักฐาน (evidence)

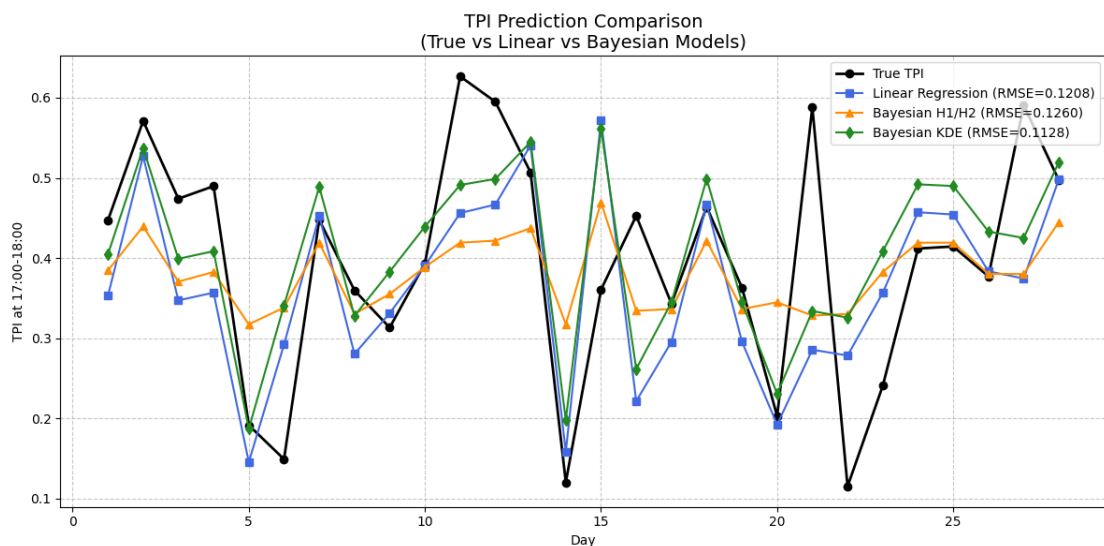
ในกราฟ เส้นสีเขียวแสดงค่า expected value ที่คำนวณจาก posterior distribution ซึ่งได้รับการอัปเดตด้วยข้อมูลจริงรายวัน ขณะที่เส้นสีม่วงแสดง expected value ที่ได้จาก prior distribution ซึ่งถูกประมาณด้วย KDE โดยไม่ถูกปรับเปลี่ยนตามข้อมูลใหม่ และเส้นสีดำแสดงค่าจริงของ TPI ที่สังเกตได้จากข้อมูล

จะเห็นว่า แบบจำลองสามารถตอบสนองต่อข้อมูลที่มีลักษณะผันผวนหรือไม่เป็นเชิงเส้นได้ดี โดยค่าที่ทำนายได้มีแนวโน้มใกล้เคียงกับค่าจริงในหลายช่วงเวลา เช่น วันที่ 9, 12, 16, 17, 18 และ 19 ซึ่งค่าทำนายมีการเปลี่ยนแปลงไปตามข้อมูลใหม่ได้อย่างยืดหยุ่น ไม่เกาะอยู่กับค่าคงที่แบบ prior เฉพาะจุด

อย่างไรก็ตาม แม้แบบจำลองนี้จะมีความแม่นยำที่ดีขึ้นเมื่อเทียบกับ prior เพียงอย่างเดียว แต่ในบางกรณียังคงพบความคลาดเคลื่อนระหว่างค่าทำนายกับค่าจริง เช่น วันที่ 20 ซึ่งค่าที่ทำนายได้อยู่ที่ประมาณ 0.32 ขณะที่ค่าจริงอยู่ที่ 0.59 ความคลาดเคลื่อนในลักษณะนี้อาจเกิดจากการใช้ค่า expected value เพียงอย่างเดียวเป็นตัวแทนของค่าที่คาดว่าจะเกิดขึ้น ซึ่งอาจไม่เพียงพอในการสะท้อนลักษณะข้อมูลที่มีการกระจายกว้าง

