การประยุกต์ใช้วิธีการแปลงเพื่อประเมินภัยแล้ง

Application of Transformation Techniques to Evaluate Drought

ธนโชติ ไชยโต[1,2*] มานัดถุ์ คำกอง[2] และ พุฒิพงษ์ พุกกะมาน[2]

Tanachot Chaito[1,2*], Manad Khamkong[2] and Putipong Bookkamana[2]

[1]*ศูนย์วิจัยวิทยาศาสตร์พหุวิทยาการ คณะวิทยาศาสตร์ มหาวิทยาลัยเชียงใหม่*

[2]*ภาควิชาสถิติ คณะวิทยาศาสตร์ มหาวิทยาลัยเชียงใหม่*

[1]*Multidisciplinary Science Research Centre, Faculty of Science, Chiang Mai University*

[2]*Department of Statistics, Faculty of Science, Chiang Mai University*

Received : 2 March 2016

Accepted : 2 June 2016

Published online : 21 June 2016

บทคัดย่อ

การศึกษาครั้งนี้มีวัตถุประสงค์เพื่อเปรียบเทียบรูปแบบการแปลงข้อมูล และหาวิธีการแปลงข้อมูลที่เหมาะสมเพื่อนำมาประยุกต์ใช้ในการหาค่าดรรชนีความแห้งแล้งของฝนที่ต่างจากค่าปรกติในการประเมินระดับความแห้งแล้ง ผลการศึกษาพบว่า ในส่วนของการจำลองวิธีการแปลงที่เหมาะสมมากที่สุดในการแปลงข้อมูลที่มีการแจกแจงแกมมาให้มีการแจกแจงใกล้เคียงกับการแจกแจงปรกติคือ วิธีการแปลงข้อมูลแบบถอดรากที่สี่ สำหรับข้อมูลที่มีการแจกแจงล็อกนอร์มอลวิธีการแปลงข้อมูลที่เหมาะสมมากที่สุดคือ วิธีการแปลงข้อมูลแบบลอการิทึม และสำหรับข้อมูลที่มีการแจกแจงไวบูลวิธีการแปลงข้อมูลที่เหมาะสมมากที่สุดคือ วิธีการแปลงข้อมูลแบบ Haynes ในส่วนของข้อมูลจริงผู้วิจัยได้เลือกสถานีตรวจวัดปริมาณน้ำฝนของอำเภอบ้านโฮ่งและอำเภอลี้ จังหวัดลำพูน ในการประเมินความแห้งแล้งพบว่า น้ำฝนตามฤดูกาลของประเทศไทยที่อำเภอบ้านโฮ่งมีการแจกแจงไวบูล ส่วนอำเภอลี้มีการแจกแจงแกมมา อำเภอบ้านโฮ่งมีฝนแล้งรุนแรงมากที่สุด ในปี พ.ศ. 2548 และ 2557 และอำเภอลี้มีฝนแล้งรุนแรงมากที่สุด ในปี พ.ศ. 2523 2547 2548 และ 2557

คำสำคัญ : การแปลงข้อมูล ดรรชนีความแห้งแล้งของฝนที่ต่างจากค่าปรกติ

*Corresponding author. E-mail : tanachot.boy@gmail.com

Abstract

The objective of this research was to compare data transformation methods, as well as to find the best data transformation technique for applying the standardized precipitation index to evaluate drought. Results of this study were as follows: part I was simulated data, Fourth-Root transformation was the best normal approximate transformation of gamma data. The Logarithm transformation was the best normal approximate transformation of lognormal data. The Haynes transformation was the best normal approximate transformation of Weibull data. Part II was real data, based on rain gauging stations of Ban Hong and Lee, Lamphun, Thailand to evaluate drought. Weibull distribution was found to be the most appropriate distribution to represent the Thai season rainfall amount in Ban Hong. Gamma distribution was found to be the most appropriate distribution to represent the Thai season rainfall amount in Lee. Ban Hong had extremely dry conditions in 2005 and 2014. Lee had extremely dry conditions in 1980, 2004, 2005 and 2014.

Keywords Transformation, Standardized Precipitation Index

บทนำ

ภัยแล้งเป็นภัยธรรมชาติที่เกิดจากการขาดแคลนน้ำในพื้นที่ใดพื้นที่หนึ่งเป็นเวลานาน จนก่อให้เกิดความแห้งแล้งและส่งผลกระทบต่อเศรษฐกิจและสังคม สำหรับภัยแล้งในประเทศไทยส่วนใหญ่เกิดจากฝนแล้งและทิ้งช่วง ซึ่งฝนแล้งสามารถเกิดได้ 2 ช่วง (Thai Meteorological Department, 2015) ได้แก่ ช่วงแรกเริ่มต้นจากฤดูหนาวต่อเนื่องถึงฤดูร้อนโดยเริ่มประมาณกลางเดือนตุลาคมเป็นต้นไป ซึ่งจะมีปริมาณน้ำฝนลดลงจนกระทั่งเข้าสู่ฤดูฝนในช่วงกลางเดือนพฤษภาคมของปีถัดไป ส่วนช่วงที่สองเริ่มประมาณปลายเดือนมิถุนายนถึงเดือนกรกฎาคมซึ่งจะมีฝนทิ้งช่วงเกิดขึ้น

