

# การประมาณข้อมูลฐานพลังงานอ้างอิงที่น่าเชื่อถือสำหรับโรงงานอุตสาหกรรมในช่วงขยายขนาดกำลังการผลิต โดยวิธีการถดถอยของกระบวนการเกาส์เซียน

## Estimation of a Reliable Energy Baseline for Industrial Plants during the Expansion of Production Capacity by Gaussian Process Regression

โตมร สุนทรนภา

ภาควิชาวิศวกรรมไฟฟ้า คณะวิศวกรรมศาสตร์ มหาวิทยาลัยสยาม

38 ถนนเพชรเกษม แขวงบางหว้า เขตภาษีเจริญ กรุงเทพมหานคร 10160 E-mail: tomorn.soo@siam.edu

### บทคัดย่อ

บทความนี้เสนอการประมาณข้อมูลฐานพลังงานอ้างอิงที่น่าเชื่อถือสำหรับโรงงานผลิตน้ำตัวอย่าง ระหว่างปี 2016-2017 โดยใช้วิธีการถดถอยของกระบวนการเกาส์เซียน (Gaussian process regression: GPR) กับข้อมูลพลังงานของโรงงานในช่วงขยายขนาดกำลังการผลิต ระหว่างปี 2011-2016 เป็นข้อมูลฝึกสอน โดยตัวแปรสำคัญ ได้แก่ ค่าพลังงานไฟฟ้า MWh ค่าผลผลิต ลบ.ม.น้ำ และค่าน้ำดิบ ลบ.ม.น้ำ ในช่วงรายวัน นอกจากนี้ได้ทำการเปรียบเทียบผลการคำนวณกับวิธีการถดถอยเชิงเส้น (Linear regression: LR) และวิธี Multilayer perceptron (MLP) จากผลการทดลองพบว่า วิธี GPR ให้ผลค่าความไม่แน่นอนที่น้อยกว่า วิธี LR และ MLP ส่งผลให้เกิดความน่าเชื่อถือของแบบจำลองมากยิ่งขึ้น

**คำสำคัญ:** วิธีการถดถอยของกระบวนการเกาส์เซียน ข้อมูลฐานพลังงานอ้างอิง วิธีการถดถอยเชิงเส้น เพอร์เซ็ปตรอนหลายชั้น

### Abstract

This paper proposes the estimation of a reliable energy baseline for the sample water treatment plant during 2016-2017. By using the Gaussian process regression: GPR) method calculates with the factory energy data during the expansion capacity period produced during 2011-2016 to be the training data. Key variables are energy unit, MWh, production, m<sup>3</sup> and raw water, m<sup>3</sup> in daily cycle. In addition, the author has made a comparison of GPR with linear regression (LR) and multilayer perceptron (MLP). Based on testing, the results show that the GPR method gives less uncertainty than the LR and MLP method, resulting in more model reliability.

**Keywords:** Gaussian Process Regression, Energy Baseline, Linear Regression, Multilayer Perceptron

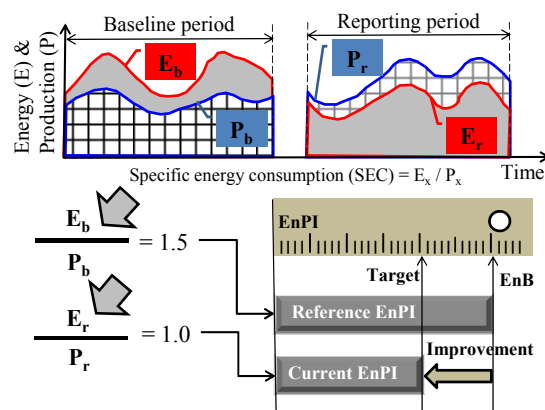
### 1. บทนำ

ปัจจุบันการอนุรักษ์พลังงานสำหรับอาคาร และโรงงานควบคุม ได้มีบทบาทสำคัญสำหรับการจัดการสมัยใหม่ที่มีความตระหนักทางด้าน

