การทำนายจำนวนผู้ติดเชื้อโรคโควิด-19 ในประเทศไทยด้วยแบบจำลองทางคณิตศาสตร์

Predicting Number of COVID-19 Cases in Thailand Based on Mathematical Model

สุวิทย์ กิระวิทยา

ภาควิชาวิศวกรรมไฟฟ้าและคอมพิวเตอร์ คณะวิศวกรรมศาสตร์ มหาวิทยาลัยนเรศวร

28 มีนาคม พ.ศ. 2563

suwitki@gmail.com

บทคัดย่อ

บทความนี้แสดงแนวทางการนำแบบจำลองทางคณิตศาสตร์มาใช้ในการทำนายจำนวนผู้ติดเชื้อโรคโควิด-19 ที่กำลังระบาดอย่างมากในประเทศไทย โดยการวิเคราะห์แบบจำลอง SIR ทำให้เราสามารถดึงพารามิเตอร์ที่สำคัญที่บ่งบอกระดับความรุนแรงของการระบาดออกมาได้ ซึ่งข้อมูลของการระบาดในประเทศจีนที่มีแนวโน้มใกล้จะสิ้นสุดลงแล้วได้ถูกนำมาวิเคราะห์ก่อนเพื่อเป็นข้อมูลอ้างอิงเบื้องต้น จากนั้นจึงนำผลที่ได้มาเปรียบเทียบกับผลการศึกษาข้อมูลจำนวนผู้ติดเชื้อในประเทศไทยและทำนายการเปลี่ยนแปลงของจำนวนผู้ติดเชื้อในประเทศไทยในอนาคต จากการศึกษาแบบจำลองนี้ทำให้เราสามารถคาดการณ์กรณีที่ดีที่สุดได้ แต่สำหรับกรณีที่แย่ที่สุดที่อาจเกิดขึ้นได้จะขึ้นกับจำนวนผู้ติดเชื้อต่อวันสูงสุดที่ยังไม่ทราบค่า ณ เวลาขณะนี้

Abstract

This article explains a method to predict the number of COVID-19 cases in Thailand by using a conventional mathematical model. By analyzing the SIR model, we can extract important model parameters, which indicate the degree of the pandemicity. Firstly, recorded cases of China, which are nearly complete, have been considered. Then, the result is compared with the study of the infected cases in Thailand. The predictions are made for the cases in Thailand. Based on this study, the best-case scenario can be depicted while the worst case can be predicted only when the maximum number of infected cases per day is revealed.

1. บทนำ

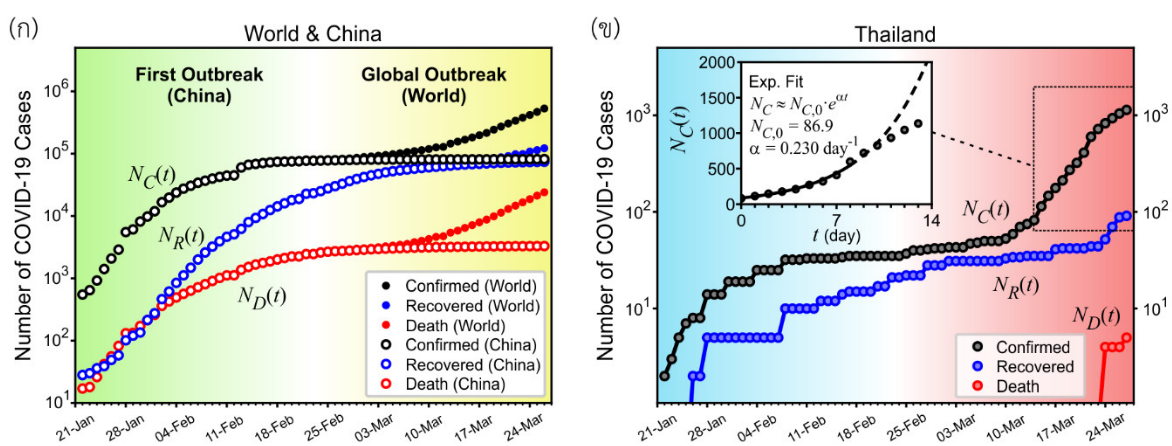
โรคโควิด-19 (COVID-19) เกิดจากโคโรนาไวรัสสายพันธุ์ใหม่ที่ชื่อว่า SARS-CoV-2 ที่ถือกำเนิดขึ้นในเมืองวู่ฮั่น (Wuhan) มณฑลฮูไป่ (Hubei) ประเทศจีน [1] การอุบัติใหม่ของเชื้อโรคนี้ได้ถูกรายงานไปยังองค์การอนามัยโลกในวันที่ 31 ธันวาคม ค.ศ. 2019 [2] และต่อมาเมื่อโรคนี้ได้ระบาดรุนแรงมากในช่วงเดือนมกราคม-มีนาคม ทางองค์การอนามัยโลกจึงได้กำหนดให้โรคโควิด-19 นี้เป็นโรคระบาดร้ายแรง (Pandemic) ในวันที่ 12 มีนาคม พ.ศ. 2563 ซึ่งสำหรับเชื้อไวรัสนี้จะระบาดโดยการติดต่อผ่านลมหายใจและสารคัดหลั่ง เช่น น้ำมูกและน้ำลาย เนื่องจากประเทศไทยรับนักท่องเที่ยวจากประเทศจีนเป็นจำนวนมาก ทำให้ประเทศไทยเป็นประเทศแรก ๆ ในโลกที่พบผู้ติดเชื้อโรคโควิด-19 นอกประเทศจีนในช่วงต้นเดือนมกราคม ซึ่งสถานการณ์ปัจจุบันของประเทศไทย (รายงาน ณ วันที่ 28 มีนาคม พ.ศ. 2563) คือ มีผู้ติดเชื้อสะสม 1,245 ราย (รักษาหายแล้ว 100 ราย กำลังรักษา 1,139 ราย และเสียชีวิตรวม 6 ราย) และทุกฝ่าย (ภาครัฐ ภาคเอกชนและประชาชน) กำลังดำเนินมาตรการต่าง ๆ เพื่อยับยั้งการระบาดของโรคนี้

บทความนี้จะนำเสนอแนวทางในการคาดการณ์จำนวนผู้ติดเชื้อที่เปลี่ยนแปลงไปตามเวลาในกรณีที่ดีที่สุดสำหรับประเทศไทย โดยใช้แบบจำลองทางคณิตศาสตร์ที่เรียกว่าแบบจำลอง SIR [3-5] ซึ่งข้อมูลของการระบาดในประเทศจีนจาก [2] ถูกนำมาวิเคราะห์และบรรยาย เพื่อนำมาเป็นข้อมูลอ้างอิงเบื้องต้นในการพิจารณาข้อมูลประเทศไทย โดยพารามิเตอร์ที่ดึงออกมาได้จากการฟิตข้อมูลจริงด้วยแบบจำลองจะถูกนำมาใช้บ่งบอกสถานการณ์และใช้ทำนายจำนวนผู้ติดเชื้อในอนาคต ซึ่งผลที่ได้นี้ทำให้เราสามารถคาดการณ์กรณีดีที่สุดได้บนพื้นฐานของแบบจำลองทางคณิตศาสตร์ ในขณะเรายังไม่สามารถคาดการณ์กรณีที่แย่ที่สุดได้ เนื่องจากยังไม่มีข้อมูลค่าสูงสุดของจำนวนผู้ติดเชื้อต่อวัน (ในขณะที่เขียนบทความ จำนวนผู้ติดเชื้อต่อวันยังคงมีแนวโน้มที่เพิ่มมากขึ้นทุกวัน)