การประเมินระดับความแห้งแล้งมีหลายวิธี โดยส่วนใหญ่มักพิจารณาในรูปแบบของดรรชนีที่บ่งชี้ระดับความรุนแรงของสภาวะความแห้งแล้ง ดรรชนีที่บ่งชี้ระดับความรุนแรงของสภาวะความแห้งแล้งที่ได้รับความนิยมและนำไปใช้อย่างแพร่หลายได้แก่ ดรรชนีความแห้งแล้งของฝนที่ต่างจากค่าปรกติ (Standardized Precipitation Index: SPI) ซึ่งได้พัฒนาขึ้นจากแนวคิดของ Mckee (1993) เพื่อเฝ้าดูสภาวะแห้งแล้งในช่วงเวลาต่าง ๆ ที่กำหนดโดยดูจากปริมาณน้ำฝนสะสมในแต่ละช่วงเวลาที่สนใจ สำหรับการศึกษาเพื่อหาค่าดรรชนี SPI จะต้องใช้ปริมาณน้ำฝนรวมเป็นหลักโดยพิจารณาใช้ฟังก์ชันการแจกแจงสะสม (Cumulative Distribution Function : CDF) ของปริมาณน้ำฝนสะสมแล้ว**ทำการแปลง (Transform)** ให้เป็นค่าปรกติมาตรฐานหรือดรรชนี SPI ที่สามารถบอกถึงระดับความชุ่มชื้นและความแห้งแล้งของปริมาณน้ำฝนในแต่ละพื้นที่ (Agro - meteorological Academic Group, 2011) โดยการแจกแจงของปริมาณน้ำฝน Mckee ได้กำหนดปริมาณน้ำฝนโดยทั่วไปจะมีการกระจายในรูปแบบการแจกแจงแกมมา (Gamma Distribution) แต่ในบางครั้งการแจกแจงของปริมาณน้ำฝนไม่ได้มีการแจกแจงแกมมาเสมอไป Zhang (2009) ได้กล่าวว่า การแจกแจงล็อกนอร์มอล (Lognormal Distribution) เป็นการแจกแจงที่ดีที่สุดสำหรับการคำนวณค่าดรรชนี SPI ในพื้นที่ลุ่มแม่น้ำเพิร์ล (Pearl) ในสาธารณรัฐประชาชนจีน และต่อมา Yusof และ Hui-Mean (2012) ได้กำหนดปริมาณน้ำฝนมีการกระจายในรูปแบบการแจกแจงไวบูล (Weibull Distribution) เนื่องจากการแจกแจงไวบูลมีลักษณะการแจกแจงแบบหางยาวซึ่งเหมาะสมกับปริมาณน้ำฝนในพื้นที่รัฐโจหัว (Johor) ซึ่งตั้งอยู่ทางตอนใต้ของคาบสมุทรมาเลเซีย

โดยที่การแปลงข้อมูลเป็นการนำรูปแบบทางคณิตศาสตร์มาใช้ เพื่อปรับให้ชุดของข้อมูลมีลักษณะตามข้อตกลงเบื้องต้น ซึ่งจะทำให้ผลการวิเคราะห์มีความถูกต้องแม่นยำและน่าเชื่อถือ การแปลงข้อมูลมีผลต่ออำนาจการทดสอบคือช่วยให้อำนาจการทดสอบเพิ่มขึ้น สถิติทดสอบที่วิเคราะห์ด้วยข้อมูลที่ทำการแปลงแล้วจะทำให้ค่าอำนาจการทดสอบสูงกว่าสถิติทดสอบที่วิเคราะห์ด้วยข้อมูลที่ไม่ได้ทำการแปลงข้อมูลและสถิติไม่อิงพารามิเตอร์ (Rasmussen และ Dunlap, 1991) ดังนั้นในการศึกษาครั้งนี้ ผู้วิจัยสนใจที่จะศึกษารูปแบบการแปลงข้อมูลสำหรับแปลงข้อมูลที่มีลักษณะเบ้ขวาให้เป็นข้อมูลที่มีการแจกแจงปรกติหรือใกล้เคียงการแจกแจงปรกติมากที่สุด พร้อมทั้งประยุกต์ใช้วิธีการแปลงข้อมูลในการหาค่าดรรชนีความแห้งแล้ง

วิธีดำเนินการวิจัย

1. ดรรชนีความแห้งแล้งของฝนที่ต่างจากค่าปรกติ

ดรรชนีความแห้งแล้งของฝนที่ต่างจากค่าปรกติ (SPI) ได้พัฒนาขึ้นจากแนวคิดของ Mckee (1993) โดยพิจารณาใช้ฟังก์ชันการแจกแจงสะสมของปริมาณน้ำฝนสะสมแล้วทำการแปลงให้เป็นค่าปรกติมาตรฐานหรือดรรชนี SPI ถ้าค่าดรรชนี SPI มีค่าบวกแสดงว่าปริมาณน้ำฝนในสถานีที่สังเกตการณ์มากกว่าปริมาณน้ำฝนเฉลี่ย และถ้าค่าดรรชนี SPI มีค่าลบแสดงว่าปริมาณน้ำฝนในสถานีที่สังเกตการณ์น้อยกว่าปริมาณน้ำฝนเฉลี่ย สำหรับระดับความรุนแรงของสภาวะความแห้งแล้งตามดรรชนีความแห้งแล้งของฝนที่ต่างจากค่าปรกติ มีเกณฑ์ดังตารางที่ 1

ตารางที่ 1 เกณฑ์การแบ่งระดับความรุนแรงของสภาวะความแห้งแล้งตามดรรชนีความแห้งแล้งของฝนที่ต่างจากค่าปรกติ

ค่าดรรชนี SPI	**ระดับความรุนแรงของสภาวะความแห้งแล้ง**
มากกว่าหรือเท่ากับ 2	ฝนชุกมากที่สุด
1.50 ถึง 1.99	ฝนชุกมาก
1.00 ถึง 1.49	ฝนชุกปานกลาง
-0.99 ถึง 0.99	ฝนใกล้เคียงค่าปรกติ
-1.00 ถึง -1.49	ฝนแล้งปานกลาง
-1.50 ถึง -1.99	ฝนแล้งรุนแรง
น้อยกว่าหรือเท่ากับ -2	ฝนแล้งรุนแรงมากที่สุด

ที่มา: Mckee (1993)

2. การแจกแจงที่เกี่ยวข้อง

การแจกแจงแกมมา(Gamma Distribution) มี 2 พารามิเตอร์ โดยมีฟังก์ชันความหนาแน่นน่าจะเป็นดังนี้

$$f(x;\alpha,\beta) = \frac{x^{(\alpha-1)}e^{-x/\beta}}{\beta^{\alpha}\Gamma(\alpha)} \quad \text{โดยที่} \quad x > 0 \ \text{และ} \ \alpha,\beta > 0 \qquad (1)$$

โดยที่ $\Gamma(\alpha)=\int_0^{\infty} x^{\alpha-1}e^{-x}dx$ คือ ฟังก์ชันแกมมา α คือ พารามิเตอร์แสดงรูปร่าง (Shape Parameter) และ β คือ พารามิเตอร์แสดงสเกล (Scale Parameter)

การแจกแจงล็อกนอร์มอล (Lognormal Distribution) มี 2 พารามิเตอร์ โดยมีฟังก์ชันความหนาแน่นน่าจะเป็นดังนี้

$$f(x;\mu,\sigma)=\frac{1}{x\sigma\sqrt{2\pi}}e^{-\frac{(\ln x-\mu)^2}{2\sigma^2}} \quad \text{โดยที่} \quad x>0,\mu\in R,\sigma^2>0 \tag{2}$$