การใช้พลังงาน เพื่อให้บรรลุเป้าหมายในการประหยัดพลังงาน และลดปริมาณก๊าซคาร์บอนไดออกไซด์ที่ปลดปล่อยออกมาจากการผลิตพลังงานไฟฟ้าของโรงจักรไฟฟ้า

การอนุรักษ์พลังงานโดยใช้พลังงานทดแทน (Renewable energy) เช่นพลังงานแสงอาทิตย์ พลังงานชีวมวล พลังงานลม และอื่นๆ ได้กำลังมีการพัฒนาในประสิทธิภาพ และได้มีการสนับสนุนเงินลงทุนรวมถึงผลตอบแทนในรูปแบบต่างๆจากหน่วยงานภาครัฐ เช่นโครงการรับซื้อไฟฟ้าจากการผลิตพลังงานไฟฟ้าจากเซลล์แสงอาทิตย์ (Solar energy) และพลังงานลม (Wind energy) ยิ่งกว่านั้นยังได้มีการสนับสนุนให้ภาคเอกชน และภาครัฐเข้ามามีส่วนร่วมในการจัดการพลังงานผ่านโครงการบริษัทจัดการพลังงาน หรือที่เรียกย่อๆว่า ESCOs (Energy Service Companies) เพื่อให้เกิดการนำเทคโนโลยีมาใช้ และเกิดผลประหยัดที่เป็นจริงในทางปฏิบัติให้มากยิ่งขึ้น

สำหรับประเทศไทย และต่างประเทศ การจัดการพลังงานผ่าน ESCOs จำเป็นต้องมีการเข้าไปตรวจวัด และทำการประเมินผลประหยัด (Measurement & Verification) ตามขั้นตอนแนะนำใน [1] ซึ่งในส่วนของ การประเมินผลประหยัดจะต้องมีการสร้างข้อมูลฐานพลังงานอ้างอิง (Energy baseline: EnB) และสร้างตัวชี้วัดสมรรถนะด้านพลังงาน (Energy performance indicators: EnPI) ดังอ้างอิงใน ISO 50006:2014 [2] เพื่อนำมาเปรียบเทียบผลประหยัดโครงการช่วงก่อนใช้มาตรการ (Pre-retrofit) และช่วงหลังใช้ (Post-retrofit) ซึ่งแสดงแนวทางได้ดังรูปที่ 1



รูปที่ 1 แนวทางการหาผลประหยัดพลังงานด้วย EnPI ตาม ISO 50006:2014

ทั่วไปการสร้าง EnB จะใช้วิธี LR ตามมาตรฐาน ISO 50001:2011 [3] หรือ ASHRAE Guideline 14-2002 [4] และตามคำแนะนำใน ISO 50006:2014 ให้สามารถใช้ Nonlinear regression หรือ วิถีทาง Engineering based theory ได้เช่นกัน โดยข้อเสียของวิธี LR คือไม่เหมาะกับข้อมูลพลังงานที่มีความสัมพันธ์ในลักษณะ Nonlinear อย่างเช่น โรงงานผลิตน้ำตัวอย่าง ที่มีการขยายกำลังการผลิต มีการทดสอบเพิ่ม-ลด แรงดันอัดเข้าเส้นท่อ และมีการเปลี่ยนแปลงโครงสร้างการใช้พลังงานของอุปกรณ์ ส่งผลให้บางปีได้ข้อมูลพลังงานที่มีความสัมพันธ์กับผลผลิตอย่างไม่เป็นเชิงเส้น นำมาซึ่งการสร้าง EnB ที่มีลักษณะชี้้นำผิด (Misleading) หรือเป็น EnB ที่ไม่น่าเชื่อถือได้

