บทที่ 2
เอกสารและงานวิจัยที่เกี่ยวข้อง

งานวิจัยที่เกี่ยวข้องกับ MCDM

Filar et al. (1999) ได้ประยุกต์ใช้ MCDM ในการประเมินค่าสิ่งแวดล้อมเพื่อลดการปลดปล่อยสารพิษ โดยใช้ข้อมูลจากฐานข้อมูลการปลดปล่อยสารพิษ (Toxic Release Inventory, TRI) ประเมินค่าสิ่งแวดล้อมของรัฐต่าง ๆ 54 รัฐ ในสหรัฐอเมริกา ตั้งแต่ปี 1987 – 1994 โดยใช้สารมลพิษสำคัญ 17 ตัวเป็นตัวชี้วัด และกำหนดน้ำหนักที่เหมาะสมแก่ตัวชี้วัด โดยใช้เอนโทรปี (Entropy) ซึ่งเป็นตัวบ่งบอกปริมาณความผันแปรของข้อมูลในแต่ละตัวชี้วัด ผลการวิจัยพบว่า รัฐที่มีคุณภาพสิ่งแวดล้อมดีที่สุดคือ รัฐที่ 5, 10 และ 32 ตามลำดับ และรัฐที่แย่ที่สุด คือรัฐที่ 46 และเมื่อตัดรัฐ 2 รัฐ คือรัฐที่ 5 และ 10 ที่มีปริมาณสารพิษเป็นศูนย์หลายชนิดออก และตัดสารเคมี 2 ชนิด ที่ค่าเป็นศูนย์ในหลาย ๆ รัฐออก วิเคราะห์ข้อมูลที่เหลือ พบว่า รัฐที่ดีสุด และแย่ที่สุดคือรัฐที่ 32 และ 46 ตามลำดับ สำหรับอันดับอื่น ๆ มีการเปลี่ยนแปลงเล็กน้อย นั่นแสดงว่า ตัวชี้วัดที่มีความผันแปรน้อยจะมีผลต่อการจัดอันดับน้อย เป็นไปตามหลักการถ่วงน้ำหนักโดยใช้เอนโทรปี

Maitra et al. (2002) ได้ประยุกต์ใช้ MCDM ในการประเมินค่าสิ่งแวดล้อมโดยใช้ข้อมูลด้านอากาศ น้ำ และคุณภาพดินใน 50 รัฐของสหรัฐอเมริกา และใน 106 ประเทศทั่วโลก มีการถ่วงน้ำหนักตัวชี้วัด 2 แบบ คือแบบใช้เอนโทรปี และแบบใช้สัมประสิทธิ์ การแปรผันตัวอย่าง (Sample coefficient variation) ซึ่งต่างก็เป็นค่าที่ใช้วัดความแปรผันของข้อมูลในแต่ละตัวชี้วัดแต่ใช้วิธีการคำนวณต่างกัน ผลการวิจัยพบว่า เมื่อใช้น้ำหนักทั้งสองแบบแล้ว ผลการเปรียบเทียบโดยเฉลี่ยปรากฏว่า ในสหรัฐอเมริกา รัฐที่มีสิ่งแวดล้อมดีที่สุดคือ รัฐหมายเลข 499 และแย่ที่สุดคือรัฐหมายเลข 060 สำหรับใน 106 ประเทศทั่วโลก ประเทศที่สิ่งแวดล้อมดีที่สุดคือโมลโดวา (Moldova) และแย่ที่สุดคือ คอสตาริกา (Costa Rica)

Lertprapai et al. (2004) ได้ประยุกต์ใช้ MCDM กับข้อมูลมลภาวะทางอากาศในกรุงเทพมหานคร ตั้งแต่ปี 1998 – 2001 โดยใช้ปริมาณสารมลพิษในอากาศ 4 ชนิด คือ คาร์บอนไดออกไซด์ ไนโตรเจนไดออกไซด์ ซัลเฟอร์ไดออกไซด์ และโอโซน เป็นตัวชี้วัดเปรียบเทียบคุณภาพอากาศจากสถานีตรวจวัดคุณภาพอากาศจำนวน 10 สถานี ในกรุงเทพมหานคร ผลการวิจัยพบว่า ในปี 1998, 2000 และ 2001 สถานีที่มีคุณภาพอากาศดีที่สุดคือ กรมอุตุนิยมวิทยา สำหรับในปี 1999 คือที่ทำการไปรษณีย์ราษฎร์บูรณะ ในขณะที่สถานีที่แย่ที่สุดในทุกปีคือ