โดยที่ μ คือ พารามิเตอร์แสดงตำแหน่ง (Location Parameter) และ σ^2 คือ พารามิเตอร์แสดงสเกล (Scale Parameter)

การแจกแจงไวบูล (Weibull Distribution) มี 2 พารามิเตอร์ โดยมีฟังก์ชันความหนาแน่นน่าจะเป็นดังนี้

$$f(x;\alpha,\beta)=\frac{\alpha}{\beta}\left(\frac{x}{\beta}\right)^{\alpha-1}e^{-\left(\frac{x}{\beta}\right)^{\alpha}} \quad \text{โดยที่} \quad x>0 \text{ และ } \alpha,\beta>0 \tag{3}$$

โดยที่ α คือ พารามิเตอร์แสดงรูปร่าง (Shape Parameter) และ β คือ พารามิเตอร์แสดงสเกล (Scale Parameter)

การแจกแจงปรกติ (Normal Distribution) มี 2 พารามิเตอร์ โดยมีฟังก์ชันความหนาแน่นน่าจะเป็นดังนี้

$$f(x;\mu,\sigma^2)=\frac{1}{\sqrt{2\pi\sigma^2}}e^{-\frac{1}{2}\left(\frac{x-\mu}{\sigma}\right)^2} \quad , \quad x,\mu\in R,\sigma^2>0 \tag{4}$$

โดยที่ μ คือ พารามิเตอร์แสดงตำแหน่ง (Location Parameter) และ σ^2 คือ พารามิเตอร์แสดงสเกล (Scale Parameter)

3. การแปลงข้อมูล

จากการศึกษารูปแบบการแปลงข้อมูลจากงานวิจัยที่เกี่ยวข้อง วิธีการแปลงข้อมูลที่สามารถแปลงข้อมูลที่มีลักษณะเบ้ขวาให้มีการแจกแจงใกล้เคียงการแจกแจงปรกติ มีดังต่อไปนี้

1. **วิธีการแปลงข้อมูลแบบบ็อกซ์-ค็อกซ์ (Box and Cox Power Transformation: BC)** Box และ Cox ได้เสนอวิธีการแปลงข้อมูลสำหรับการแก้ไขปัญหาค่าสังเกตที่ไม่มีการแจกแจงปรกติให้มีลักษณะการแจกแจงปรกติและมีความแปรปรวนคงที่ ซึ่งวิธีการแปลงข้อมูลแบบบ็อกซ์-ค็อกซ์ (Box และ Cox, 1964) คือ

$$Y = \begin{cases} \dfrac{x^{\lambda}-1}{\lambda} & ,\lambda \neq 0 \\ \ln(x) & ,\lambda = 0 \end{cases} \qquad \text{สำหรับ } x > 0 \tag{5}$$

โดยที่ λ คือ พารามิเตอร์สำหรับการแปลง สามารถประมาณได้ด้วยวิธีภาวะน่าจะเป็นสูงสุด (Maximum Likelihood Estimation : MLE)

$$\frac{-n\left(\sum_{i=1}^{n} x_i^{2\lambda}\ln x_i - \left(\frac{1}{n}\right)\left(\sum_{i=1}^{n} x_i^{\lambda}\right)\left(\sum_{i=1}^{n} x_i^{\lambda}\ln x_i\right)\right)}{\left[\sum_{i=1}^{n} x_i^{2\lambda} - \frac{1}{n}\left(\sum_{i=1}^{n} x_i^{\lambda}\right)^2\right]} + \frac{n}{\lambda} + \sum_{i=1}^{n}\ln(x_i) = 0 \tag{6}$$

2. **วิธีการแปลงข้อมูลแบบเลขชี้กำลัง (Exponential Transformation: EP)** Manly ได้เสนอวิธีการแปลงข้อมูลแบบเลขชี้กำลัง ซึ่งการแปลงด้วยวิธีนี้จะมีประสิทธิภาพอย่างมากในข้อมูลที่มีลักษณะการแจกแจงเบ้ฐานนิยมเดี่ยว (Skew Unimodal Distribution) ซึ่งวิธีการแปลงข้อมูลแบบเลขชี้กำลัง (Manly, 1971) คือ

$$Y = \begin{cases} \dfrac{\exp(\lambda x)-1}{\lambda} & ,\lambda \neq 0 \\ x & ,\lambda = 0 \end{cases} \tag{7}$$

โดยที่ λ คือพารามิเตอร์สำหรับการแปลง สามารถประมาณได้ด้วยวิธี MLE

$$\frac{-n\left(\sum_{i=1}^{n} e^{2x_i\lambda}x_i - \left(\frac{1}{n}\right)\left(\sum_{i=1}^{n} e^{x_i\lambda}\right)\left(\sum_{i=1}^{n} e^{x_i\lambda}x_i\right)\right)}{\left[\sum_{i=1}^{n} e^{2x_i\lambda} - \frac{1}{n}\left(\sum_{i=1}^{n} e^{x_i\lambda}\right)^2\right]} + \frac{n}{\lambda} + \sum_{i=1}^{n} x_i = 0 \tag{8}$$

3. **วิธีการแปลงข้อมูลแบบ Yeo และ Johnson (Yeo and Johnson Transformation: YJ)** Yeo และ Johnson ได้เสนอรูปแบบการแปลงข้อมูลโดยพัฒนามาจากวิธีการแปลงของ Box และ Cox ซึ่งวิธีการแปลงข้อมูลแบบ Yeo และ Johnson (Yeo และ Johnson, 2000) คือ

$$Y=\begin{cases}\dfrac{(x+1)^{\lambda}-1}{\lambda} & ,x\geq 0,\lambda\neq 0\\ \ln(x+1) & ,x\geq 0,\lambda=0\end{cases} \tag{9}$$

โดยที่ λ คือ พารามิเตอร์สำหรับการแปลง สามารถประมาณได้ด้วยวิธี MLE

$$\frac{-n\left(\sum_{i=1}^{n}(x_i+1)^{2\lambda}\ln(x_i+1)-\left(\frac{1}{n}\right)\left(\sum_{i=1}^{n}(x_i+1)^{\lambda}\right)\left(\sum_{i=1}^{n}(x_i+1)^{\lambda}\ln(x_i+1)\right)\right)}{\left[\sum_{i=1}^{n}(x_i+1)^{2\lambda}-\frac{1}{n}\left(\sum_{i=1}^{n}(x_i+1)^{\lambda}\right)^{2}\right]}+\frac{n}{\lambda}+\sum_{i=1}^{n}\ln(x_i+1)=0 \tag{10}$$