สำหรับงานวิจัยทางการสร้าง EnB ได้มีการนำวิถีทางเครือข่ายประสาทเทียม (Artificial Neural Networks: ANNs) มาทำนาย EnB เช่น วิธี Multilayer Perceptron (MLP) หรือ Feed-forward Neural Networks (FNN) [5] ได้ทำการศึกษาร่วมกับวิธี Support Vector Machine (SVM) และวิธี Radial Basis Function Networks (RBFN) ซึ่งผลการวิจัยพบว่าวิธี MLP ให้ผลความไม่แน่นอน (Uncertainty) ที่น้อยที่สุด แต่อย่างไรก็ตามวิถีทางด้าน ANNs จำเป็นต้องมีข้อมูลที่มากพอในขบวนการ Training เพื่อให้ได้ผลการทำนายที่ดี ดังนั้นจึงได้มีการนำเสนอวิถีทางด้าน GPR ที่มีความสามารถทำนายข้อมูลในลักษณะ Nonlinear ที่มีจำนวนไม่มากได้ และให้ผลการทำนายที่ดีกว่าวิธี LR หรือ ANNs ดังงานวิจัย [6]

งานวิจัยนี้จึงได้นำเสนอการประยุกต์ใช้วิธี GPR กับข้อมูลพลังงานของโรงงานผลิตน้ำมหาสวัสดิ์ ที่มีลักษณะของการขยายกำลังการผลิตระหว่างปี 2011 -2016 มาใช้เป็น Training data และทำการทำนาย EnB เพื่อนำมาใช้เปรียบเทียบกับข้อมูลพลังงาน ระหว่างปี 2016-2017 เพื่อเปรียบเทียบผลประหยัดของมาตรการอนุรักษ์พลังงาน หรือลักษณะการใช้พลังงาน ช่วง Pre-retrofit และ Post-retrofit

## 2. Gaussian Process Regression (GPR) Models

ในส่วนนี้ขอกรีนนำหลักการพื้นฐานของการถดถอยกระบวนการเกาส์เซียนเพื่อให้เกิดความเข้าใจและนำไปประยุกต์ใช้กับปัญหาต่อไป

### 2.1 พื้นฐานของ GPR

กำหนดให้มีชุดข้อมูลฝึกสอน  $\mathcal{D} = (X, y)$  ประกอบด้วยเวกเตอร์อินพุตจำนวน  $N$  และมีขนาดมิติเป็น  $D$  นั่นคือ  $X = \{x_n\}_{n=1}^N$  เมื่อ  $x_n \in \mathbb{R}^D$  และมีเอาต์พุตที่สอดคล้องกับอินพุตเป็น  $y = \{y_n\}_{n=1}^N$  เมื่อ  $y_n \in \mathbb{R}$  โดยความสัมพันธ์ระหว่างอินพุต-เอาต์พุต สามารถเขียนได้ดังสมการ (1)

$$y_n = f(x_n) + \varepsilon_n \quad (1)$$

เมื่อ  $f(\cdot)$  คือฟังก์ชันแฝง (Latent function) และ  $\varepsilon_n$  คือ สัญญาณรบกวนแบบเกาส์ที่มีการเพิ่มค่าเฉลี่ยศูนย์เข้ากับค่าความแปรปรวน  $\sigma_n^2$  นั่นคือ  $\varepsilon_n \sim \mathcal{N}(0, \sigma_n^2)$  โดยสมมติฐานว่า  $\{\varepsilon_n\}_{n=1}^N$  เป็นลำดับที่เป็นอิสระต่อกันและมีการกระจายที่เหมือนกัน (Independent and identically

distributed: i.i.d) วัตถุประสงค์หลักคือการจำลองแบบฟังก์ชัน  $f(\cdot)$  ซึ่งทำการจับคู่อินพุต  $x$  กับค่าเอาต์พุต  $y$  ที่สอดคล้อง โดยสมมติฐานสำคัญใน GPR คือ ทุกเซตของค่าฟังก์ชันจะต้องมีการกระจายแบบเกาส์เซียนหลายตัวแปร [7] ดังสมการ (2)

$$p(f | x_1, x_1, \dots, x_n) = \mathcal{N}(0, K) \quad (2)$$

จากสมการ (2)  $f = [f(x_1), f(x_2), \dots, f(x_n)]^T$  และ  $0$  เป็นเวกเตอร์มิติ  $N \times 1$  ที่มีค่าเป็น 0 ทั้งหมด และ  $K$  เป็นเมตริกซ์ Covariance ที่มีสมาชิก  $K_{ij} = k(x_i, x_j)$  เป็นค่าฟังก์ชันเคอเนลสำหรับทุกๆ คู่ของอินพุตฝึกสอน สำหรับงานวิจัยนี้ได้ใช้ ฟังก์ชันเคอเนลแบบ Squared exponential (SE) ที่สามารถกำหนดค่าเองได้ ดังสมการ (3)

$$k_s(x_i, x_j) = v_0^2 \exp\left[-\frac{1}{2} \sum_{d=1}^D \left(\frac{x_{id} - x_{jd}}{l_d}\right)^2\right] \quad (3)$$