2. ข้อมูลผู้ติดเชื้อ

จากการยกระดับการเฝ้าระวังโรคโควิด-19 ทำให้ทางองค์การอนามัยโลกทำการรายงาน บันทึกและปรับปรุงข้อมูลสถานการณ์และจำนวนผู้ติดเชื้อสะสมทุกวัน ผ่านฐานข้อมูลเปิด [2] ซึ่งข้อมูลที่ถูกบันทึกนี้ สามารถนำมาพล็อตรวมหรือแยกตามประเทศและภูมิภาคได้ โดยรูปที่ 1(ก) และ (ข) แสดงการเปลี่ยนแปลงตามเวลาของจำนวนผู้ติดเชื้อสะสม $N_C(t)$ จำนวนผู้ที่รักษาหายสะสม $N_R(t)$ และจำนวนผู้เสียชีวิตสะสม $N_D(t)$ ของโลกและประเทศจีนและประเทศไทย ตามลำดับ การแสดงที่นำเสนอนี้แสดงในล็อกสเกลเพื่อให้เห็นการเปลี่ยนแปลงในระดับที่ใหญ่ขึ้น จากรูปที่ 1(ก) จะเห็นได้ว่า การระบาดในจีนซึ่งเป็นการระบาดครั้งแรกในโลกมีแนวโน้มลดลงแล้ว คือมีจำนวนผู้ติดเชื้อเพิ่มขึ้นน้อยกว่าการเพิ่มของจำนวนผู้ที่รักษาหาย ในขณะที่การระบาดในประเทศอื่น ๆ ในโลก นอกเหนือจากประเทศจีน กำลังเริ่มขึ้นทวีความรุนแรงขึ้นและรุนแรงมากกว่าการระบาดในจีนที่อยู่ตัวแล้ว

รูปที่ 1(ข) แสดงถึงการเปลี่ยนแปลงจำนวนผู้ติดเชื้อสะสม $N_C(t)$ จำนวนผู้ที่รักษาหายสะสม $N_R(t)$ และจำนวนผู้เสียชีวิตสะสม $N_D(t)$ ในประเทศไทย ที่มีการรายงานทุกวัน ซึ่งแนวโน้มการเปลี่ยนแปลงของจำนวนเหล่านี้ มีลักษณะเช่นเดียวกับของประเทศอื่น ๆ โดยการระบาดรุนแรงที่เกิดขึ้นในประเทศไทย มีสาเหตุหลักมาจากเหตุการณ์การแพร่เชื้อที่สนามมวย [6] ซึ่งทราบทั่วกันในวันที่ 13 มีนาคม พ.ศ. 2563 และหลังจากนั้นก็มีการพบผู้ติดเชื้อรวมมากกว่าหนึ่งร้อยราย (กรอบเส้นประในรูปที่ 1(ข)) โดยจากเหตุการณ์นี้ หน่วยงานต่าง ๆ ที่เกี่ยวข้องก็ได้มีการออกมาตรการต่าง ๆ เพื่อยับยั้งการระบาดนี้



รูปที่ 1 การเปลี่ยนแปลงจำนวนผู้ติดเชื้อสะสม $N_C(t)$ ผู้ที่รักษาหายสะสม $N_R(t)$ และผู้เสียชีวิตสะสม $N_D(t)$ ของ (ก) โลกและประเทศจีนและ (ข) ประเทศไทย (ข้อมูล ณ วันที่ 27 มีนาคม 2563 จาก [2,6])

รูปเล็กในรูปที่ 1(ข) แสดงการเพิ่มขึ้นของจำนวนผู้ติดเชื้อในประเทศไทยในสเกลเชิงเส้น โดยหากกำหนดให้ วันที่ 14 มีนาคม เป็นวันที่ 0 ($t = 0$) จากกราฟจะเห็นได้ว่า ประเทศไทยมีจำนวนผู้ติดเชื้อมากกว่าหนึ่งร้อยรายในวันที่ 15 มีนาคมและเพิ่มขึ้นอย่างรวดเร็วตั้งแต่ช่วงวันที่ 15-22 มีนาคม โดยอัตราการเพิ่มจำนวนนี้ลดน้อยลงในช่วงวันที่ 23-27 มีนาคม ดังนั้นเราจึงคาดคะเนได้ว่าจำนวนผู้ติดเชื้อสะสม $N_C(t)$ จะเพิ่มขึ้นแบบเอกซ์โพเนนเชียลในช่วงแรกของการระบาด ซึ่งจากการฟิตจำนวนผู้ติดเชื้อตามเวลา ด้วยสมการ

$$N_C(t) = N_{C,0} \cdot e^{\alpha t} \tag{1}$$

โดย $N_{C,0}$ คือจำนวนผู้ติดเชื้อที่เวลา $t = 0$ และ α คือค่าอัตราการเพิ่มจำนวนผู้ติดเชื้อต่อวัน (day^{-1})

การฟิตข้อมูลจำนวนผู้ติดเชื้อของประเทศไทยในระหว่างวันที่ 14-22 มีนาคม ได้ผลคือ $N_{C,0} = 86.9$ รายและ $\alpha = 0.230$ ต่อวัน โดยผลการฟิตนี้แสดงเป็นเส้นทึบในรูปเล็กในรูปที่ 1(ข) และจากการพล็อตสมการที่ได้นี้ออกไปนอกช่วงที่ฟิต (เส้นประในรูปเล็กในรูปที่ 1(ข)) จะพบว่า การเพิ่มขึ้นของจำนวนผู้ติดเชื้อ*มิได้*เพิ่มขึ้นแบบเอกซ์โพเนนเชียลในอัตรา 0.230 ต่อวัน แล้วแต่มีแนวโน้มการเพิ่มที่ลดลง ซึ่งผลที่สังเกตเห็นนี้อาจเป็นผลมาจากมาตรการการยับยั้งการระบาดที่เริ่มบังคับใช้มากขึ้น โดยจากผลที่ได้รับนี้ เราจึงควรพิจารณาข้อมูลนี้ด้วยแบบจำลองที่ถูกต้องมากยิ่งขึ้นการคาดการณ์ว่าการเพิ่มขึ้นจะเป็นแบบเอกซ์โพเนนเชียลด้วยอัตราการเพิ่มคงที่