การเคหะชุมชนดินแดง นอกจากนี้ยังได้หาความสัมพันธ์ของการจัดอันดับโดยวิธีต่าง ๆ โดยใช้สัมประสิทธิ์สหสัมพันธ์แบบสเปียร์แมนพบว่าการจัดอันดับแต่ละวิธีมีความสัมพันธ์กัน

Lertprapai et al. (2004) ได้พิจารณาปัญหาการประมาณค่าสัดส่วนของการแจกแจงแบบทวินาม ($B(n,\theta)$) คือเปรียบเทียบตัวประมาณค่ามาตรฐาน 2 ตัว ได้แก่ $T_1 = x/n$ และ $T_2 = (x+\sqrt{n}/2)/(n+\sqrt{n})$ เมื่อทราบค่าของ n และ $0<\theta<1$ โดยใช้วิธี MCDM และถ่วงน้ำหนัก 3 แบบ คือ แบบน้ำหนักเท่ากันทุกตัวชี้วัด แบบใช้เอนโทรปี และแบบใช้สัมประสิทธิ์การแปรผัน ผลการวิจัยพบว่า ในตัวอย่างขนาดเล็ก ตัวประมาณค่าสัดส่วน T_2 ดีกว่า T_1

นอกจากนี้ ยังได้เปรียบเทียบตัวประมาณค่า $\theta(1-\theta)$ 3 ตัว ได้แก่ $T_1 = (x/n)(1-x/n)$, $T_2 = x(n-x)/[n/(n-1)]$ และ $T_3 = [x(n-x)+n\sqrt{n}/2+n/4]/(n+\sqrt{n})^2$ ผลการวิจัยพบว่า T_1 เป็นตัวประมาณค่าที่ดีที่สุด

Tiensuwan et al. (2006) ได้เปรียบเทียบการประมาณค่าเฉลี่ยของประชากรปกติตัวแปรเดียว ($N(\mu,\sigma^2)$) ที่ไม่ทราบค่าความแปรปรวน ภายใต้การทดสอบสมมติฐานเกี่ยวกับค่าเฉลี่ยของประชากร โดยเปรียบเทียบตัวประมาณค่าเฉลี่ย 4 ชนิด ได้แก่ ตัวประมาณค่าที่ไม่เอนเอียงคือตัวประมาณค่าภาวะความน่าจะเป็นสูงสุด (Maximum Likelihood Estimator) และตัวประมาณค่าที่เอนเอียงอีก 3 ชนิด คือ ตัวประมาณค่าแบบกำกัด (Restricted Estimator, RE) ตัวประมาณค่าการทดสอบเบื้องต้น (Preliminary Test Estimator, PTE) และตัวประมาณค่าแบบย่อ (Shrinkage Estimator, SE) โดยใช้วิธี MCDM ผลปรากฏว่า ตัวประมาณค่าที่ไม่เอนเอียงเป็นตัวประมาณค่าที่ดีที่สุด และในบรรดาตัวประมาณค่าที่เอนเอียงทั้งหมด PTE และ SE ดีกว่า RE

Tiensuwan et al. (2006) ได้เปรียบเทียบตัวประมาณค่าที่ไม่เอนเอียง 4 แบบ ของค่าเฉลี่ยร่วมของประชากรสองกลุ่ม ได้แก่ $\hat{\mu}_1 = (\overline{x}/s_1^2+\overline{y}/s_2^2)/(1/s_1^2+1/s_2^2)$, $\hat{\mu}_2 = \overline{x}+D[(s_1^2+D^2)/(s_1^2+s_2^2+D^2)]$, $\hat{\mu}_3 = \overline{x}+D\min\{s_1^2/(s_1^2+s_2^2), s_2^2/(s_1^2+s_2^2)\}$ และ $\hat{\mu}_4 = \overline{x}+D[s_1/(s_1+s_2)]$ เมื่อ $D=\overline{y}-\overline{x}$ โดยใช้ MCDM ผลการวิจัยพบว่า ตัวประมาณค่าที่ดีที่สุดคือ $\hat{\mu}_4$ รองลงมาคือ $\hat{\mu}_1$