โดยสรุป แบบจำลอง Bayesian ร่วมกับ KDE สามารถสะท้อนแนวโน้มของข้อมูลจริงได้ดี และเหมาะสมอย่างยิ่งกับข้อมูลที่ไม่เป็นเชิงเส้น หรือไม่ได้มีการแจกแจงแบบปกติ เช่น ข้อมูล TPI ที่มีลักษณะกระจายแบบอสมมาตรและเปลี่ยนแปลงตลอดเวลา ทั้งนี้ประสิทธิภาพของแบบจำลองยังขึ้นอยู่กับลักษณะของ prior และคุณภาพของข้อมูลที่ใช้ในการอัปเดตด้วย

#### 4.3 ผลการเปรียบเทียบทั้ง 3 แบบจำลอง



รูปที่ 4.6 กราฟการเปรียบเทียบระหว่างค่า TPI จริง และค่าที่ทำนายจากแบบจำลองด้วยวิธีการ Linear Regression, Bayesian ( $H_1/H_2$ ), Bayesian Continuous Distribution KDE ทั้งหมด 27 วัน ของถนนประมาณ 100 เมตร ก่อนถึงแยกซอยบ้านกล้วยใต้

รูปที่ 4.6 แสดงถึงการเปรียบเทียบกันของค่า TPI ที่ทำนายได้ในแต่ละแบบจำลองกับค่าจริง TPI เพื่อเปรียบเทียบประสิทธิภาพของทั้ง 3 แบบจำลอง พร้อมค่า Root Mean Square Error (RMSE) พบว่า ทั้ง 3 แบบจำลองให้ประสิทธิภาพที่ใกล้เคียงกันมาก

แบบจำลอง Bayesian KDE ซึ่งแสดงจากกราฟเส้นสีเขียวจะทำนายได้ดีในช่วงที่ข้อมูลมาการเปลี่ยนแปลงอย่างฉับพลัน เช่น วันที่ 2, 5 และ 17 ของกลุ่มตัวอย่าง

แบบจำลอง Linear Regression ซึ่งแสดงจากกราฟเส้นสีฟ้า จะเห็นได้ว่าแบบจำลองนี้จะค่อนข้างทำนายค่าได้ดีในหลายๆจุด แต่จะมีการคาดการณ์ที่ไม่ค่อยใกล้เคียงในบางจุด เช่น วันที่ 11 และ 21 จะกลุ่มตัวอย่างวัน ซึ่งมีค่าต่างกับค่าจริง TPI ค่อนข้างมาก เนื่องด้วยข้อมูลย้อนหลัง (historical data) จะมีการกระจายแบบไม่ต่อเนื่อง และมีความผันผวนข้อมูลสูง

แบบจำลอง Bayesian ( $H_1/H_2$ ) มีค่าที่ทำนายได้ใกล้เคียงกับค่าทำจากแบบจำลองอื่นๆ เช่นกัน แต่ส่วนใหญ่แล้ว ค่าที่ทำนายได้จะห่างจากค่าจริง TPI กว่าแบบจำลองอื่นๆ

อย่างไรก็ตาม จากรูปที่ 4.6 ถึงแม้ว่า แบบจำลองด้วยวิธี Linear Regression มีแนวโน้มที่ทำนายค่าได้ดีกว่าแบบจำลองอื่นๆ แต่เมื่อนำค่า Root Mean Square Error (RMSE) มาเปรียบเทียบกับกันแล้ว แบบจำลองด้วยวิธี Bayesian KDE สามารถทำนายออกมาได้ดีกว่า โดยเรียงลำดับ RMSE ได้ดังนี้ Bayesian KDE < Linear Regression < Bayesian ( $H_1/H_2$ ) เป็น 0.1128, 0.1208 และ 0.1260 ตามลำดับ

โดยสรุป แบบจำลอง Linear Regression มีแนวโน้มทำนายค่าได้ค่อนข้างดีในหลายช่วงเวลา แต่ยังไม่เหมาะกับข้อมูลที่มีการผันผวนสูง ขณะที่แบบจำลอง Bayesian KDE ให้ค่าความคลาดเคลื่อนต่ำที่สุด (RMSE = 0.1128) และสามารถตอบสนองต่อข้อมูลที่มีแนวโน้มเปลี่ยนแปลงฉับพลันได้อย่างมีประสิทธิภาพ กล่าวได้ว่าแบบจำลองนี้มีความเหมาะสมในการพยากรณ์ค่าดัชนี TPI ทั้งในช่วงที่ข้อมูลมีลักษณะปกติและช่วงที่มีความผันผวนสูง ซึ่งแสดงให้เห็นถึงศักยภาพของการนำการพยากรณ์ TPI ไปประยุกต์ใช้ในเชิงปฏิบัติ เช่น การบริหารจัดการสัญญาไฟฟ้า การแจ้งเตือนช่วงเวลาเสี่ยง หรือการวางแผนการเดินทางของผู้ใช้ถนนในพื้นที่ที่มีแนวโน้มแออัดซ้ำซ้อน