3. แบบจำลอง SIR

แบบจำลองทางคณิตศาสตร์ที่นิยมนำมาใช้ในการอธิบายการเปลี่ยนแปลงของจำนวนผู้ที่เกี่ยวข้องกับการระบาดของโรคคือแบบจำลอง SIR [3-5] ซึ่งเป็นแบบจำลองที่เริ่มจากการจัดแบ่งกลุ่ม (Compartment) แล้วอธิบายการเปลี่ยนแปลงจำนวนของประชากรในกลุ่มนั้น ๆ ด้วยการสมการเชิงอนุพันธ์อันดับหนึ่ง โดยในแบบจำลอง SIR นี้จะมีการแบ่งกลุ่มผู้เกี่ยวข้องเป็น 3 กลุ่ม คือ กลุ่มผู้ที่เสี่ยงต่อการติดโรค S (Susceptible) กลุ่มผู้ที่ติดโรค I (Infected) และ กลุ่มผู้ที่หายจากการติดโรค R (Recovered หรือ Removed) โดยกลุ่มสุดท้ายนี้จะหมายรวมถึงผู้ที่รักษาหายและผู้ป่วยที่เสียชีวิตจากโรคด้วย (ดูรูปที่ 2(ก))

แนวคิดเบื้องต้นของแบบจำลอง SIR นี้ คือ เมื่อเกิดโรคระบาดแล้ว การระบาดก็จะเพิ่มจำนวนผู้ติดโรคและทำให้ผู้ที่เสี่ยงต่อการติดโรคลดจำนวนลง (คือ ผู้ที่เสี่ยงกลายไปเป็นผู้ติดโรค) ซึ่งอัตราการลดจำนวนผู้ที่เสี่ยงต่อการติดโรคจะแปรผันตรงกับจำนวนผู้ติดโรคแล้วและแปรผันตรงกับจำนวนผู้ที่เสี่ยงต่อการติดโรค ณ ขณะนั้นด้วย โดยเมื่อจำนวนผู้ติดโรคเพิ่มมากขึ้น ก็จะมีผู้ติดโรคบางส่วนได้รับการรักษาให้หายหรือเสียชีวิตลงไป โดยอัตราการเพิ่มของจำนวนผู้ที่หายจากการติดโรคนี้จะแปรผันตรงกับจำนวนผู้ติดโรค ณ ขณะนั้น

ดังนั้น หากเรากำหนดตัวแปรให้ จำนวนประชากรโดยรวมที่เกี่ยวกับการแพร่ของโรคระบาดคือ N_T (สมมุติให้มีค่าคงที่) สัดส่วนจำนวนผู้ที่เสี่ยงต่อการติดโรคต่อประชากร คือ $S(t)$ สัดส่วนจำนวนผู้ที่ติดโรคต่อประชากร คือ $I(t)$ และ สัดส่วนจำนวนผู้ที่หายจากการติดโรค คือ $R(t)$ (จะได้ว่า $S(t)+I(t)+R(t)=1$ เสมอ และจำนวนประชากรในแต่ละกลุ่มคือผลคูณของ N_T กับ $S(t)$, $I(t)$, และ $R(t)$) จากแบบจำลอง SIR ที่บรรยายในย่อหน้าที่แล้ว เราจะสามารถเขียนความสัมพันธ์ของจำนวนต่าง ๆ ได้เป็นสมการเชิงอนุพันธ์แบบไม่เชิงเส้น คือ

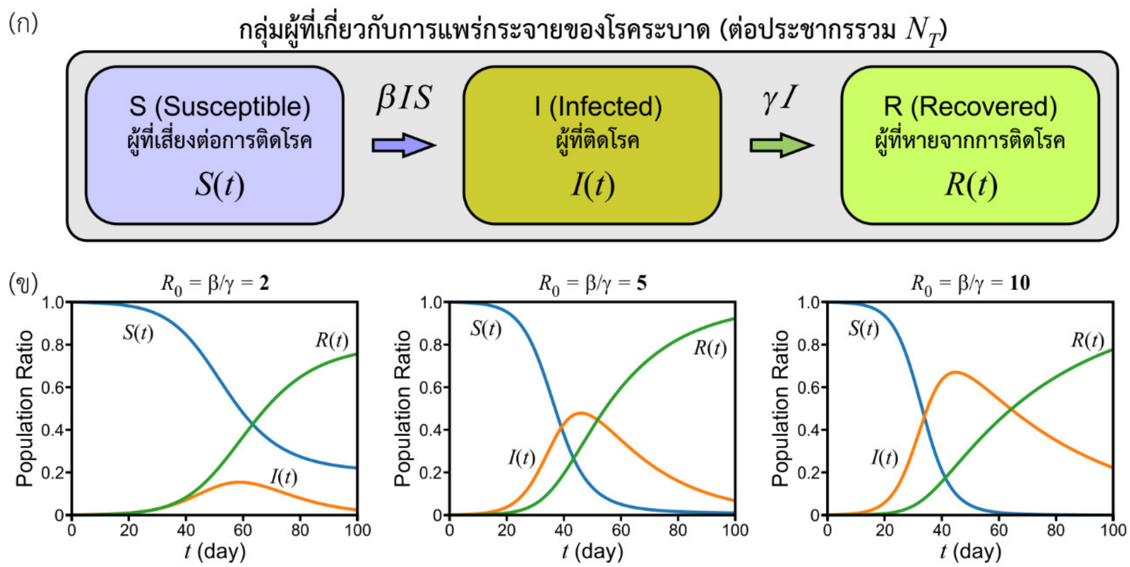
$$\frac{dS}{dt} = -\beta \cdot I \cdot S \tag{2}$$

$$\frac{dI}{dt} = \beta \cdot I \cdot S - \gamma \cdot I \tag{3}$$

$$\frac{dR}{dt} = \gamma \cdot I \tag{4}$$

โดย β คือค่าเฉลี่ยของจำนวนครั้งที่แต่ละบุคคลสัมผัสกันในหนึ่งช่วงเวลาคูณกับความน่าจะเป็นที่เชื้อโรคจะถูกส่งผ่านจากการสัมผัสระหว่างผู้ที่เสี่ยงต่อการติดโรคและผู้ติดโรค และ γ คืออัตราเฉลี่ยของการรักษาให้หายจากการติดโรค (รักษาหายหรือเสียชีวิต) โดยหากคิดในหน่วยต่อวันแล้ว จะได้ว่า γ^{-1} คือจำนวนวันเฉลี่ยที่ต้องใช้ในการรักษาผู้ติดโรคนั่นเอง