Lertprapai and Tiensuwan (2009) ประยุกต์ใช้ MCDM กับข้อมูลมลภาวะทางอากาศในกรุงเทพมหานคร ตั้งแต่ปี 2002 – 2004 โดยใช้ปริมาณสารมลพิษทางอากาศที่สำคัญ 5 ชนิด เป็นตัวชี้วัดเปรียบเทียบคุณภาพอากาศ 7 สถานี ผลการวิจัยพบว่า โดยส่วนใหญ่แล้วสถานีที่ดีที่สุดคือ โรงเรียนนนทรีวิทยา สถานีที่แย่ที่สุดคือการเคหะชุมชนดินแดง นอกจากนี้ ยังได้วิเคราะห์เปรียบเทียบแยกตามตามฤดูกาลพบว่า ฤดูร้อน ฝน และหนาว สถานีที่ดีที่สุดคือ โรงเรียนสิงหราชพิทยาคม โรงเรียนนนทรีวิทยา และสนามกีฬาเคหะชุมชนห้วยขวาง ตามลำดับ สำหรับสถานีที่แย่

ที่สุดคือ การเคหะชุมชนดินแดง บ้านพักสถานีตำรวจจราจร และมหาวิทยาลัยราชภัฏบ้านสมเด็จเจ้าพระยา ตามลำดับ

กระบวนการ MCDM

กระบวนการ MCDM เป็นกระบวนการการตัดสินใจที่มีการประเมินทางเลือกโดยพิจารณาจากเกณฑ์ หรือตัวชี้วัด หรือลักษณะเฉพาะ (Attribute) หลายอย่าง ในแต่ละทางเลือกจะมีค่าที่วัดได้ตามตัวชี้วัดต่าง ๆ ปัญหาที่ผู้ตัดสินใจพบคือทางเลือกหนึ่งมักจะมีค่าที่ดีในบางตัวชี้วัดในขณะที่บางตัวชี้วัดมีค่าไม่ดีนัก หลักการของ MCDM คือรวมตัวชี้วัดหลายตัวให้เป็นตัวชี้วัดเพียงตัวเดียวแล้วตัดสินใจเลือกทางเลือกโดยใช้ค่าของตัวชี้วัดดังกล่าว (Zeleny, 1982)

Yoon and Hwang (1995) ได้อธิบายถึงเทคนิคหนึ่งของ MCDM ที่ง่ายต่อการเข้าใจคือเทคนิคการจัดอันดับความพึงพอใจโดยใช้ความคล้ายคลึงกับผลเฉลยในอุดมคติ (Technique for Order Preference by Similarity to Ideal Solution, TOPSIS) เทคนิคนี้มีแนวคิดที่ว่า ทางเลือกที่ถูกเลือกควรมีระยะทางใกล้ผลเฉลยในอุดมคติบวก (Positive-Ideal Solution) มากที่สุด และไกลจากผลเฉลยในอุดมคติลบ (Negative-Ideal Solution) มากที่สุด ดังนั้นจึงได้กำหนดตัวชี้วัดจากระยะทางทั้งสองนี้

ผลเฉลยในอุดมคติบวก คือทางเลือกที่ประกอบด้วยค่าที่ดีที่สุดของแต่ละตัวชี้วัด ในกรณีที่เป็นตัวชี้วัดเชิงผลประโยชน์ (benefit) เช่น ผลกำไร คะแนนสอบ ค่านี้คือค่าสูงสุดของตัวชี้วัดนั้น ๆ กรณีที่เป็นตัวชี้วัดเชิงค่าใช้จ่าย (cost) เช่น ค่าใช้จ่าย ปริมาณน้ำตาลในเลือด หรือปริมาณสารพิษที่ปลดปล่อยมาในอากาศ ค่านี้คือค่าต่ำสุดของตัวชี้วัดนั้น ๆ