ถ้าข้อมูลมีการแจกแจงไวบูลสามารถใช้**วิธีการแปลงข้อมูลแบบ Haynes (Haynes Transformation)** ในการแปลงข้อมูลให้มีการแจกแจงใกล้เคียงการแจกแจงปรกติ วิธีการแปลงข้อมูลแบบ Haynes ได้พัฒนามาจากวิธีการแปลงของ Box และ Cox โดยกำหนดพารามิเตอร์สำหรับการแปลง (λ) มีค่าเท่ากับ $10^{-0.6230+(1.004\times\log_{10}(\text{Shape}))}$ ซึ่งวิธีการแปลงข้อมูลแบบ Haynes มีรูปแบบการแปลงข้อมูลดังนี้ (Haynes, 2011)

$$Y=\begin{cases}\dfrac{x^{10^{-0.6230+(1.004\times\log_{10}(\text{Shape}))}}-1}{10^{-0.6230+(1.004\times\log_{10}(\text{Shape}))}} & ,\lambda\neq 0\\ \ln(x) & ,\lambda=0\end{cases} \tag{11}$$

โดยที่ Shape คือ พารามิเตอร์แสดงรูปร่าง (Shape Parameter)

จากการศึกษางานวิจัยที่เกี่ยวข้องถ้าข้อมูลมีการแจกแจงแกมมา Krishnamoorthy และคณะ (2008) ได้เสนอว่าวิธีการแปลงข้อมูลที่เหมาะสมสำหรับการแปลงข้อมูลที่มีการแจกแจงแกมมาให้มีการแจกแจงใกล้เคียงการแจกแจงปรกติ มีดังต่อไปนี้

- **วิธีการแปลงข้อมูลแบบถอดรากที่สาม** (Cube-Root Transformation) ซึ่งมีรูปแบบการแปลงข้อมูลดังนี้

$$Y=x^{\frac{1}{3}} \tag{12}$$

- **วิธีการแปลงข้อมูลแบบถอดรากที่สี่** (Fourth-Root Transformation) ซึ่งมีรูปแบบการแปลงข้อมูลดังนี้

$$Y = x^{\frac{1}{4}} \tag{13}$$

ถ้าข้อมูลมีการแจกแจงล็อกนอร์มอล Ngamjarus (2001) ได้กล่าวว่า สำหรับการแจกแจงล็อกนอร์มอลวิธีที่เหมาะสมในการแปลงข้อมูลให้มีการแจกแจงใกล้เคียงการแจกแจงปรกติ คือ **วิธีการแปลงข้อมูลแบบลอการิทึม** (Logarithm Transformation) ซึ่งมีรูปแบบการแปลงข้อมูลดังนี้

$$Y = \log(x) \tag{14}$$

4. การทดสอบภาวะสารูปดี (Goodness-of-fit Tests)

4.1 สถิติทดสอบแอนเดอร์สัน-ดาร์ลิ่ง (Anderson-Darling Test: AD) Anderson และ Darling ได้เสนอสถิติทดสอบภาวะสารูปดีซึ่งเป็นการเปรียบเทียบระหว่างฟังก์ชันการแจกแจงสะสมของค่าสังเกตกับฟังก์ชันการแจกแจงสะสมที่คาดหมาย ซึ่งมีสูตรในการคำนวณดังนี้

$$A^2 = -n - \frac{1}{n}\sum_{i=1}^{n}(2i-1)\left[\ln F(x_i) + \ln(1 - F(x_{n+1-i})\right] \tag{15}$$

โดยที่ n คือ ขนาดตัวอย่าง และ $F(x_i)$ คือ ฟังก์ชันการแจกแจงสะสมของการแจกแจงที่คาดหมาย ของ x_i ซึ่งมีการเรียงลำดับจากน้อยไปหามาก

4.2 เกณฑ์สารสนเทศของอะกะอิเกะหรือเกณฑ์เอไอซี (Akaike Information Criterion: AIC) Akaike ได้เสนอเกณฑ์การคัดเลือกตัวแบบซึ่งสร้างมาจากความแปรปรวนของข้อสนเทศคูลส์แบล็ค-ไลท์เบอร์ (Kullback Leibler) ระหว่างตัวแบบที่แท้จริงกับตัวแบบที่เหมาะสม การคัดเลือกตัวแบบโดยใช้เกณฑ์เอไอซีจะเลือกตัวแบบที่ให้ค่าเอไอซีต่ำสุดเป็นตัวแบบที่ดีที่สุด ซึ่งมีสูตรในการคำนวณดังนี้

$$AIC = 2k - 2\ln L \tag{16}$$

โดยที่ k คือ จำนวนพารามิเตอร์ในตัวแบบ และ L คือ ค่าสูงสุดของฟังก์ชันภาวะน่าจะเป็นสูงสุดสำหรับการประมาณ

5. ข้อมูลที่ใช้ในการวิจัย

5.1 ข้อมูลจากการจำลอง

5.1.1 จำลองข้อมูลให้มีการแจกแจงและพารามิเตอร์ระดับต่างๆ ดังนี้

1) การแจกแจงแกมมา α = 0.5, 4.0, 6.0 และ β = 1.0, 3.0

2) การแจกแจงล็อกนอร์มอล μ = 0.5, 4.0, 6.0 และ σ = 0.5, 1.5

3) การแจกแจงไวบูล α = 0.5, 1.0, 2.0 และ β = 1.0, 3.0

โดยสุ่มตัวอย่างขนาด 15, 30, 50 และ 100

5.1.2 ตรวจสอบว่าข้อมูลที่จำลองในข้อ 5.1.1 มีการแจกแจงแกมมา การแจกแจงล็อกนอร์มอลและการแจกแจงไวบูลหรือไม่ โดยสถิติทดสอบ AD ที่ระดับนัยสำคัญ 0.05

5.1.3 ตรวจสอบว่าข้อมูลที่จำลองในข้อ 5.1.1 มีการแจกแจงปรกติหรือไม่ โดยสถิติทดสอบ AD ที่ระดับนัยสำคัญ 0.05