โดยฟังก์ชันเคอเนลถูกจำแนกโดย Hyperparameters กำหนดสัญลักษณ์เป็น  $\Theta$  โดยมีค่า  $v_i$  สำหรับ  $i \in \{0, 1, 2\}$  เป็นค่าความแปรปรวนของสัญญาณและเป็นจำนวนความแปรปรวนของฟังก์ชันแฝงจากค่าเฉลี่ยของมัน โดย  $l_d$  แทน Characteristic length scale สำหรับแต่ละมิติของอินพุต ซึ่งเป็นตัวหาความสัมพันธ์ที่สำคัญของแต่ละตัวแปรอินพุตในการประมาณเอาต์พุตเป้าหมาย ค่า Characteristic length scale ที่มีค่าเล็กกว่าจะแสดงเป็นนัยว่าขนาดมิติของอินพุตที่สอดคล้องจะมีผลกระทบมากขึ้นต่อเอาต์พุต ซึ่งหมายถึงการมีความสัมพันธ์มากขึ้น

การกำหนดฟังก์ชันเคอเนลเป็นส่วนสำคัญสำหรับ GPR เพราะเป็นการเข้ารหัส Prior assumption เกี่ยวกับคุณสมบัติของฟังก์ชันแฝง เช่น ความเรียบ ความเป็นคาบเวลา หรือ Non-stationary ที่เราจะใช้สร้างแบบจำลอง โดยฟังก์ชันเคอเนล

เอาต์พุตในสมการ (1) ได้สมมติว่ามีการผสม Gaussian noise ด้วยค่าความแปรปรวน  $\sigma_n^2$  ดังนั้นเราสามารถจัดรูปในเทอมของฟังก์ชันเคอเนลได้เป็น

$$k(x_i, x_j) = k_s(x_i, x_j) + \sigma_n^2 \delta_{ij} \quad (4)$$

เมื่อ  $\delta_{ij}$  คือ Kronecker delta ซึ่งกำหนดให้มีค่าเป็น 1 เมื่อถ้า  $i = j$  และกำหนดเป็น 0 ในกรณีอื่น จากเงื่อนไขการกระจายของฟังก์ชัน  $y$  เมื่อกำหนดให้มีค่าฟังก์ชันแฝง  $f$  และอินพุต  $X$  สามารถเขียนได้เป็น

$$p(y | f, X) = \mathcal{N}(y, \sigma_n^2 I) \quad (5)$$

เมื่อ  $I$  เป็นเมตริกซ์หนึ่งหน่วยมิติ  $N \times N$  โดยการใช้สมการ (2) และ (5) ความน่าจะเป็น Marginal distribution ของ  $y$  สามารถเขียนได้เป็น

$$p(y | X) = \int p(y | f, X) p(f | X) df = \mathcal{N}(0, K + \sigma_n^2 I) \quad (6)$$

จาก (6) ค่า Marginal log-likelihood ของ  $y$  สามารถเขียนได้เป็น

$$\log p(y | X, \Theta) = -\frac{1}{2} y^T (K + \sigma_n^2 I)^{-1} y - \frac{1}{2} \log |K + \sigma_n^2 I| - \frac{N}{2} \log 2\pi \quad (7)$$

เมื่อ  $|\cdot|$  คือค่าดีเทอร์มิแนนต์ของเมตริกซ์ สำหรับตัวแปร Hyperparameters ได้ถูก Optimized โดยทำการหาค่า Maximized value