ผลเฉลยในอุดมคติลบ คือทางเลือกที่ประกอบด้วยค่าแย่ที่สุดของแต่ละตัวชี้วัด ในกรณีที่เป็นตัวชี้วัดเชิงผลประโยชน์ ค่านี้คือค่าต่ำสุดของตัวชี้วัดนั้น ๆ กรณีที่เป็นตัวชี้วัดเชิงค่าใช้จ่าย ค่านี้คือค่าสูงสุดของตัวชี้วัดนั้น ๆ

Yoon and Hwang (1995) ได้พิจารณาปัญหา MCDM ที่มีทางเลือก m ทางเลือก และถูกประเมินโดย n เกณฑ์ ให้อยู่ในระบบเรขาคณิตที่มี m จุด ในปริภูมิ n มิติ ให้ x_{ij} แทนข้อมูลของทางเลือกที่ i ตัวชี้วัดที่ j ขั้นตอนของ TOPSIS มีดังต่อไปนี้

1. คำนวณค่าบรรทัดฐาน (Normalized Rating) เป็นขั้นตอนการทำเวกเตอร์ให้เป็นบรรทัดฐาน (Vector Normalization) โดยการคำนวณ ดังนี้

$$r_{ij} = x_{ij} / (\sum_{i=1}^{m} x_{ij}^2)^{1/2} \quad , i = 1, 2, \ldots, m \; , \; j = 1, 2, \ldots, n$$

2. คำนวณค่าบรรทัดฐานถ่วงน้ำหนัก (Weighted Normalized Rating)

$v_{ij} = w_j r_{ij}$, $i = 1, 2, \ldots, m,\ j = 1, 2, \ldots, n$

โดยที่ w_j เป็นน้ำหนักของเกณฑ์ที่ j

3. กำหนดผลเฉลยในอุดมคติบวก และผลเฉลยในอุดมคติลบ ใช้สัญลักษณ์ A^* และ A^- กำหนดจากค่าบรรทัดฐานถ่วงน้ำหนัก ดังนี้

$$A^* = \{v_1^*, v_2^*, \ldots, v_n^*\}$$
$$= \{(\max_i v_{ij} \mid j \in J_1), (\min_i v_{ij} \mid j \in J_2) \mid i = 1,\ 2,\ \ldots,\ m\}$$
$$A^- = \{v_1^-, v_2^-, \ldots, v_n^-\}$$
$$= \{(\min_i v_{ij} \mid j \in J_1), (\max_i v_{ij} \mid j \in J_2) \mid i = 1,\ 2,\ \ldots,\ m\}$$

โดยที่ J_1 เป็นตัวชี้วัดเชิงผลประโยชน์ และ โดยที่ J_2 เป็นตัวชี้วัดเชิงค่าใช้จ่าย

4. คำนวณระยะทางระหว่างทางเลือกกับผลเฉลยในอุดมคติบวก (S_i^*) และระยะทางระหว่างทางเลือกกับผลเฉลยในอุดมคติลบ (S_i^-) ดังนี้

$$S_i^* = [\sum_{j=1}^{n} (v_{ij} - v_j^*)^2]^{1/2},\ i\ =\ 1,\ 2,\ \ldots,\ m$$

และ
$$S_i^* = [\sum_{j=1}^{n} (v_{ij} - v_j^-)^2]^{1/2},\ i\ =\ 1,\ 2,\ \ldots,\ m$$

5. คำนวณความคล้ายคลึงกับผลเฉลยในอุดมคติบวก

$$C_i^* = S_i^- / (S_i^* + S_i^-),\ i\ =\ 1,\ 2,\ \ldots,\ m$$

สังเกตได้ว่า $0 \le C_i^* \le 1$ โดยที่ $C_i^* = 0$ เมื่อ $A_i = A^-$ และ $C_i^* = 1$ เมื่อ $A_i = A^*$

6. จัดอันดับความพึงพอใจ โดยการเลือกทางเลือกที่มีค่า C_i^* สูงสุด หรือจัดอันดับทางเลือกตามลำดับของค่า C_i^*