5.1.4 ทำการแปลงข้อมูลในแต่ละวิธีการแปลงข้อมูล สำหรับวิธีการแปลงแบบ BC EP และ YJ จะประมาณค่าพารามิเตอร์การแปลงด้วยการประยุกต์ใช้วิธีซีแคนต์ (Secant Method) ในแต่ละกรณี เนื่องจากการประมาณค่าพารามิเตอร์ด้วยวิธีภาวะน่าจะเป็นสูงสุดมีความซับซ้อนเกินไปและไม่สามารถประมาณค่าพารามิเตอร์ได้เมื่อค่าที่ประมาณนั้นเข้าใกล้ศูนย์เนื่องจากพารามิเตอร์สำหรับการแปลงที่เราต้องการประมาณค่ามีค่าเข้าใกล้ศูนย์ (Lakhana, 2014)

5.1.5 ตรวจสอบว่าข้อมูลที่ผ่านการแปลงในแต่ละวิธีมีการแจกแจงปรกติหรือไม่ โดยใช้สถิติทดสอบ AD ที่ระดับนัยสำคัญ 0.05

5.1.6 ทำซ้ำข้อ 5.1.1 ถึงข้อ 5.1.5 จำนวน 1000 ครั้ง

5.1.7 พิจารณาร้อยละของการยอมรับสมมติฐานหลักซึ่งคำนวณดังนี้

$$\frac{\text{จำนวนชุดข้อมูลที่มีการแจกแจงปรกติหลังการแปลงข้อมูล}}{1000} \times 100$$

5.1.8 สรุปผล

5.2.สำหรับข้อมูลจริง

5.2.1 นำข้อมูลปริมาณน้ำฝนรายเดือนจากสถานีสำรวจปริมาณน้ำฝนของศูนย์อุทกวิทยาและบริหารน้ำภาคเหนือตอนบน กรมชลประทาน มาทำการสะสมเป็นปริมาณน้ำฝนราย 4 เดือน (ปริมาณน้ำฝนตามฤดูกาลของประเทศไทย) โดยที่ปริมาณน้ำฝนตามฤดูกาลแบ่งตามฤดูกาลของประเทศไทยซึ่งแบ่งได้ 3 ฤดูกาล ได้แก่ ฤดูร้อนประมาณเดือนกุมภาพันธ์ถึงเดือนพฤษภาคม ฤดูฝนประมาณเดือนมิถุนายนถึงเดือนกันยายน และฤดูหนาวประมาณเดือนตุลาคมถึงเดือนมกราคมของปีถัดไป โดยในงานวิจัยนี้จะใช้ข้อมูลปริมาณน้ำฝนจากสถานีตรวจวัดปริมาณน้ำฝนของอำเภอลี้และอำเภอบ้านโฮ่ง จังหวัดลำพูน

5.2.2 ตรวจสอบว่าข้อมูลปริมาณน้ำฝนตามฤดูกาล ในข้อ 5.2.1 มีการแจกแจงแกมมา การแจกแจงล็อกนอร์มอล หรือการแจกแจงไวบูล โดยสถิติทดสอบ AD ที่ระดับนัยสำคัญ 0.05 และเกณฑ์ AIC

5.2.3 ตรวจสอบว่าข้อมูลปริมาณน้ำฝนตามฤดูกาล ในข้อ 5.2.1 มีการแจกแจงปรกติหรือไม่ โดยใช้สถิติทดสอบ AD ที่ระดับนัยสำคัญ 0.05

5.2.4 นำข้อมูลปริมาณน้ำฝนตามฤดูกาลมาทำการแปลงข้อมูลโดยวิธีที่ดีที่สุดในแต่ละการแจกแจง

5.2.5 ตรวจสอบว่าข้อมูลปริมาณน้ำฝนตามฤดูกาลที่แปลงแล้วในข้อ 5.2.4 มีการแจกแจงปรกติหรือไม่ โดยใช้สถิติทดสอบ AD ที่ระดับนัยสำคัญ 0.05

5.2.6 ถ้าข้อมูลปริมาณน้ำฝนตามฤดูกาลที่แปลงแล้วมีการแจกแจงปรกติจะนำข้อมูลปริมาณน้ำฝนตามฤดูกาลที่แปลงแล้วไปประเมินภัยแล้งโดยใช้วิธีดรรชนี SPI

ผลการวิจัยและวิจารณ์ผล

ผลการจำลองข้อมูลสามารถเปรียบเทียบร้อยละของการยอมรับสมมติฐานข้อมูลมีการแจกแจงปรกติได้ดังนี้

1. การแจกแจงแกมมาวิธีการแปลงข้อมูลแบบถอดรากที่สามและวิธีการแปลงข้อมูลแบบถอดรากที่สี่ให้ร้อยละของการยอมรับสมมติฐานสูงสุด แต่วิธีการแปลงข้อมูลแบบถอดรากที่สาม วิธีการแปลงข้อมูลแบบบ็อกซ์-ค็อกซ์ วิธีการแปลงข้อมูลแบบเลขชี้กำลังและวิธีการแปลงข้อมูลแบบ Yeo และ Johnson จะไม่เหมาะสมเมื่อข้อมูลที่จำลองมีความเบ้ที่มากขึ้นและขนาดตัวอย่างที่เพิ่มขึ้น ดังตารางที่ 2

2. การแจกแจงไวบูลวิธีการแปลงข้อมูลแบบ Haynes และวิธีการแปลงข้อมูลแบบบ็อกซ์-ค็อกซ์ให้ร้อยละของการยอมรับสมมติฐานสูงสุด ในส่วนของวิธีการแปลงข้อมูลแบบ Yeo และ Johnson และวิธีการแปลงข้อมูลแบบเลขชี้กำลังจะไม่เหมาะสมเมื่อความเบ้มากขึ้นและขนาดตัวอย่างเพิ่มขึ้น ดังตารางที่ 3

3. การแจกแจงล็อกนอร์มอลวิธีการแปลงข้อมูลแบบลอการิทึมให้ร้อยละของการยอมรับสมมติฐานสูงสุด รองลงมาได้แก่ วิธีการแปลงข้อมูลแบบบ็อกซ์-ค็อกซ์และวิธีการแปลงข้อมูลแบบ Yeo และ Johnson ดังตารางที่ 4