## บทที่ 5

### สรุปผลการวิจัย อภิปรายผล และข้อเสนอแนะ

ในบทนี้จะนำเสนอการสรุปผลการวิจัยที่ได้จากการทดลองตามกระบวนการที่ได้ดำเนินการไว้ในบทก่อนหน้านี้ โดยเฉพาะผลการวิเคราะห์ค่าดัชนี TPI และผลการพยากรณ์ที่ได้จากแบบจำลองทั้งสามรูปแบบ ได้แก่ แบบจำลอง Linear Regression, Bayesian แบบจำแนกกลุ่ม ( $H_1/H_2$ ) และ Bayesian แบบ Continuous Distribution โดยใช้ Kernel Density Estimation (KDE) เพื่อแสดงให้เห็นถึงความสามารถของแต่ละแบบจำลองในการทำนายค่าดัชนี TPI ในช่วงเวลาถัดไป

นอกจากนี้ ยังมีการอภิปรายผลการวิจัยโดยเปรียบเทียบความเหมาะสมของแต่ละแบบจำลองกับลักษณะข้อมูลจริงที่พบในแต่ละช่วงเวลา และอธิบายข้อจำกัดเชิงโครงสร้างของแบบจำลอง พร้อมทั้งเชื่อมโยงกับลักษณะการจราจรในช่วงเวลาที่มีความแออัด ซึ่งมีการวิเคราะห์ไว้ในบทที่ 4

#### 5.1 สรุปผลงานวิจัย

งานวิจัยนี้มีวัตถุประสงค์เพื่อวิเคราะห์และพยากรณ์ลักษณะการจราจรบนถนนพระราม 4 โดยใช้ข้อมูล GPS จาก Grab Taxi ซึ่งนำมาประมวลผลเป็นค่าดัชนีประสิทธิภาพการจราจร (Traffic Performance Index: TPI) ในช่วงเวลา 1 ชั่วโมง แล้วนำไปใช้ในแบบจำลองเชิงพยากรณ์ในช่วงเวลาถัดไป

จากการวิเคราะห์ในบทที่ 4.1 พบว่า ช่วงเวลาที่มีแนวโน้มการจราจรติดขัดสูงที่สุดคือช่วงเวลา 16:00–17:00 น. ซึ่งมีค่า TPI ต่ำกว่าช่วงเวลาอื่นอย่างชัดเจน ข้อมูลดังกล่าวสะท้อนถึงลักษณะของสภาพจราจรในช่วงเย็นของวันทำงาน ที่มักมีความหนาแน่นของรถยนต์เพิ่มขึ้นจากการเดินทางกลับบ้านและการใช้ถนนร่วมกับเส้นทางสายหลักอื่น การทำความเข้าใจช่วงเวลาที่มีความเสี่ยงสูงจึงเป็นพื้นฐานสำคัญของการพัฒนาแบบจำลองพยากรณ์

ในงานวิจัยนี้ได้ทำการสร้างแบบจำลองขึ้นทั้งหมด 3 แบบ ได้แก่ แบบจำลอง Linear Regression, Bayesian (แบบจำแนกกลุ่ม  $H_1/H_2$ ), และ Bayesian (แบบ Continuous Distribution โดยใช้ Kernel Density Estimation – KDE) ซึ่งแต่ละแบบมีข้อดีและข้อจำกัดแตกต่างกันไป

ผลการวิจัยแสดงให้เห็นว่าแบบจำลอง Linear Regression มีความสามารถในการพยากรณ์ได้แม่นยำในช่วงเวลาที่ข้อมูลมีความเสถียร และมีแนวโน้มเชิงเส้นที่ชัดเจน ส่วนแบบจำลอง Bayesian แบบ  $H_1/H_2$  สามารถสะท้อนค่ากลุ่มได้ดี แต่ยังมีข้อจำกัดเมื่อต้องพยากรณ์ข้อมูลที่มีการเปลี่ยนแปลงฉับพลัน ในขณะที่แบบจำลอง Bayesian แบบ Continuous Distribution ที่ใช้ KDE มี