Maitra et al. (2002) ได้สรุปวิธีการ TOPSIS เพื่อให้สะดวกในการนำไปใช้ โดยพิจารณาปัญหา MCDM ที่มีทางเลือก K ทาง และตัวชี้วัดทางสิ่งแวดล้อม N ตัว เขียนให้อยู่ในรูปเมตริกซ์ขนาด $K \times N$ ซึ่งมี K แถว และ N สดมภ์ โดยกำหนด x_{ij} แทนสมาชิกในเมตริกซ์แถวที่ i สดมภ์ที่ j ซึ่งหมายถึงค่าของทางเลือกที่ i ตัวชี้วัดที่ j สำหรับผลเฉลยในอุดมคติบวกหรือแถวในอุดมคติ (Ideal Row) เขียนโดยย่อว่า *IDR* และผลเฉลยในอุดมคติลบหรือแถวในอุดมคติลบ (Negative Ideal Row) เขียน โดยย่อว่า *NIDR* หาได้จาก

$$IDR = (\min_i x_{i1}, \min_i x_{i2}, \ldots, \min_i x_{iN}) = (u_1, u_2, \ldots, u_N)$$

และ
$$NIDR = (\max_i x_{i1}, \max_i x_{i2}, \ldots, \max_i x_{iN}) = (v_1, v_2, \ldots, v_N)$$

สำหรับระยะทางสามารถวัดได้โดยใช้ L_1- norm หรือ L_2- norm ดังนี้

ภายใต้ L_1 - norm ระยะทางจากทางเลือกที่ i ไปยัง *IDR* และระยะทางจากทางเลือกที่ i ไปยัง *NIDR* คำนวณได้จาก

$$L_1(i,IDR)=\sum_{j=1}^{N}[|x_{ij}-u_j|w_j/(\sum_{i=1}^{K}|x_{ij}|)]$$

และ
$$L_1(i,NIDR)=\sum_{j=1}^{N}[|x_{ij}-v_j|w_j/(\sum_{i=1}^{K}|x_{ij}|)]$$

เมื่อ w_j เป็นค่าถ่วงน้ำหนักที่เหมาะสม แถวต่าง ๆ จะถูกเปรียบเทียบโดยใช้ดัชนีโดยรวม ซึ่งคำนวณจาก

$$L_1(Index_i)=L_1(i,IDR)/[L_1(i,IDR)+L_1(i,NIDR)],\ i=1,...,K$$

ในทำนองเดียวกัน ภายใต้ L_2 – norm เราคำนวณ

$$L_2(i,IDR)=\{\sum_{j=1}^{N}[(x_{ij}-u_j)^2 w_j^2/(\sum_{i=1}^{K}x_{ij}^2)]\}^{1/2}$$

และ
$$L_2(i,NIDR)=\{\sum_{j=1}^{N}[(x_{ij}-v_j)^2 w_j^2/(\sum_{i=1}^{K}x_{ij}^2)]\}^{1/2}$$

และเปรียบเทียบแถวโดยใช้ดัชนีโดยรวม ซึ่งคำนวณจาก

$$L_2(Index_i)=L_2(i,IDR)/[L_2(i,IDR)+L_2(i,NIDR)],\ i=1,...,K$$

เราสามารถจัดอันดับทางเลือกจากค่าดัชนีโดยรวมนี้ ซึ่งทางเลือกที่ดีกว่าจะมีค่า *L(Index)* ต่ำ

กรณีที่ตัวชี้วัดเป็นลักษณะต่อเนื่อง คือมีตัวชี้วัด j เปลี่ยนไปเป็นลักษณะต่อเนื่อง Lertprapai et al. (2004) ได้กำหนดการคำนวณระยะทางจากทางเลือกที่ i ไปยัง *IDR* และ *NIDR* เป็นดังนี้