ตารางที่ 2 ร้อยละของการยอมรับสมมติฐานข้อมูลมีการแจกแจงปรกติของการแปลงการแจกแจงแกมมา

n	α	β	ความเบ้	$x^{1/3}$	$x^{1/4}$	BC	EP	YJ	n	α	β	ความเบ้	$x^{1/3}$	$x^{1/4}$	BC	EP	YJ
15	0.5	1	2.83	97.5	**99.6**	97.9	46.2	67.0	50	0.5	1	2.83	84.0	**97.8**	88.2	0.8	4.6
	4	1	1.00	**100.0**	99.8	99.6	99.9	99.8		4	1	1.00	**99.8**	99.3	99.2	97.1	99.5
	6	1	0.82	**100.0**	**100.0**	99.8	100.0	99.8		6	1	0.82	**100.0**	99.7	100.	98.8	100.
	0.5	3	2.83	97.5	**99.6**	97.9	46.2	56.2		0.5	3	2.83	84.0	**97.8**	88.2	0.8	1.1
	4	3	1.00	**100.0**	99.8	99.6	99.8	99.8		4	3	1.00	**99.8**	99.3	99.2	97.1	99.0
	6	3	0.82	**100.0**	**100.0**	99.8	100.0	99.9		6	3	0.82	**100.0**	99.7	100.	97.6	99.9
30	0.5	1	2.83	97.5	**98.4**	93.2	11.6	30.3	100	0.5	1	2.83	61.7	**94.6**	57.5	0.0	0.0
	4	1	1.00	**100.0**	99.6	99.6	99.6	99.8		4	1	1.00	**99.8**	98.8	99.7	96.2	99.3
	6	1	0.82	99.2	**99.4**	99.1	99.1	99.2		6	1	0.82	**100.0**	99.7	99.3	99.2	99.7
	0.5	3	2.83	92.3	**98.4**	93.2	11.6	18.9		0.5	3	2.83	61.7	**94.5**	57.5	0.0	0.0
	4	3	1.00	**100.0**	99.6	99.6	99.6	99.8		4	3	1.00	**99.8**	98.8	99.7	96.2	99.0
	6	3	0.82	99.2	**99.4**	99.1	99.1	99.1		6	3	0.82	**100.0**	99.7	99.3	99.2	99.7

ตารางที่ 3 ร้อยละของการยอมรับสมมติฐานข้อมูลมีการแจกแจงปรกติของการแปลงการแจกแจงไวบูล

n	α	β	ความเบ้	Haynes	BC	EP	YJ
15	0.5	1	5.20	**99.5**	98.0	28.6	60.0
	1	1	0.65	**99.9**	96.6	81.2	93.2
	2	1	0.03	**99.3**	69.4	15.3	43.9
	0.5	3	5.20	**99.5**	98.0	28.6	76.3
	1	3	0.65	**99.9**	96.6	81.2	96.7
	2	3	0.03	**99.3**	69.4	15.3	54.4
30	0.5	1	5.20	98.9	**99.2**	1.8	16.7
	1	1	0.65	**99.6**	**99.6**	74.4	88.2
	2	1	0.03	**100.0**	98.5	96.5	98.7
	0.5	3	5.20	98.8	**99.2**	1.8	31.9
	1	3	0.65	**99.6**	99.6	74.4	93.8
	2	3	0.03	**100.0**	98.5	96.5	99.4
50	0.5	1	5.20	98.2	**98.8**	0.0	1.7
	1	1	0.65	98.3	**98.9**	46.9	76.5
	2	1	0.03	**99.6**	96.4	87.2	94.1
	0.5	3	5.20	98.2	**98.8**	0.0	9.5
	1	3	0.65	98.3	**98.9**	46.9	89.1
	2	3	0.03	**99.6**	96.4	87.2	95.8
100	0.5	1	5.20	**95.7**	93.8	0.0	0.0
	1	1	0.65	**95.7**	93.8	6.0	33.3
	2	1	0.03	**99.0**	90.9	73.2	83.4
	0.5	3	5.20	**95.6**	93.8	0.0	0.1
	1	3	0.65	**95.7**	93.8	6.0	64.2
	2	3	0.03	**99.0**	90.9	73.2	87.7

ตารางที่ 4 ร้อยละของการยอมรับสมมติฐานข้อมูลมีการแจกแจงปรกติของการแปลงการแจกแจงล็อกนอร์มอล

n	μ	σ	ความเบ้	Log	BC	EP	YJ
15	0.5	0.5	1.75	**100.0**	99.4	95.4	100.0
	4	0.5	1.75	**100.0**	99.4	99.2	99.5
	6	0.5	1.75	**100.0**	99.4	99.9	99.5
	0.5	1.5	33.47	**100.0**	98.9	36.3	93.2
	4	1.5	33.47	**100.0**	98.9	70.1	99.8
	6	1.5	33.47	**100.0**	98.9	59.7	99.2
30	0.5	0.5	1.75	**99.9**	98.3	96.8	99.6
	4	0.5	1.75	**99.9**	98.3	96.8	98.3
	6	0.5	1.75	**99.9**	98.3	96.8	98.0
	0.5	1.5	33.47	**100.0**	94.5	14.0	86.1
	4	1.5	33.47	**100.0**	94.5	0.3	96.8
	6	1.5	33.47	**100.0**	94.5	14.0	94.7
50	0.5	0.5	1.75	**100.0**	93.2	85.5	97.6
	4	0.5	1.75	**100.0**	93.2	85.5	93.7
	6	0.5	1.75	**100.0**	93.2	97.2	93.2
	0.5	1.5	33.47	**100.0**	87.7	0.6	69.3
	4	1.5	33.47	**100.0**	87.7	0.6	91.7
	6	1.5	33.47	**100.0**	87.7	0.9	88.6
100	0.5	0.5	1.75	**100.0**	97.0	84.5	100.0
	4	0.5	1.75	**100.0**	97.0	84.5	99.50
	6	0.5	1.75	**100.0**	97.0	92.4	99.5
	0.5	1.5	33.47	**100.0**	96.8	0.0	93.2
	4	1.5	33.47	**100.0**	96.8	0.0	99.8
	6	1.5	33.47	**100.0**	96.8	0.0	99.2

สำหรับข้อมูลจริงมีผลการวิเคราะห์ดังต่อไปนี้ สถานีตรวจวัดปริมาณน้ำฝนของอำเภอบ้านโฮ่ง จังหวัดลำพูน พบว่าปริมาณน้ำฝนตามฤดูกาลมีการแจกแจงไวบูล ส่วนสถานีตรวจวัดปริมาณน้ำฝนของอำเภอลี้ จังหวัดลำพูน พบว่าข้อมูลปริมาณน้ำฝนตามฤดูกาลมีการแจกแจงแกมมา ดังตารางที่ 5

ตารางที่ 5 การทดสอบการแจกแจงของข้อมูลปริมาณน้ำฝนตามฤดูกาลของอำเภอบ้านโฮ่งและอำเภอลี้ จังหวัดลำพูน