จากชุดสมการ (2) – (4) และความจริงที่ว่า β และ γ มีค่ามากกว่าศูนย์ ทำให้สังเกตได้ว่าจำนวนผู้ที่เสี่ยงต่อการติดโรคต่อหน่วยประชากรจะมีค่าลดลงเสมอ นั่นคือ $S(t)$ เป็นฟังก์ชันลดทางเดียว (Monotonically Decreasing Function) ในขณะที่สัดส่วนจำนวนผู้ที่หายจากการติดโรคจะมีค่าเพิ่มขึ้นเสมอ นั่น $R(t)$ เป็นฟังก์ชันเพิ่มทางเดียว (Monotonically Increasing Function) โดยจากความจริงที่ว่า ผลรวมของประชากรทั้งสามกลุ่มจะมีค่าคงที่ เราจะแสดงได้ว่าที่เวลานาน ๆ ($t \to \infty$) ประชากรทั้งหมดจะกลายเป็นผู้ที่หายจากการติดโรค ($S(t \to \infty) = 0$ และ $R(t \to \infty) = 1$) และการศึกษาการเปลี่ยนแปลงของปริมาณเพียงสองตัวเช่น $I(t)$ และ $R(t)$ ในแบบจำลอง SIR นี้ก็เพียงพอแล้ว โดยสำหรับฟังก์ชัน $I(t)$ จะเป็นไปได้หลายรูปแบบซึ่งขึ้นกับอัตราการเพิ่มและลดของจำนวนผู้ติดโรคนี้ (สมการ (3)) ลูกศรในรูปที่ 2(ก) ถูกกำกับด้วยสัญลักษณ์ของอัตราการเปลี่ยนแปลงของประชากรทั้งสามกลุ่มนี้ โดย ณ ช่วงเวลาเริ่มต้น เราอาจกล่าวได้ว่า $S(t \approx 0) \approx 1, I(t \approx 0) \approx 0$ และ $R(t \approx 0) \approx 0$



รูปที่ 2 (ก) แผนภาพแสดงความสัมพันธ์ของกลุ่มผู้ที่เกี่ยวข้องกับการแพร่กระจายของโรคระบาด (คิดต่อประชากรรวม N_T คงที่) และ (ข) ตัวอย่างผลการจำลองการเปลี่ยนแปลงของจำนวนประชากรแต่ละกลุ่ม โดยมีพารามิเตอร์คือ $R_0 = 2, 5,$ และ 10, $\beta = 0.230$ ต่อวัน (คงที่), $I(t = 0) = 10^{-3}$ (คงที่) และ $R(t = 0) = 0$ (คงที่)

การบอกระดับความรุนแรงของการระบาดตามแบบจำลอง SIR นี้ มักจะแสดงด้วยค่าสัดส่วนของค่า β และ γ ซึ่งเรียกว่า **อัตราส่วนการติดโรคในขั้นต้น (Basic Reproduction Ratio)** และจะแสดงด้วยสัญลักษณ์ R_0 ($R_0 \equiv \beta/\gamma$) โดยหาก R_0 มีค่าน้อยกว่า 1 แล้วการระบาดจะเกิดขึ้นอย่างไม่รุนแรง คือ สามารถรักษาหายได้ทันและผู้ที่เสี่ยงจะกลายเป็นผู้ที่หายจากการติดโรคอย่างค่อยเป็นค่อยไป แต่หาก R_0 มีค่ามากกว่า 1 แล้วจะหมายความว่าการระบาดจะเกิดขึ้น รูปที่ 2(ข) แสดงตัวอย่างการเปลี่ยนแปลงของ $S(t)$, $I(t)$ และ $R(t)$ ที่ค่า $R_0 = 2, 5$ และ $R_0 = 10$ ตามลำดับ โดยในการจำลองนี้ กำหนดให้ $\beta = 0.23$ ต่อวัน, $I(0) = 10^{-3}$ และ $R(t = 0) = 0$ คงที่ โค้ดสำหรับการจำลองนี้เขียนด้วยภาษาไพธอนและดัดแปลงมาจาก [5]

นอกจากพารามิเตอร์ β และ γ ที่เป็นพารามิเตอร์สำคัญและเขียนแสดงอย่างชัดแจ้งในแบบจำลอง SIR นี้ เรายังมีพารามิเตอร์คือจำนวนประชากรรวม N_T ด้วย โดยในโมเดลทางทฤษฎีนั้นเรามักจะกำหนดให้มีค่าคงที่ ในขณะที่ในการระบาดจริงของโรคระบาด ค่า N_T อาจจะแปรเปลี่ยนตามปริมาณของการระบาดด้วย ดังนั้นในการฟิตข้อมูลจริงกับแบบจำลอง SIR นี้ พารามิเตอร์ที่สำคัญอีกตัวหนึ่งคือ N_T โดยในการยับยั้งการระบาดของโรคคือการลดค่า N_T ให้น้อยที่สุดนั่นเอง

4. การประมาณพารามิเตอร์ของแบบจำลองจากข้อมูลจริง

4.1 ความสัมพันธ์ระหว่างข้อมูลที่บันทึกและตัวแปรในแบบจำลอง SIR

ข้อมูลที่บันทึกและพล็อตแสดงในหัวข้อที่ 2 คือ จำนวนผู้ติดเชื้อสะสม $N_C(t)$ จำนวนผู้ที่รักษาหายสะสม $N_R(t)$ และจำนวนผู้เสียชีวิตสะสม $N_D(t)$ ซึ่งเป็นข้อมูลที่ขึ้นกับเวลาและสามารถนำมาเชื่อมโยงกับตัวแปรในแบบจำลอง SIR ได้คือ

$$S(t) = 1 - I(t) - R(t) \tag{5}$$

$$I(t) = \frac{N_C(t) - N_R(t) - N_D(t)}{N_T} \tag{6}$$

$$R(t) = \frac{N_R(t) + N_D(t)}{N_T} \tag{7}$$

$$N_T = N_C(t \to \infty) \tag{8}$$

โดยในการเขียนสมการ (8) นั้น เรามีสมมติฐานว่า ณ เวลานานมากหลังการระบาด การระบาดจะสิ้นสุดลงและสามารถทราบจำนวนผู้ติดเชื้อสะสมรวมทั้งหมดได้

นอกจากความสัมพันธ์เชิงเส้นที่แสดงในสมการ (5) – (8) ในการจำลองการระบาดของโรค เราจะต้องมีเงื่อนไขเริ่มต้น (Initial Condition) ที่นำมาจากข้อมูลจริงด้วย โดยเราสามารถกำหนดจุดเริ่มต้น ณ จุดเวลา $t = 0$ ได้ตามลักษณะของข้อมูลจริงที่มี ซึ่งก็คือการแทน $t = 0$ ในตัวแปรในสมการ (5) – (7) และจะได้ว่าเงื่อนไขเริ่มต้นคือ

$$S(0) = 1 - I(0) - R(0) \tag{9}$$

$$I(0) = (N_C(0) - N_R(0)) / N_T \tag{10}$$

$$R(0) = (N_R(0) + N_D(0)) / N_T \tag{11}$$

และจากข้อมูลจริงที่มีอยู่ เราสามารถประมาณค่าพารามิเตอร์ $\beta-\gamma$ ได้จากการฟิตข้อมูลช่วงเริ่มต้นของการระบาดด้วยสมการเอกซ์โพเนนเชียล (ดังที่กระทำกับข้อมูลประเทศไทยและแสดงในรูปเล็กในรูปที่ 1(ข) เพื่อให้ได้ค่า α) โดยเนื่องจากว่า ที่ $t \approx 0$ เรามี $S(t \approx 0) \approx 1$ และจากสมการ (3) จะได้ว่า

$$\frac{dI}{dt} = \beta \cdot I \cdot S - \gamma \cdot I \approx (\beta - \gamma) \cdot I \qquad (12)$$