$$L_1(i,IDR)=\int_{all\ r}[x_i(r)-u(r)]w(r)dr$$

$$L_1(i,NIDR)=\int_{all\ r}[x_i(r)-v(r)]w(r)dr$$

$$L_2(i,IDR)=\left\{\int_{all\ r}[x_i(r)-u(r)]^2[w(r)]^2\,dr\right\}^{1/2}$$

$$L_2(i,NIDR)=\left\{\int_{all\ r}[x_i(r)-v(r)]^2[w(r)]^2\,dr\right\}^{1/2}$$

เมื่อ $x_i(r)$ เป็นฟังก์ชันขึ้นกับตัวแปร $r,\ i=1,2,...,K$

$u(r)=\min_i\{x_i(r)\}$ และ $v(r)=\max_i\{x_i(r)\}$

นอกจากนี้ ภายใต้ L_1 - norm ทางเลือกที่ i จะดีกว่า ทางเลือกที่ j ถ้า

$$\int_{all\ r} x_i(r)w(r)dr < \int_{all\ r} x_j(r)w(r)dr$$

และ ภายใต้ L_2 – norm ทางเลือกที่ i จะดีกว่า ทางเลือกที่ j ถ้า

$$\int_{all\ r} [x_i(r)-u(r)]^2[w(r)]^2\,dr \Big/ \int_{all\ r} [v(r)-x_i(r)]^2[w(r)]^2\,dr$$
$$< \int_{all\ r} [x_j(r)-u(r)]^2[w(r)]^2\,dr \Big/ \int_{all\ r} [v(r)-x_j(r)]^2[w(r)]^2\,dr$$

สำหรับงานวิจัยนี้จะใช้หลักการของ Maitra et al. (2002) กรณีข้อมูลที่มีตัวชี้วัดชนิดไม่ต่อเนื่อง และใช้หลักการของ Lertprapai et al. (2004) กรณีข้อมูลที่มีตัวชี้วัดชนิดต่อเนื่อง

การถ่วงน้ำหนัก

การถ่วงน้ำหนักเป็นการให้ความสำคัญของเกณฑ์หรือตัวชี้วัดแต่ละตัว (Zeleny, 1982) โดยพิจารณาจากข้อมูลของเกณฑ์ที่ใช้ในการตัดสินใจนั้น ๆ เช่น ใช้เอนโทรปี (Entropy) หรือสัมประสิทธิ์การแปรผันของตัวอย่าง เป็นต้น

เอนโทรปีเป็นค่าที่ Shannon and Weaver (1947) ใช้วัดสารสนเทศในทฤษฎีการสื่อสาร สำหรับปัญหา MCDM นั้น Filar et al. (1999) ให้เอนโทรปี (ϕ) เป็นค่าที่ใช้วัดการใกล้ชิดกันของข้อมูลในแต่ละเกณฑ์หรือตัวชี้วัด ซึ่งเอนโทรปีของตัวชี้วัดที่ j คำนวณได้จาก

$$\phi_j = -\sum_{i=1}^{K}[p_{ij}\ln(p_{ij})]/[ln(K)] \qquad \text{เมื่อ } p_{ij} = x_{ij} \Big/ \sum_{i=1}^{K} x_{ij}$$

หากเกณฑ์ใดมีค่า ϕ น้อยเกณฑ์นั้นมีความสำคัญต่อการประเมินผลทางเลือกมากควรให้น้ำหนักมาก ดังนั้น น้ำหนักสำหรับเกณฑ์หรือตัวชี้วัดที่ j จึงกำหนดเป็นดังนี้

$$w_j = (1-\phi_j) \Big/ \sum_{j=1}^{N}(1-\phi_j),\ j = 1,\ldots,N$$

สำหรับกรณีข้อมูลที่มีตัวชี้วัดเป็นลักษณะต่อเนื่อง Lertprapai et al. (2004) ได้ปรับปรุงสูตรการคำนวณค่าถ่วงน้ำหนักโดยใช้เอนโทรปี ดังนี้

$$w(r) = [1-\phi(r)] \Big/ \int_{all\ r}[1-\phi(r)]dr$$

โดยที่ $$\phi(r) = -(1/\ln K)\sum_{i=1}^{K}\{[x_i(r)/\sum_{i=1}^{K}x_i(r)]\cdot\ln[x_i(r)/\sum_{i=1}^{K}x_i(r)]\}$$

นอกจากนี้ Maitra et al. (2002) ยังได้เสนอการถ่วงน้ำหนักโดยใช้สัมประสิทธิ์การแปรผันตัวอย่างของข้อมูลในแต่ละเกณฑ์หรือตัวชี้วัด โดยมีสูตรการคำนวณดังนี้

$$w_j = CV_j = s_j / \overline{x}_j \quad \text{เมื่อ} \quad s_j = [\sum_{i=1}^{K} (x_{ij} - \overline{x}_j)^2 / (K-1)]^{1/2} \text{ และ } \overline{x}_j = \sum_{i}^{K} x_{ij} / K$$