สถานี	ปริมาณน้ำฝน (มม.)		มัธยฐาน	ส่วนเบี่ยงเบนคลอไทล์	ความเบ้	การแจกแจงแกมมา		การแจกแจงไวบูล		การแจกแจงล็อกนอร์มอล	
	ต่ำสุด	สูงสุด				AD	AIC	AD	AIC	AD	AIC
อ.บ้านโฮ่ง	4.0	863.5	293.9	134.30	0.65876	0.4733	2042.8	0.4811	2038.6	1.6045*	2073.0
อ.ลี้	31.3	907.1	276.0	133.35	0.63654	0.7223	2139.3	0.9159*	2139.0	1.1371*	2152.0

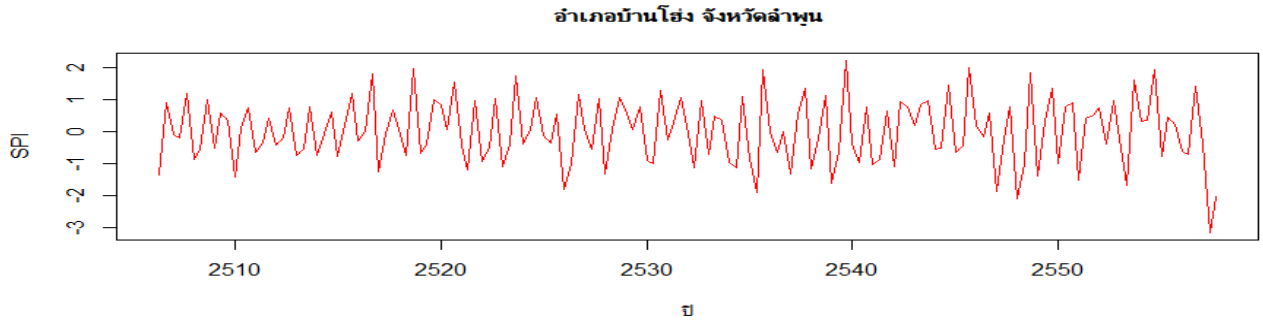
*มีนัยสำคัญทางสถิติที่ 0.05, มม. คือ มิลลิเมตร

จากการจำลองข้อมูลถ้าข้อมูลมีการแจกแจงไวบูลวิธีการแปลงข้อมูลที่เหมาะสมสำหรับการแปลงข้อมูลให้มีการแจกแจงปรกติคือ วิธีการแปลงข้อมูลแบบ Haynes ส่วนข้อมูลมีการแจกแจงแกมมาคือ วิธีการแปลงข้อมูลแบบถอดรากที่สี่ ดังนั้นสถานีตรวจวัดปริมาณน้ำฝนของอำเภอบ้านโฮ่งจะทำการแปลงด้วยวิธีการแปลงข้อมูลแบบ Haynes และสถานีตรวจวัดปริมาณน้ำฝนของอำเภอลี้จะทำการแปลงด้วยวิธีการแปลงข้อมูลแบบถอดรากที่สี่ จากตารางที่ 5 พบว่าสถานีตรวจวัดปริมาณน้ำฝนของอำเภอบ้านโฮ่งก่อนการแปลงข้อมูลปริมาณน้ำฝนตามฤดูกาลไม่มีการแจกแจงปรกติ หลังจากแปลงด้วยวิธีการแปลงข้อมูลแบบ Haynes ทำให้ข้อมูลปริมาณน้ำฝนตามฤดูกาลมีการแจกแจงปรกติ และสถานีตรวจวัดปริมาณน้ำฝนของอำเภอลี้ก่อนการแปลงข้อมูลปริมาณน้ำฝนตามฤดูกาลไม่มีการแจกแจงปรกติ หลังจากแปลงด้วยวิธีการแปลงข้อมูลแบบถอดรากที่สี่ทำให้ข้อมูลปริมาณน้ำฝนตามฤดูกาลมีการแจกแจงปรกติ ดังตารางที่ 6

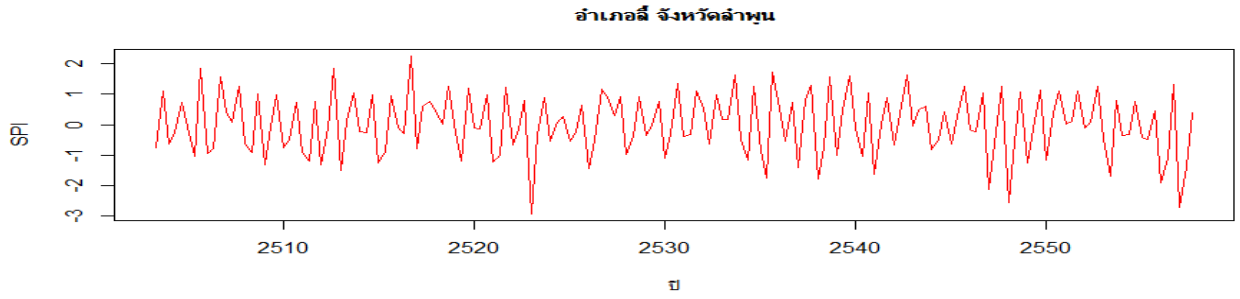
ตารางที่ 6 การทดสอบการแปลงข้อมูลปริมาณน้ำฝนตามฤดูกาลของประเทศไทย ของอำเภอบ้านโฮ่งและลี้ จังหวัดลำพูน

สถานี	วิธีการแปลง	ก่อนแปลงข้อมูล		หลังการแปลงข้อมูล	
		AD	p-value	AD	p-value
อ.บ้านโฮ่ง	Haynes	2.1063	<0.0001**	0.4734	0.23920
อ.ลี้	$x^{1/4}$	2.6023	<0.0001**	0.7374	0.05365

**มีนัยสำคัญทางสถิติที่ 0.01



ภาพที่ 1 ค่าดรรชนี SPI ของสถานีอำเภอบ้านโฮ่ง จังหวัดลำพูน ตั้งแต่ปี พ.ศ. 2506 – 2557