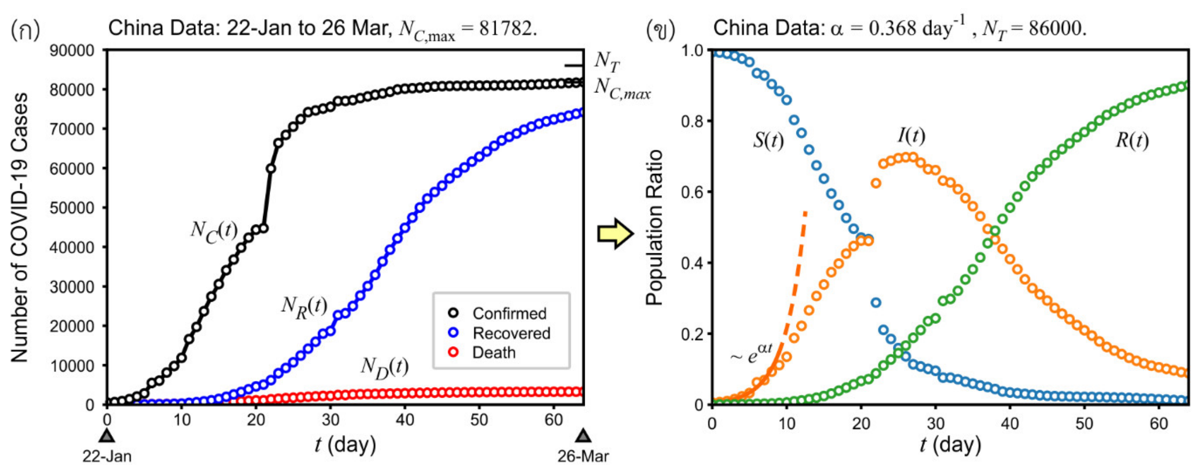
ซึ่งสมการเชิงอนุพันธ์นี้มีผลเฉลยคือ

$$I(t) = I(0) \cdot \exp((\beta - \gamma) \cdot t) \qquad (13)$$

โดยในตอนเริ่มต้นของการระบาด สัดส่วนจำนวนผู้ติดโรคยังคงมีค่าประมาณเท่ากับสัดส่วนจำนวนผู้ติดเชื้อต่อประชากรรวม ($I(t) \approx N_C(t)/N_T$) เนื่องจากจำนวนผู้ที่รักษาหายจากโรคยังมีค่าน้อยมาก (ประมาณเป็นศูนย์ได้) ดังนั้นเราจึงสามารถเขียนได้ว่า

$$\alpha = \beta - \gamma \qquad (14)$$

จากความสัมพันธ์ในชุดสมการ (5) – (8) และข้อมูลการระบาดของโรคโควิด-19 ในประเทศจีนในระหว่างวันที่ 22 มกราคม – 26 มีนาคม (65 วัน) ที่พล็อตในรูปที่ 1(ก) และพล็อตอีกครั้งในสเกลเชิงเส้นในรูปที่ 3(ก) เราสามารถเขียนแสดงเป็นสัดส่วนจำนวนผู้ที่เสี่ยงต่อการติดต่อโรค $S(t)$ สัดส่วนจำนวนผู้ที่ติดโรค $I(t)$ และ สัดส่วนจำนวนผู้ที่หายจากการติดโรค $R(t)$ ของประเทศจีน ได้ดังรูปที่ 3(ข) โดยในการพล็อตนี้ เรากำหนดให้ประชากรรวม N_T มีค่าเท่ากับ 86000 ราย โดยค่านี้เป็นค่าที่ประเมินเบื้องต้น (~1.05 เท่าของค่าสูงสุดของจำนวนผู้ติดเชื้อสะสม) เนื่องจากว่าเรายังไม่ทราบค่าที่แท้จริง (ดูสมการ (8))