วิธี ELECTRE

วิธี ELECTRE (ELimination Et Choix Traduisant la REalité) เป็นวิธีเปรียบเทียบทางเลือกคราวละ 2 ทางเลือกว่าทางเลือกใดดีกว่ากัน ไม่ได้เรียงอันดับทางเลือกทั้งหมดในคราวเดียวกัน คิดค้นโดย Roy (1971) ต่อมา Yoon and Hwang (1995) ได้สรุปแนวคิดที่ง่ายต่อการเข้าใจและวิเคราะห์เพิ่มเติมเพื่อให้สามารถจัดอันดับทางเลือกได้ ดังนี้

จากข้อมูล m ทางเลือก n เกณฑ์ เขียนแทนด้วยเมตริกซ์ $[x_{ij}]_{m \times n}$ เมื่อ x_{ij} หมายถึงข้อมูลทางเลือกที่ i เกณฑ์ที่ j

ขั้นตอนที่ 1 คำนวณค่าบรรทัดฐาน เป็นขั้นตอนการทำเวกเตอร์ให้เป็นบรรทัดฐาน (Vector Normalization) โดยการคำนวณ ดังนี้

$$r_{ij} = x_{ij} / (\sum_{i=1}^{m} x_{ij}^2)^{1/2}, i = 1,2,...,m, j = 1,2,...,n$$

ขั้นตอนที่ 2 คำนวณค่าบรรทัดฐานถ่วงน้ำหนัก

$v_{ij} = w_j r_{ij}, i = 1,2,...,m, j = 1,2,...,n$ โดยที่ w_j เป็นน้ำหนักของเกณฑ์ที่ j, และ $\sum_{j=1}^{n} w_j = 1$

ขั้นตอนที่ 3 หาเซตของความสอดคล้องและไม่สอดคล้อง (Concordance and Discordance) โดยในแต่ละสองทางเลือก A_p และ A_q ($p,q = 1,2,...,m$ และ $p \neq q$) เซตของเกณฑ์จะถูกแบ่งเป็น 2 เซตย่อยที่แตกต่างกัน คือเซตของความสอดคล้อง $C(p,q)$ จะประกอบไปด้วยเกณฑ์ที่มีค่าของทางเลือกที่ p ดีกว่าทางเลือกที่ q และเซตของความไม่สอดคล้อง $D(p,q)$ จะประกอบไปด้วยเกณฑ์ที่มีค่าของทางเลือกที่ p แย่กว่าทางเลือกที่ q เขียนสัญลักษณ์ได้ดังนี้ $C(p,q) = \{j \mid v_{pj} \geq v_{qj}\}$ และ $D(p,q) = \{j \mid v_{pj} < v_{qj}\}$ สังเกตได้ว่า $C(p,q) \neq D(q,p)$ ถ้ามีบาง j ที่ทำให้ $v_{pj} = v_{qj}$

ขั้นตอนที่ 4 หาดัชนีความสอดคล้องและไม่สอดคล้อง ซึ่งเป็นดัชนีที่ใช้วัดระดับความสอดคล้องและความไม่สอดคล้องของการเปรียบเทียบทางเลือกที่ p ดีกว่าทางเลือกที่ q ($A_p \rightarrow A_q$) คำนวณโดย

$$C_{pq} = \sum_{j^*} w_{j^*}, \quad j^* \in C(p,q)$$

และ
$$D_{pq} = (\sum_{j^\circ} |v_{pj^\circ} - v_{qj^\circ}|)/(\sum_{j} |v_{pj} - v_{qj}|), \; j^\circ \in D(p,q)$$

ขั้นตอนที่ 5 เปรียบเทียบทางเลือกคราวละสองทางเลือก ทางเลือก p จะดีกว่า q ถ้า $C_{pq} \geq \overline{C}$ และ $D_{pq} < \overline{D}$ เมื่อ $\overline{C}$ และ $\overline{D}$ เป็นค่าเฉลี่ยของ C_{pq} และ D_{pq} ตามลำดับ