ของ สมการ Marginal log-likelihood function ใน (7) โดยการหาค่า Gradient ของ (7) เทียบกับค่า Hyperparameter  $\Theta$  ที่  $i$  แสดงได้ดัง (8)

$$\frac{\partial \log p(y | X, \Theta)}{\partial \theta_i} = -\frac{1}{2} \text{tr} \left( (K + \sigma_n^2 I)^{-1} \frac{\partial (K + \sigma_n^2 I)}{\partial \theta_i} \right) + \frac{1}{2} y^T (K + \sigma_n^2 I)^{-1} \frac{\partial (K + \sigma_n^2 I)}{\partial \theta_i} (K + \sigma_n^2 I)^{-1} y \quad (8)$$

ซึ่งสามารถนำไปทำการ Optimization แบบวิธี Gradient ได้เพื่อหาค่าที่ดีที่สุดของ Hyperparameters ที่ Maximize (7) และฟังก์ชันเป้าหมายไม่จำเป็นต้อง Convex เป็นผลให้วิธี Gradient อาจ Converge ในการหาค่าเฉลยแบบ Local optimum วิธีที่เป็นไปได้สำหรับปัญหานี้ ทำโดยการเริ่มหาค่าด้วย Gradient based searches และต่อมาทำการเลือก Optimal point ที่ให้ค่า Marginal log-likelihood ที่ใหญ่ที่สุด อย่างไรก็ตามเราควรพิจารณาในการคำนวณฟังก์ชัน Log-likelihood และการใช้วิธี Gradient จะประกอบด้วยการทำงานอินเวอร์สเมตริกซ์  $K + \sigma_n^2 I$  ด้วยขนาดมิติ  $N \times N$  ซึ่งต้องการเวลาคำนวณเป็น  $O(N^3)$  ดังนั้น วิธีการใช้งาน GPR จะเหมาะสมกับปัญหาที่มีตัวอย่างข้อมูลฝึกสอนประมาณ 2-3 พัน ตัวอย่าง แต่สำหรับข้อมูลฝึกสอนที่มากกว่านี้สามารถใช้วิธีประมาณ Sparse approximation กับ ปัญหา GPR ปกติ โดยเลือกข้อมูลย่อยขนาดเล็กๆ ที่ใช้แทนข้อมูลฝึกสอนตัวอย่างสามารถนำมาใช้งานได้อย่างมีประสิทธิภาพ

เมื่อทำการหา Optimal hyperparameters ได้แล้ว ต่อมาเราจะพิจารณาค่า Joint distribution ของ  $y$  และ  $y_*$  ดัง (9)

$$p(y, y_* | X, x_*, \Theta) = N \left( \begin{bmatrix} 0 \\ 0 \end{bmatrix}, \begin{bmatrix} K + \sigma_n^2 I & k_* \\ k_*^T & k_{**} + \sigma_n^2 \end{bmatrix} \right) \quad (9)$$

เมื่อ  $k_* = [k(x_1, x_*), \dots, k(x_N, x_*)]^T$  และ  $k_{**} = k(x_*, x_*)$  เป้าหมายหลักของ GPR คือการหา Predictive distribution สำหรับเวกเตอร์อินพุตใหม่  $x_*$  และเมื่อใช้ Marginal joint distribution ใน (9) ผ่านข้อมูลฝึกสอนเอาต์พุต  $y$  เราจะได้ค่า Predictive distribution ของเอาต์พุตทดสอบ  $y_*$  ซึ่งเป็นการกระจายแบบเกาส์เซียน นั่นคือ  $p(y_* | X, y, x_*, \Theta) = N(\mu_*, \Sigma_*)$  ด้วยค่า Mean และ Covariance ดังนี้

$$\mu_* = k_*^T (K + \sigma_n^2 I)^{-1} y \quad (10)$$

$$\Sigma_* = \sigma_n^2 + k_{**} - k_*^T (K + \sigma_n^2 I)^{-1} k_* \quad (11)$$