ความยืดหยุ่นสูงสุด และสามารถตอบสนองต่อข้อมูลที่มีความแปรปรวนหรือไม่เป็นเชิงเส้นได้ดี โดยมีความคลาดเคลื่อนน้อยที่สุด ( $RMSE = 0.1128$ )

จากผลการเปรียบเทียบทั้งสามแบบจำลอง พบว่าแบบจำลอง Bayesian KDE มีความแม่นยำสูงสุด เหมาะสมต่อการพยากรณ์ค่าดัชนี TPI ทั้งในกรณีที่ข้อมูลมีแนวโน้มปกติและข้อมูลที่เปลี่ยนแปลงอย่างรวดเร็ว โดยเฉพาะในช่วงเวลาที่มีความเสี่ยงต่อการจลาจรติดขัดสูง ซึ่งแสดงให้เห็นถึงศักยภาพของแบบจำลองในการประยุกต์ใช้กับระบบบริหารจัดการจราจรในพื้นที่จริง

## 5.2 อภิปรายผลงานวิจัย

จากการศึกษาค่า TPI ในแต่ละช่วงเวลา พบว่าช่วงเวลา 16:00–17:00 น. เป็นช่วงที่มีค่าดัชนีต่ำกว่าช่วงอื่นอย่างสม่ำเสมอ ซึ่งสามารถสะท้อนปัญหาการจราจรติดขัดในช่วงเร่งด่วนได้อย่างชัดเจน การพยากรณ์ค่าดัชนี TPI ในช่วง 17:00–18:00 น. จึงมีความสำคัญอย่างยิ่งในการช่วยคาดการณ์แนวโน้มของความแออัดที่อาจเกิดขึ้นต่อเนื่อง

แบบจำลองแต่ละรูปแบบมีจุดเด่นและข้อจำกัดที่แตกต่างกัน โดยแบบจำลอง Linear Regression มีข้อได้เปรียบในด้านความเรียบง่ายและใช้ข้อมูลน้อย แต่ไม่สามารถติดตามพฤติกรรมของข้อมูลที่มีความผันผวนหรือไม่เป็นเชิงเส้นได้ดี ส่วนแบบจำลอง Bayesian  $H_1/H_2$  ซึ่งใช้การแบ่งกลุ่มตาม threshold สามารถคาดการณ์ค่ากลุ่มโดยรวมได้แม่นยำในกรณีที่ค่าจริงไม่เปลี่ยนแปลงมากนัก แต่เมื่อเกิดเหตุการณ์ที่มีการเปลี่ยนแปลงฉับพลันแบบจำลองจะไม่สามารถตอบสนองได้ทันทั่วทั้งที่ เนื่องจากใช้โครงสร้างกลุ่มแบบคงที่

ในขณะที่แบบจำลอง Bayesian KDE ซึ่งใช้การประมาณค่าความหนาแน่นแบบไม่พารามิเตอร์ สามารถตอบสนองต่อข้อมูลได้อย่างยืดหยุ่น และสะท้อนค่าที่ใกล้เคียงกับค่าจริงได้ดีแม้ในช่วงที่ข้อมูลมีความแปรปรวนสูง โดยเฉพาะในกรณีที่ TPI กระโดดขึ้นหรือลงอย่างรวดเร็ว เช่นที่พบในช่วงวันที่ 9, 12, 19 และ 22 อย่างไรก็ตาม แบบจำลองนี้ยังมีข้อจำกัดเมื่อค่าที่ทำนายมีลักษณะ extreme เนื่องจากการใช้ค่า expected value เพียงอย่างเดียว อาจไม่สามารถสะท้อนการกระจายของ posterior ได้ครบถ้วน