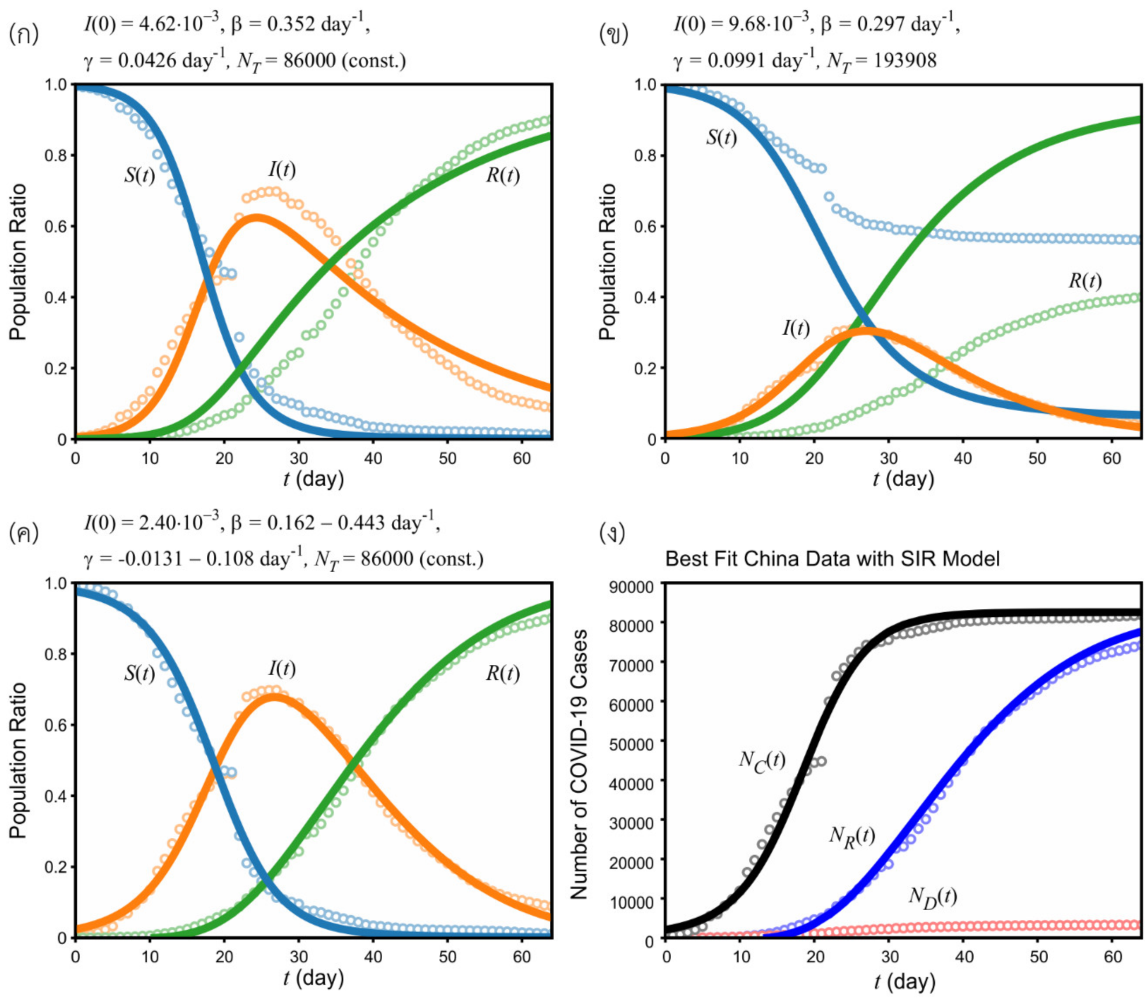


รูปที่ 3 (ก) การเปลี่ยนแปลงจำนวนผู้ติดเชื้อ $N_C(t)$ ผู้ที่รักษาหาย $N_R(t)$ และผู้ตาย $N_D(t)$ และ (ข) สัดส่วนจำนวนผู้ที่เสี่ยงต่อการติดต่อโรค $S(t)$ สัดส่วนจำนวนผู้ที่ติดโรค $I(t)$ และ สัดส่วนจำนวนผู้ที่หายจากการติดโรค $R(t)$ ของประเทศจีน โดยกำหนดให้ประชากรรวม N_T มีค่าเท่ากับ 86000 ราย

4.2 ข้อมูลของประเทศจีน

จากข้อมูลการเปลี่ยนแปลงของค่า $I(t)$ ของประเทศจีนที่มีค่าเพิ่มขึ้นอย่างรวดเร็วในระยะแรกที่แสดงในรูปที่ 3(ข) นั้น เราสามารถฟิตด้วยเอกซ์โพเนนเชียลฟังก์ชันและดึงค่าอัตราการเพิ่มของผู้ที่ติดโรค α ออกมาได้ ซึ่งจากข้อมูลจริงนี้จะได้ว่า สำหรับข้อมูลใน 9 วันแรก (22-30 มกราคม พ.ศ. 2563) ของประเทศจีน จะได้ $\alpha = 0.368$ ต่อวัน และ $I(0) = 5.388{\cdot}10^{-3}$ โดยค่า α ที่ฟิตได้นี้มีค่ามากกว่าของประเทศไทย (ดูรูปเล็กในรูปที่ 1(ข)) อาจเนื่องจากประเทศจีนเป็นประเทศแรกที่มีการระบาดและทำให้การป้องกันตัวของประชากรเกิดขึ้นช้าและการระบาดเกิดขึ้นในวงกว้างแล้ว

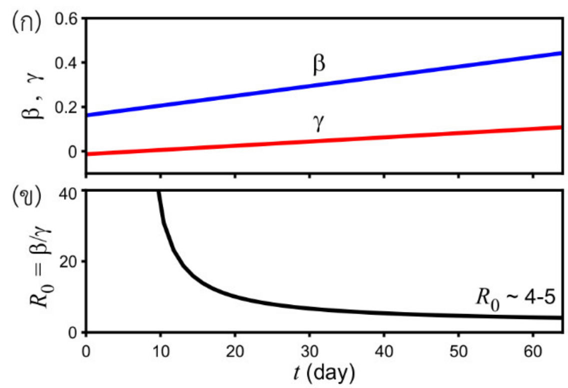
สำหรับการฟิตข้อมูลเชิงเลขด้วยแบบจำลอง SIR นั้น เราทำได้ด้วยการใช้โมดูลสำเร็จรูปชื่อ optimize ในแพ็กเกต Scipy ของภาษาไพธอนที่แจกจ่ายให้ใช้ฟรี [7,8] และทำการฟิตฟังก์ชัน $I(t)$ เพียงฟังก์ชันเดียว แล้วนำผลที่ได้มาพล็อตฟังก์ชันอื่นด้วย จากการฟิตนี้ได้ผลแสดงในรูปที่ 4 ซึ่งผลการฟิตที่แสดงในรูปที่ 4(ก) นี้ มีพารามิเตอร์ที่แปรเปลี่ยนได้ในแบบจำลองมี 3 ตัว ได้แก่ $I(0)$, β และ γ โดยเรากำหนด N_T คงที่ คือ 86000 และกำหนดค่าเริ่มต้นของการฟิตค่าเหล่านี้จากการฟิตเบื้องต้นที่แสดงในหัวข้อที่แล้วและให้ $\gamma = 0.1\beta$ (หรือ $R_0 = 10$) จากรูปจะเห็นได้ว่าผลการฟิตมีแนวโน้มที่ถูกต้องแต่ผลเชิงเลขยังไม่ถูกต้องมากนัก



รูปที่ 4 ผลการฟิตข้อมูลประเทศจีนด้วย (ก) 3 พารามิเตอร์และ N_T = 86000 คงที่ (ข) 4 พารามิเตอร์ (ค) 5 พารามิเตอร์ (ดูคำอธิบายในเนื้อหาบทความ) และ (ง) ผลการฟิตในรูป (ค) ที่แปลงข้อมูล $I(t)$ และ $R(t)$ กลับเป็น $N_C(t)$ และ $N_R(t)$ แล้ว

สำหรับในรูปที่ 4(ข) เราได้เพิ่มพารามิเตอร์อีกตัวสำหรับการฟิต คือค่า N_T ซึ่งการเพิ่ม N_T เป็นพารามิเตอร์นี้ทำให้ผลการพล็อตข้อมูล $S(t), I(t)$ และ $R(t)$ เปลี่ยนแปลงไปด้วย (ดูสมการ (5)-(7)) จากรูปที่ 4(ข) จะเห็นได้ว่า ผลการฟิต $I(t)$ เป็นที่น่าพอใจอย่างยิ่ง โดยการฟิตนี้ให้ค่า $N_T = 193908$ แต่ทั้งนี้ค่านี้มีค่ามากเกินไปอย่างมาก และทำให้ผลการฟิตสำหรับ $S(t)$ และ $R(t)$ ไม่ถูกต้อง