จาก (10) สังเกตได้ว่าค่า Mean  $\mu_*$  ของค่า Predictive distribution ที่หาได้จะเป็นการรวมแบบเชิงเส้นของ Noisy output  $y$  และวิธีนี้เป็นการประมาณที่มีประสิทธิภาพของเอาต์พุตทดสอบ นอกจากนี้ ค่า Variance ของ Predictive distribution ใน (11) เป็นการวัดความแปรปรวน โดยการใส่ (10) และ (11) ค่า Confidence interval  $100(1-\alpha)\%$  คำนวณได้เป็น

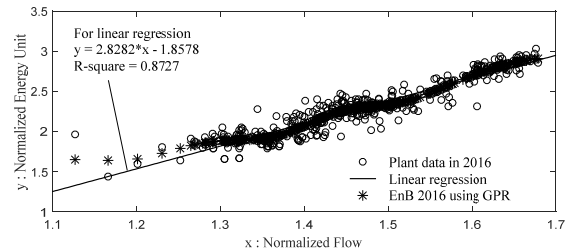
$$[\mu_* - z_{(1-\alpha)/2} \Sigma_*, \mu_* + z_{(1-\alpha)/2} \Sigma_*] \quad (12)$$

เมื่อ  $\alpha \in [0, 1]$  แทนระดับ Confidence level และ  $z_{(1-\alpha)/2}$  เป็นค่าวิกฤตของการกระจาย ค่า Confidence interval ให้ช่วงของค่าซึ่งน่าจะบรรจุไปด้วยค่าจริงของเอาต์พุตทดสอบ โดยเฉพาะอย่างยิ่งค่า Variance ที่มีค่าเล็กจะส่งผลทำให้ช่วง Confidence interval มีค่าแคบ และยังเป็นการบ่งชี้ว่า

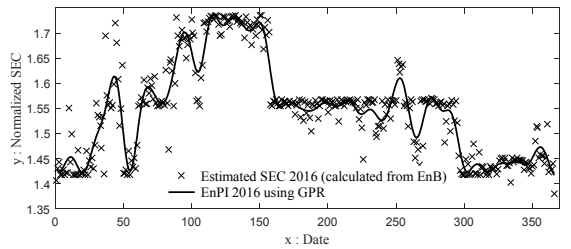
การประมาณมีความแม่นยำมากขึ้นสำหรับเอาต์พุตทดสอบ ดังนั้น GPR ให้การทำนายด้วยการกระจายความน่าจะเป็นที่ได้เปรียบกว่าวิธี Support Vector Machine (SVM) หรือ Neural Network (NN) และวิธีการแบบ Non-probabilistic machine learning อื่นๆ

### 3. ผลการประมาณ EnB และ EnPI ด้วย GPR

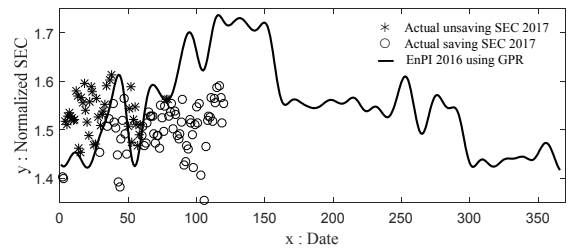
การประเมินพลังงานในปี 2017 จะใช้ข้อมูลพลังงานไฟฟ้า และปริมาณน้ำสูบน้ำในปี 2016 มาใช้สร้าง EnB ได้ผลดังรูปที่ 2 และเนื่องจากความสัมพันธ์ของพลังงานไฟฟ้าที่ใช้มีความสัมพันธ์แบบไม่เชิงเส้น กับปริมาณน้ำสูบน้ำเนื่องจากผลของแรงดันส่งมอบ (Delivery pressure) ดังนั้น งานวิจัยนี้จึงได้นำ EnB มาคำนวณค่าพลังงานจำเพาะต่อปริมาณผลผลิต SEC ซึ่งเป็น EnPI ตัวหนึ่งได้ผลดังรูปที่ 3 ผู้วิจัยได้เขียนโปรแกรมเปรียบเทียบผลประหยัดเพื่อทำการประเมินการใช้พลังงานในปี 2017 โดยแสดงให้เห็นถึงวันที่ประหยัด และไม่ประหยัด แสดงได้ดังรูปที่ 4 และผลประหยัดสะสมของค่า SEC แสดงได้ดังรูปที่ 5