จากการทดลองทั้ง 3 แบบจำลอง จึงสามารถสรุปได้ว่าแบบจำลอง Bayesian KDE เหมาะสมที่สุดสำหรับข้อมูลการจราจรที่มีความไม่แน่นอน หรือไม่มีลักษณะจำเพาะแบบเชิงเส้น โดยเฉพาะในช่วงเวลาที่มีความเสี่ยงสูงต่อการติดขัดตามที่พบในบทที่ 4.1 และสามารถนำไปต่อยอดในระบบพยากรณ์การจราจรที่มีความยืดหยุ่นสูงได้ในอนาคต

จากผลการพยากรณ์ที่ได้ แสดงให้เห็นว่าการนำดัชนี TPI มาใช้เป็นตัวแปรเป้าหมายในแบบจำลอง สามารถสะท้อนลักษณะของสภาพจราจรได้อย่างมีประสิทธิภาพ โดยเฉพาะเมื่อใช้ร่วมกับแบบจำลองที่สามารถปรับตัวตามข้อมูล เช่น Bayesian KDE จึงชี้ให้เห็นว่าแนวทางการ

ทำนายค่าดัชนี TPI ล่วงหน้า มีศักยภาพสูงในการใช้เป็นเครื่องมือสำหรับการบริหารจัดการจราจรอย่างเป็นระบบในระดับพื้นที่

### 5.3 ข้อเสนอแนะ

#### 5.3.1 ข้อเสนอแนะในการนำผลวิจัยไปใช้

ผลการวิจัยในครั้งนี้สามารถนำไปประยุกต์ใช้ในการพัฒนาระบบสนับสนุนการตัดสินใจด้านการจราจรในเขตเมือง โดยเฉพาะในด้านการเฝ้าระวังแนวโน้มความแออัดล่วงหน้าในช่วงเวลาถัดไป ทั้งนี้ การนำข้อมูล GPS ของยานพาหนะมาแปลงเป็นค่าดัชนี TPI ซึ่งอยู่ในช่วง 0 ถึง 1 จะช่วยให้สามารถประเมินระดับความหนาแน่นของการจราจรในเชิงเปรียบเทียบข้ามวันและช่วงเวลาได้อย่างเป็นระบบ และสามารถใช้เป็นเครื่องมือวิเคราะห์ความเปลี่ยนแปลงของสภาพการจราจรในพื้นที่เฉพาะจุดได้แบบต่อเนื่อง

แบบจำลองที่ใช้ในการพยากรณ์ เช่น Linear Regression และ Bayesian Updating ยังสามารถนำไปประยุกต์ใช้ในระบบที่มีการอัปเดตข้อมูลแบบ real-time เพื่อประเมินความแออัดของถนนในช่วงเวลาถัดไปได้อย่างมีประสิทธิภาพ เช่น ใช้ร่วมกับระบบควบคุมและบริหารจัดการสัญญาณไฟจราจรอัตโนมัติ (adaptive signal control), ระบบแจ้งเตือนความหนาแน่นล่วงหน้าในศูนย์ควบคุมจราจร หรือระบบแนะนำเส้นทางในแอปพลิเคชันนำทาง

นอกจากนี้ ผลการวิเคราะห์ค่าดัชนี TPI ยังสามารถนำไปใช้ในการวางแผนและประเมินผลของมาตรการจราจรในระดับพื้นที่ เช่น การปรับเวลาไฟจราจร การเปิดหรือปิดเส้นทางบางช่วงเวลา หรือการประเมินผลก่อนและหลังดำเนินมาตรการ (before-after analysis) โดยใช้ค่าดัชนี TPI เป็นเครื่องชี้วัดที่เชื่อถือได้

#### 5.3.2 ข้อเสนอแนะในการทำวิจัยครั้งต่อไป

ในการศึกษาครั้งนี้ใช้ข้อมูล GPS จำนวนจำกัด และวิเคราะห์ค่าดัชนี TPI ในภาพรวมของแต่ละช่วงเวลา โดยยังไม่ได้แยกวิเคราะห์ตามลักษณะเฉพาะของวันต่าง ๆ ซึ่งในความเป็นจริงพบว่าแต่ละวันมีพฤติกรรมการจราจรที่แตกต่างกันอย่างชัดเจน เช่น วันศุกร์ มักมีปริมาณรถหนาแน่นในช่วงเย็นจากการเดินทางกลับภูมิลำเนาหรือออกนอกเมือง ขณะที่ วันหยุด มักมีจำนวนผู้ใช้ถนนช่วงเช้าลดลง แต่กลับมีความแออัดเพิ่มขึ้นในช่วงสายหรือเย็นจากกิจกรรมด้านอื่น เช่น การท่องเที่ยว หรือทำธุระส่วนตัว