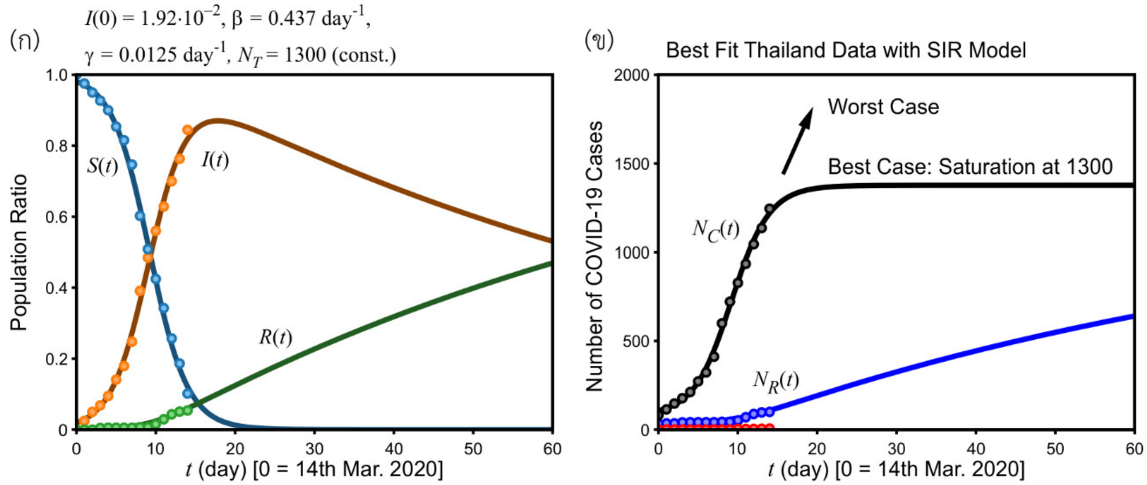
เนื่องจากได้ผลการจำลองต่างจากข้อมูลจริงอยู่มากทำให้ต้องปรับปรุงแบบจำลองจากสมมติฐานเพิ่มเติมที่ได้จากการพิจารณาสภาพความเป็นจริง คือในความเป็นจริง เมื่อเวลาผ่านไป การระบาดเพิ่มมากขึ้น คนทั่วไปจะระวังตัวมากขึ้นมิให้เป็นกลุ่มเสี่ยงซึ่งจะสะท้อนให้ เกิดอัตราส่วนการติดโรคในขั้นต้น R_0 ที่ลดลง ดังนั้นหากเราแทนค่าคงที่ β และ γ ด้วยฟังก์ชันที่แปรตามเวลา ก็น่าจะทำให้ได้ผลการฟิตที่ถูกต้องมากยิ่งขึ้น ดังนั้นเราจึงกำหนดให้ฟิตข้อมูลด้วยค่า β และ γ ที่ขึ้นกับเวลาแบบเชิงเส้นแทนที่จะเป็นค่าคงที่ (คือให้แต่ละตัวแปรขึ้นกับค่าคงที่สองตัวคือค่าเริ่มต้นและความชันของการเปลี่ยนแปลงค่าตามเวลา) ผลการฟิตตามสมมติฐานนี้แสดงในรูปที่ 4(ค) สำหรับการเปลี่ยนแปลงของ β และ γ และ R_0 ตามเวลาของข้อมูลประเทศจีนแสดงในรูปที่ 5(ก) และ 5(ข) ตามลำดับ โดยค่า R_0 ที่สูงมากในช่วงแรกในรูปที่ 5(ข) สะท้อนให้เห็นถึงความรุนแรงอย่างมากของการระบาดของโรคนี้



รูปที่ 5 (ก) กราฟแสดงการเปลี่ยนแปลงของค่า β และ γ ที่ทำให้ได้ผลการฟิตที่ดีที่สุด (รูปที่ 4(ค)) และ (ข) ค่า R_0 ที่ขึ้นกับเวลาในการระบาดของโรค

4.3 ข้อมูลของประเทศไทย

สำหรับข้อมูลการติดเชื้อโรคโควิด-19 ของประเทศไทยที่พล็อตแสดงในรูปที่ 1(ข) เราสามารถนำมาแปลงเป็นข้อมูล $S(t), I(t)$ และ $R(t)$ เพื่อนำมาฟิตกับแบบจำลอง SIR ในลักษณะเดียวกับข้อมูลของประเทศจีนได้ แต่เพื่อให้การฟิตถูกต้องยิ่งขึ้นเราจะต้องทำการนำจำนวนฐานออกเสียก่อน ทั้งนี้เพราะก่อนการระบาดครั้งใหญ่ในประเทศไทย เราได้มีผู้ติดเชื้อ (82 ราย) ผู้ที่รักษาหาย (35 ราย) และผู้เสียชีวิต (1 ราย) อยู่ก่อนแล้ว โดยรูปที่ 6(ก) แสดงข้อมูลที่แปลงแล้วนี้ (นับตั้งแต่วันที่ 14 มีนาคม พ.ศ. 2563 เป็นต้นไป) พร้อมกับผลการฟิตด้วยแบบจำลอง SIR โดยการฟิตนี้ เรากำหนดให้ $N_T = 1300$ ซึ่งเป็นค่าประชากรรวมที่เกี่ยวข้องกับการระบาดนี้*ที่น้อยที่สุดที่เป็นไปได้*โดยประมาณเทียบกับข้อมูลปัจจุบัน และการใช้พารามิเตอร์นี้คือกรณีที่ดีที่สุดที่เป็นไปได้ (คือมีจำนวนผู้ที่เกี่ยวข้องกับการระบาดนี้น้อยที่สุด)



รูปที่ 6 ผลการฟิตข้อมูลประเทศไทยด้วย (ก) 3 พารามิเตอร์และ $N_T = 1300$ คงที่
และ (ข) ผลการฟิตในรูป (ก) ที่แปลงข้อมูล $I(t)$ และ $R(t)$ กลับเป็น $N_C(t)$ และ $N_R(t)$ แล้ว

จากการฟิตข้อมูลประเทศไทยที่นำเสนอนี้ เราจะได้ว่าค่า $\beta = 0.437$ ต่อวันและ $\gamma = 0.0125$ ต่อวัน และจะได้ว่า $R_0 = \beta/\gamma = 34.9$ ซึ่งเป็นค่าที่สูงมาก และบ่งบอกว่า การระบาดนี้เกิดขึ้นอย่างรุนแรงมาก (เช่นเดียวกับประเทศจีนในช่วงแรก) ข้อสังเกตหนึ่งที่ได้จากการพิจารณาผลการฟิตคือ ค่า γ ที่ได้มีค่าน้อย ซึ่งสอดคล้องกับผลการฟิตข้อมูลประเทศจีนที่แสดงในรูปที่ 5(ก) เช่นกัน ทั้งนี้เราอาจบอกได้ว่า ในระยะแรกของการระบาด ผู้ติดเชื้อแต่ละคนต้องจะใช้เวลามากในการรักษาให้หายจากโรค (โดยเวลารักษาเฉลี่ยนี้เป็นส่วนกลับของค่า γ) และในช่วงเวลาต่อไปค่า γ นี้จะมีเพิ่มขึ้น ดังนั้นเราอาจบอกได้ว่า อัตราเร็วในการหายจากโรคที่ได้จากการประมาณนอกช่วงออกไป ($t \geq 20$ วัน) จะมีค่าน้อยกว่าความเป็นจริง และสำหรับการฟิตโดยให้ N_T เป็นพารามิเตอร์ที่นำมาใช้ฟิต หรือการกำหนดให้ β และ/หรือ γ มีค่าไม่คงที่ ด้วยนั้น เป็นการกำหนดที่เกินจำเป็นสำหรับข้อมูลการระบาดในประเทศไทยที่อยู่ในช่วงเริ่มต้น ดังนั้นผู้เขียนจึงยังมิได้นำมาพิจารณาในบทความนี้