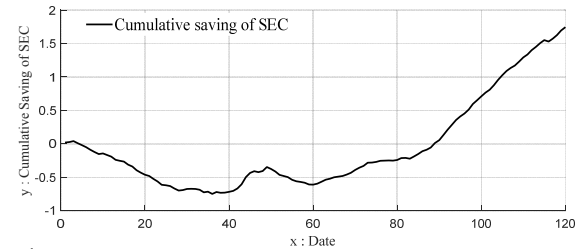
รูปที่ 2 ผลการประมาณ EnB รายวันของปี 2016 สำหรับใช้เทียบปี 2017



รูปที่ 3 ผลการประมาณ EnPI รายวันของปี 2016



รูปที่ 4 ผล Validation EnPI รายวันของปี 2016 (Reference) เทียบกับปี 2017 (Actual)



รูปที่ 5 ผลประหยัดสะสมของ SEC รายวัน (1 Jan - 30 Apr 2017) เทียบกับ EnPI 2016

ตารางที่ 1 เปรียบเทียบผลค่า Uncertainty ของแบบจำลอง EnB ที่นำเสนอ

Year	R <sup>2</sup> (R-Square)			CV-RMSE (%)			NMBE (%)		
	LR	MLP	GPR	LR	MLP	GPR	LR	MLP	GPR
2011	0.7161	0.9634	0.9973	5.5139	6.0324	5.4564	0.0027	-0.0257	-1.365E-14
2012	0.9253	0.9950	0.9994	4.0762	4.5843	4.0215	-0.0011	0.1294	-4.099E-14
2013	0.8800	0.9688	0.9751	5.9141	6.2420	5.2573	0.0002	-0.0533	2.4176E-5
2014	0.3157	0.9342	0.9998	8.7786	9.2554	8.7786	0.0032	0.6159	-1.0375E-4
2015	0.7972	0.9311	0.9956	5.9513	6.6499	5.7811	-0.0003	0.1472	6.3704E-5
2016	0.8727	0.9453	0.9770	5.1585	6.0324	4.6125	-0.0019	0.0154	9.2879E-5

และเพื่อทำการยืนยันถึงความสามารถของ GPR ในการประมาณฟังก์ชันที่มีความซับซ้อนและไม่เป็นเชิงเส้น ที่เหนือกว่าวิธีอื่น เช่น LR หรือวิธีทางเครือข่ายประสาทเทียม ซึ่งผู้วิจัยได้ทำการศึกษา การหา EnB ด้วยวิธี Multilayer perceptron (MLP) โดยขอนำผลการทดลอง มาใช้ร่วมประกอบการเปรียบเทียบ แสดงได้ดังตารางที่ 1

ค่า Uncertainty นี้ได้นำมาจากคำแนะนำสำหรับการสร้างแบบจำลอง EnB ตาม ASHRAE Guideline 14-2002 ดังแสดงในตารางที่ 2 โดยประกอบไปด้วยค่า Coefficient of Variation of the Root Mean Square Error (CV-RMSE) และ Net Mean Bias Error (NMBE) ดังสมการที่ (13) และ (14) โดย  $y_i$  คือข้อมูลตัวอย่าง  $\hat{y}_i$  คือข้อมูลจากการ-

$$CV-RMSE = 100 \times \left[ \frac{\sum (y_i - \hat{y}_i)^2}{(n-p)} \right]^{1/2} / \bar{y} \quad (13)$$

$$NMBE = \frac{\sum (y_i - \hat{y}_i)}{(n-p) \times \bar{y}} \times 100 \quad (14)$$