ในการศึกษาครั้งถัดไป ควรเพิ่มความละเอียดของช่วงเวลาในการวิเคราะห์ (temporal resolution) เช่น จากการใช้ข้อมูลรายชั่วโมง ให้ละเอียดขึ้นเป็นช่วงละ 15 นาที หรือ 5 นาที เพื่อให้สามารถติดตามแนวโน้มของความแออัดได้ใกล้เคียงกับสถานการณ์จริงมากขึ้น และรองรับการใช้งานในระบบพยากรณ์แบบ near real-time ได้อย่างเหมาะสม

ควรใช้ข้อมูลที่มีจำนวนมากขึ้นและแยกวิเคราะห์ตามลักษณะของวัน เช่น วันธรรมดาทั่วไป วันศุกร์ วันหยุด และวันก่อนหยุด เพื่อให้สามารถสะท้อนลักษณะเฉพาะของพฤติกรรมจราจรในแต่ละวันได้อย่างแม่นยำมากขึ้น

นอกจากนี้ ควรขยายขอบเขตการวิเคราะห์ให้ครอบคลุมถึง ข้อมูลจากถนนโดยรอบ หรือพื้นที่ที่มีความเชื่อมโยงกันในโครงข่ายจราจร เช่น หากพบว่าถนนสายหนึ่งมีความแออัดสูงในเวลา 16:00 น. ก็อาจส่งผลให้ถนนที่อยู่ใกล้เคียงเกิดการสะสมของรถ และมีแนวโน้มจะติดขัดมากขึ้นในเวลา 17:00 น. การรวมบริบทเชิงพื้นที่นี้จะช่วยเพิ่มความแม่นยำของแบบจำลองการพยากรณ์ได้อย่างมีนัยสำคัญ

อีกทั้งควรพิจารณาเพิ่ม ข้อมูลภายนอก ประเภที่สามารถสังเกตและตรวจสอบได้ชัดเจน เช่น ข้อมูลการก่อสร้างหรือการซ่อมบำรุงถนน รวมถึง สภาพอากาศ เช่น ฝนตก ซึ่งมักส่งผลกระทบโดยตรงต่อรูปแบบการจราจรในพื้นที่ใกล้เคียงและช่วงเวลาใกล้เคียง

## เอกสารอ้างอิง

1. Wei, L., Chen, P., Mei, Y., and Wang, Y. (2022). *Turn-level network traffic bottleneck identification using vehicle trajectory data*. School of Transportation Science and Engineering, Beihang University, China.
2. Huimin Wen, Jianping Sun, Xi Zhang. (2014) *Study on Traffic Congestion Patterns of Large City in China Taking Beijing as an Example*. Beijing Transportation Research Center.
3. John G. Manchuk. (2009). *Conversion of Latitude and Longitude to UTM Coordinates*.[https://www.ccgaberta.com/ccgresources/report11/2009-410\\_converting\\_latlon\\_to\\_utm.pdf](https://www.ccgaberta.com/ccgresources/report11/2009-410_converting_latlon_to_utm.pdf)
4. Adam Hayes. (2024, March 30). *Bayes' Theorem: What It Is, the Formula, and Examples*. Investopedia. <https://www.investopedia.com/terms/b/bayes-theorem.asp>
5. Forbes, J.G., Gardner, T., McGee, H., and Srinivasan R. (2012). *Methods and Practices for Setting Speed Limits: An Informational Report*. (Report No.FHWA-SA-12-004). Institute of Transportation Engineers, Washington, DC.
6. Christopher M. Bishop (2006). *Pattern Recognition and Machine Learning*. <https://www.microsoft.com/en-us/research/wp-content/uploads/2006/01/Bishop-Pattern-Recognition-and-Machine-Learning-2006.pdf>
7. Jonah Finkelstein. (2016, October 27). *85th Percentile Speed Explained*. Mike on Traffic. [https://www.mikeontraffic.com/85th-percentile-speed-explained/?utm\\_](https://www.mikeontraffic.com/85th-percentile-speed-explained/?utm_)