ผลการฟิตข้อมูลประเทศไทยด้วยแบบจำลอง SIR สามารถถูกแปลงกลับมาเป็นข้อมูลจำนวนผู้ติดเชื้อสะสม $N_C(t)$ และจำนวนผู้ที่รักษาหายสะสม $N_R(t)$ ได้ดังแสดงในรูปที่ 6(ข) โดยในกรณีที่ดีที่สุดสำหรับข้อมูลประเทศไทยที่แสดงให้เห็นในรูปนี้ คือ กรณีที่การระบาดชะลอและหยุดลงในเวลาอันสั้น อันอาจเกิดได้จากการควบคุมโรคอย่างมีประสิทธิภาพและส่งผลให้มีจำนวนผู้ที่เกี่ยวข้องกับการระบาดนี้น้อยที่สุด (คือ $N_C(t \to \infty) = N_T = 1300$ ราย) โดยการคาดการณ์นี้จะเป็นลักษณะตรงกันข้ามกับการคาดการณ์อีกกรณีที่การระบาดเกิดขึ้นต่อไปในวงกว้างและทำให้มีจำนวนผู้ติดเชื้อเพิ่มขึ้นแบบทวีคูณไปเรื่อย ๆ ซึ่งเป็นกรณีที่แย่ที่สุด (Worst Case) โดยการเพิ่มแบบทวีคูณเป็นสิ่งที่พบในหลายประเทศเช่น อิหร่านและอิตาลี ในช่วงที่มีการระบาดในระยะแรก (100-10000 คน) จากแบบจำลอง SIR นี้ เราสามารถกล่าวได้ว่า กรณีที่แย่ที่สุด (ที่ไม่น่าเกิดขึ้นได้จริง) คือ การมีจำนวนผู้ที่เกี่ยวข้องกับการระบาดสูงสุดเท่ากับจำนวนประชากรในประเทศ โดยการศึกษาแบบจำลองในกรณีนี้จะเป็นการบอกจำนวนประชากรแต่ละกลุ่มในแต่ละช่วงเวลา ของการระบาด

การทำนายด้วยแบบจำลอง SIR นี้จะแม่นยำขึ้นเมื่อได้รับข้อมูลจำนวนผู้ที่ติดโรคต่อวันสูงสุด หรือค่าสูงสุดของ $I(t)$ ทั้งนี้เพราะว่า จุดเวลานี้เป็นจุดที่อัตราการเพิ่มของผู้ที่รักษาหายจากโรคมีค่าเท่ากับอัตราการลดของผู้ที่เสี่ยงต่อการติดโรค (ดูสมการ (2)-(4)) โดยหากค่า $I(t)$ ผ่านจุดสูงสุดแล้ว เราก็จะไม่จำเป็นต้องทำนายค่าสูงสุดของ $I(t)$ (ซึ่งขึ้นกับค่าของ N_T ด้วย) และทำให้การทำนายนี้ง่ายขึ้น

5. สรุปและอภิปรายผล

บทความนี้แสดงข้อมูลจำนวนผู้ติดเชื้อโรคโควิด-19 ที่กำลังระบาดอย่างมากในโลกและในประเทศไทย และจากนั้นจึงได้เสนอแบบจำลอง SIR ที่มักจะนำมาใช้ในการอธิบายการเกิดโรคระบาด โดยการวิเคราะห์เชิงคณิตศาสตร์ของสมการในแบบจำลอง SIR นี้ทำให้เราสามารถดึงพารามิเตอร์ที่สำคัญที่บ่งบอกระดับความรุนแรงของการระบาดออกมาได้ ข้อมูลของการระบาดในประเทศจีนได้ถูกนำมาศึกษาเพื่อแสดงเป็นข้อมูลอ้างอิงก่อน จากนั้นจึงทำการศึกษาข้อมูลจำนวนผู้ติดเชื้อในประเทศไทยและทำนายการเปลี่ยนแปลงของจำนวนผู้ติดเชื้อในประเทศไทยในอนาคต จากการศึกษาแบบจำลองนี้ทำให้เราสามารถคาดการณ์กรณีที่ดีที่สุดได้สำหรับประเทศไทย คือ มีจำนวนผู้ที่เกี่ยวข้องกับการระบาดรวมเพียง 1300 คน แต่สำหรับกรณีที่แย่ที่สุดที่อาจเกิดขึ้นได้นั้นจะขึ้นกับจำนวนผู้ติดเชื้อที่เพิ่มมากขึ้นในแต่ละวัน โดยเรายังไม่สามารถคาดการณ์จำนวนผู้ติดเชื้อสูงสุดได้อย่างแม่นยำ

สำหรับตัวเลข 1300 คนที่เสนอในบทความนี้ ดูเหมือนว่าจะเป็นค่าที่ต่ำเกินจริง เนื่องจากในสถานการณ์จริง ผู้ที่ติดเชื้อได้แพร่กระจายไปในหลายสิบจังหวัดทั่วทุกภาคในประเทศไทยแล้ว ดังนั้นจึงเป็นการยากที่ R_0 จะมีค่าลดลงในระยะเวลาอันสั้น แต่หากทุกคนในประเทศร่วมกันป้องกันและกักโรค มิให้เกิดการระบาดต่อไปอย่างรวดเร็วและต่อเนื่อง (ลด N_T) การระบาดนี้ก็จะลดระดับความรุนแรงและหยุดลงในที่สุด

เอกสารอ้างอิง

[1] https://en.wikipedia.org/wiki/2019-20_coronavirus_pandemic

[2] https://github.com/CSSEGISandData/COVID-19

[3] Kermack, W. O.; McKendrick, A. G. (1927). "A Contribution to the Mathematical Theory of Epidemics". Proceedings of the Royal Society A. 115 (772): 700–721.

[4] https://en.wikipedia.org/wiki/Compartmental_models_in_epidemiology

[5] Linge, S.; Langtangen, H. P. (2020) Programming for Computations – Python, Springer: 225–239.

[6] กระทรวงสาธารณสุข แถลงข่าวความคืบหน้าสถานการณ์ไวรัสโคโรนาสายพันธุ์ใหม่ 2019 (COVID-19) ในประเทศไทย ประจำวันที่ 27 มีนาคม พ.ศ. 2563

[7] https://www.scipy.org/

[8] Johansson, R. (2015) Numerical Python: A Practical Techniques Approach for Industry, A Press: 147–168.