-ประมาณ  $\bar{y}$  คือข้อมูลเฉลี่ย  $n$  คือจำนวนข้อมูลศึกษา และ  $p$  คือจำนวนพารามิเตอร์ของแบบจำลอง ในที่นี้ใช้เฉพาะค่า Flow เป็นพารามิเตอร์ตัวแปรต้น ดังนั้น  $p=1$  จากตารางที่ 2 เป็นการเปรียบเทียบรายเดือน และรายชั่วโมง สำหรับโรงงานใดๆจะเลือกวิเคราะห์แบบใด นั้นขึ้นกับมิเตอร์วัดพลังงานที่ใช้ อาจใช้ Bill ค่าไฟจากการไฟฟ้าเป็นรายเดือนได้ หรือ หากมิเตอร์มิเตอร์ใช้ อาจวัดเป็นรายชั่วโมง หรือนำมารวมเป็นรายวันได้ โดยผู้วิจัยเลือกใช้ข้อมูลสรุปเป็นรายวัน เนื่องจากมีความสะดวกมากกว่ารายชั่วโมง และสามารถเห็นรายละเอียดได้มากกว่ารายเดือน

#### 4. อภิปราย

จากผลการทดลองจะเห็นได้ว่าในปี 2014 มีความไม่เป็นเชิงเส้นสูงมาก โดย R<sup>2</sup> มีค่าต่ำสุดที่ 0.3157 ด้วยวิธี LR ส่วนวิธี MLP และ GPR สามารถทำให้เอาต์พุตมีความเป็นเชิงเส้นได้ดีมาก ส่วนค่า Error ที่มาจาก CV-RMSE นั้น GPR มีค่าน้อยที่สุด และใกล้เคียงกับ LR ส่วน NMBE นั้น GPR มีค่าน้อยกว่าวิธีอื่นอย่างเห็นได้ชัดในระดับร้อยละ 10<sup>-14</sup>

ตารางที่ 2 ค่า Uncertainty แนะนำสำหรับ EnB ตาม ASHRAE Guideline 14-2002

EnB model uncertainty	Monthly	Hourly
CV-RMSE	15%	30%
NMBE	5%	10%

#### 5. สรุป

การประมาณข้อมูลฐานพลังงานอ้างอิงของโรงงานผลิตน้ำด้วยวิธี GPR ได้นำเสนอถึงความสามารถที่นำไปใช้ในการประเมินพลังงานรายปีสำหรับโรงงานอุตสาหกรรมทั่วไป ได้อย่างมีประสิทธิภาพ และมีความน่าเชื่อถือ เหนือกว่าวิธี LR หรือ MLP และสามารถนำวิธีการนี้ไปใช้กับข้อมูลของโรงงานในลักษณะไม่เชิงเส้น หรือมีความซับซ้อน โดยเฉพาะกรณีที่โรงงานมีการขยายกำลังการผลิตได้เป็นอย่างดี

#### เอกสารอ้างอิง

- [1] M&V Guidelines: Measurement and Verification for Performance Based Contracts Ver.4.0, prepared for the U.S. Department of Energy Federal Energy Management Program; November 2015.
- [2] ISO 50006:2014, Energy management systems, measuring energy performance using energy baselines and energy performance indicators, General principles and guidance; 2014.
- [3] ISO 50001:2011, Energy management systems, Requirements with guidance for use; 2011.
- [4] ASHRAE Guideline 14-2002: Measurement of Energy and Demand Savings, ASHRAE Inc, Atlanta, GA; 2002.
- [5] Oludolapo Akanni Olanrewaju et al., "Comparison of Artificial Intelligence Techniques for Energy Consumption Estimation," 16th annual IEEE Electrical Power and Energy Conference (EPEC), 2016, Ottawa, Canada; 12-14 October 2016.
- [6] Joseph Carpenter et al., "A Comparison of Gaussian Process Regression and Change-point Regression for the Baseline Model in Industrial Facilities," 2016 Building Performance Modeling Conference, Salt Lake City, UT; August 8-12, 2016
- [7] C.E. Rasmussen, C.K. Williams, "Gaussian Processes for Machine Learning," MIT Press, Cambridge, MA, 2006.



ดร. สุนทรนภา รัชนีปัญญามหาบัณฑิต สาขาวิศวกรรมไฟฟ้า จากจุฬาลงกรณ์มหาวิทยาลัย ปี 2545 สนใจงานวิจัยด้านอนุรักษ์พลังงาน และการเรียนรู้ด้วยเครื่อง