ภาพที่ 2 ค่าดรรชนี SPI ของสถานีอำเภอลี้ จังหวัดลำพูน ตั้งแต่ปี พ.ศ. 2503 – 2557

หลังจากทำการแปลงข้อมูลปริมาณน้ำฝนตามฤดูกาลด้วยวิธีการแปลงข้อมูลที่เหมาะสมในแต่ละการแจกแจง พบว่าข้อมูลปริมาณน้ำฝนตามฤดูกาลมีการแจกแจงปรกติ ดังนั้นจึงประเมินความแห้งแล้งโดยใช้วิธีดรรชนี SPI พบว่าอำเภอบ้านโฮ่ง จังหวัดลำพูน ในปี พ.ศ. 2548 และ 2557 มีฝนแล้งรุนแรงมากที่สุด ดังภาพที่ 1 ส่วนอำเภอลี้ จังหวัดลำพูน ในปี พ.ศ. 2523 2547 2548 และ 2557 มีฝนแล้งรุนแรงมากที่สุด ดังภาพที่ 2

สรุปผลการวิจัย

จากผลการวิจัยพบว่า ในส่วนของการจำลองข้อมูลเมื่อข้อมูลมีการแจกแจงแกมมาวิธีการแปลงข้อมูลที่เหมาะสมมากที่สุดในการแปลงข้อมูลให้มีการแจกแจงปรกติคือ วิธีการแปลงข้อมูลแบบถอดรากที่สี่ รองลงมาคือ วิธีการแปลงข้อมูลแบบถอดรากที่สาม วิธีการแปลงข้อมูลที่เหมาะสมน้อยที่สุดคือ วิธีการแปลงข้อมูลแบบเลขชี้กำลัง สำหรับข้อมูลที่มีการแจกแจงล็อกนอร์มอลวิธีการแปลงข้อมูลที่เหมาะสมมากที่สุดคือ วิธีการแปลงข้อมูลแบบลอการิทึม วิธีการแปลงข้อมูลที่เหมาะสมน้อยที่สุดคือ วิธีการแปลงข้อมูลแบบเลขชี้กำลัง ในส่วนของการแจกแจงไวบูลวิธีการแปลงข้อมูลที่เหมาะสมมากที่สุดคือ วิธีการแปลงข้อมูลแบบ Haynes ส่วนวิธีการแปลงข้อมูลที่เหมาะสมน้อยที่สุดคือ วิธีการแปลงข้อมูลแบบเลขชี้กำลัง จะเห็นได้ว่าถ้าทราบการแจกแจงว่ามีการแจกแจงแกมมาสามารถใช้วิธีการแปลงข้อมูลแบบถอดรากที่สี่ ส่วนการแจกแจงล็อกนอร์มอลสามารถใช้วิธีการแปลงข้อมูลแบบลอการิทึมและการแจกแจงไวบูลสามารถใช้วิธีการแปลงข้อมูลแบบ Haynes ถ้าไม่ทราบการแจกแจงข้อมูลสามารถใช้วิธีการแปลงข้อมูลแบบบ็อกซ์-ค็อกซ์ซึ่งสามารถแปลงข้อมูลได้ดีและสอดคล้องกับงานวิจัยของ Ngamjarus (2001) สำหรับข้อมูลจริงผู้วิจัยได้เลือกใช้สถานีตรวจวัดปริมาณน้ำฝนของอำเภอบ้านโฮ่งและอำเภอลี้ จังหวัดลำพูน พบว่าน้ำฝนตามฤดูกาลของประเทศไทยของอำเภอบ้านโฮ่งมีการแจกแจงไวบูล ส่วนอำเภอลี้มีการแจกแจงแกมมา การประเมินความแห้งแล้ง อำเภอบ้านโฮ่งมีฝนแล้งรุนแรงมากที่สุด พ.ศ. 2548 และ 2557 และอำเภอลี้มีฝนแล้งรุนแรงมากที่สุดในปี พ.ศ. 2523 2547 2548 และ 2557

กิตติกรรมประกาศ

ผู้เขียนขอขอบพระคุณ ศูนย์วิจัยวิทยาศาสตร์พหุวิทยาการ คณะวิทยาศาสตร์ มหาวิทยาลัยเชียงใหม่ ที่จัดสรรทุนสนับสนุนในการทำวิจัยครั้งนี้ ขอขอบพระคุณ ศูนย์อุทกวิทยาและบริหารน้ำภาคเหนือตอนบน กรมชลประทาน ที่ให้ความอนุเคราะห์ข้อมูลในการทำวิจัยในครั้งนี้ และขอขอบพระคุณผู้ทรงคุณวุฒิที่ช่วยปรับปรุงและเสนอแนะจนบทความนี้มีความสมบูรณ์มากขึ้น

เอกสารอ้างอิง

Agro - meteorological Academic Group Meteorological Development Bureau. (2011). *Study on Drought Index in Thailand*. Retrieved September 10, 2015, from http://tmd.go.th. (in Thai)

Box, G. E. P. and Cox, D. R. (1964). An analysis of transformations. *Journal of the Royal Statistical Society*, *26(*2), 211-252.

Bickel, P.J. and Doksum, K.A. (1981). An analysis of transformations revisited. *Journal of the American Statistical Association*. *76*, 296- 311

Krishnamoorthy, K., Mathew, T. and Mukherjee, S. (2008). Normal-Based Methods for a Gamma Distribution: Prediction and Tolerance Intervals and Stress-Strength Reliability. *American Statistical Association and the American Society for Quality*, *50*(1), 69-78.

Haynes, R. (2011). Transformation of Weibull Distribution Data A surprising result. *Smarter Solutions*, 1-10

Manly, B. F. J. (1976). Exponential Data Transformations. *Statistician*, *25*(1), 37-42.

McKee, T.B., Doesken, N.J. and Kleist, J. 1993. The relationship of drought frequency and duration on time scale. *Eighth Conf on Applied Climatology*, 179–184.

Ngamjarus, C. (2001). The Comparison of Transformation Methods for Exponential Data to Normal Data. Bangkok : King Mongkut's University of technology North Bangkok.(in Thai)

Thai Meteorological Department. (2015). *Drought.* Retrieved September 10, 2015, from http://www.tmd.go.th /info/ info.php?FileID=71 (in Thai)

Watthanacheewakul, L. (2014). A New Family of Transformations for Lifetime Data. In *Proceedings of the World Congress on Engineering.* United Kingdom: London.

Yeo, I. and Johnson, N. R. (2000). A new family of power transformations to improve normality or symmetry. *Biometrika*, *87*(2), 954-959.

Yusof, F.and Hui-Mean, F. (2012). Use of statistical distribution for drought analysis. *Applied Math Sci*, *6*(21), 1031–1051.

Zhang, Q., Xu, C.Y. and Zhang, Z. (2009). Observed changes of drought wetness episodes in the Pearl River Basin, China, using the standardized precipitation index and aridity index. *Theor Appl Climatol*, *98*,89–99.