หากต้องการจัดอันดับทางเลือกทั้งหมด เราจะใช้ดัชนีสุทธิ (Net Indexes) ของทางเลือกที่ p สองชนิด คือดัชนีความสอดคล้องสุทธิ (Net Concordance Index, C_p) และดัชนีความไม่สอดคล้องสุทธิ (Net Discordance Index, D_p) โดยทางเลือกที่ p จะดีกว่า ถ้า C_p มีค่าสูงกว่า และ D_p มีค่าต่ำกว่า ดังนั้น เราสามารถจัดอันดับทางเลือกได้โดยจัดอันดับค่า C_p จากมากไปหาน้อย และจัดอันดับค่า D_p จากน้อยไปหามาก ทางเลือกที่ดีที่สุดคือทางเลือกที่มีค่าเฉลี่ยของอันดับดีที่สุด สำหรับดัชนีสุทธิคำนวณได้จาก

$$C_p = \sum_{\substack{k=1 \\ k \neq p}}^{K} C_{pk} - \sum_{\substack{k=1 \\ k \neq p}}^{K} C_{kp}$$

และ
$$D_p = \sum_{\substack{k=1 \\ k \neq p}}^{K} D_{pk} - \sum_{\substack{k=1 \\ k \neq p}}^{K} D_{kp}$$

ข้อมูลการตรวจวัดคุณภาพอากาศ

สำนักจัดการคุณภาพอากาศและเสียง กรมควบคุมมลพิษ กระทรวงทรัพยากรธรรมชาติและสิ่งแวดล้อม ได้ติดตั้งสถานีตรวจวัดคุณภาพอากาศในบรรยากาศ ซึ่งประกอบด้วยสถานีตรวจวัดคุณภาพอากาศแบบกึ่งถาวร (สามารถเคลื่อนย้ายได้ถ้าจำเป็น) และแบบเคลื่อนที่ สิ่งที่ทำการตรวจวัดได้แก่ ฝุ่นละอองขนาดต่าง ๆ คาร์บอนมอนนอกไซด์ ซัลเฟอร์ไดออกไซด์ ออกไซด์ของไนโตรเจน ไฮโดรคาร์บอน โอโซน ความเร็วของกระแสลม อุณหภูมิของอากาศ ปริมาณการแผ่รังสีของแสงจากดวงอาทิตย์ และแหล่งอื่น ๆ ความดันของบรรยากาศ ปริมาณน้ำฝน และระดับเสียง มลพิษที่สำคัญที่ใช้ในการคำนวณดัชนีคุณภาพอากาศ (Air Quality Index, AQI) ในประเทศไทย มีอยู่ 5 ชนิด คือ โอโซน เฉลี่ย 1 ชั่วโมง ไนโตรเจนไดออกไซด์ เฉลี่ย 1 ชั่วโมง คาร์บอนมอนอกไซด์เฉลี่ย 8 ชั่วโมง ซัลเฟอร์ไดออกไซด์ เฉลี่ย 24 ชั่วโมง และฝุ่นละอองขนาดเล็กกว่า 10 ไมครอน (PM_{10}) เฉลี่ย 24 ชั่วโมง ซึ่งคำนวณโดยเทียบจากมาตรฐานคุณภาพอากาศในบรรยากาศโดยทั่วไปของสารมลพิษทางอากาศ 5 ประเภทดังกล่าว ทั้งนี้ ดัชนีคุณภาพอากาศที่คำนวณได้ของสารมลพิษทางอากาศประเภทใดมีค่าสูงสุด จะใช้เป็นดัชนีคุณภาพอากาศของวันนั้น เกณฑ์ของดัชนีคุณภาพอากาศ

สำหรับประเทศไทย คือ 0-50 คุณภาพดี 51-100 คุณภาพปานกลาง 101-200 มีผลกระทบต่อสุขภาพ 201-300 มีผลกระทบต่อสุขภาพมาก มากกว่า 300 อันตราย

สำนักจัดการคุณภาพอากาศและเสียงได้รายงานคุณภาพอากาศพื้นที่กรุงเทพมหานครปริมณฑล และต่างจังหวัด แยกตามสถานีเป็นรายวัน รายเดือน รายปี สำหรับในพื้นที่ต่างจังหวัดสถานีที่มีอยู่ในปัจจุบัน มีทั้งสิ้น 28 สถานี ดังนี้

1. ศูนย์ราชการรวม อำเภอแม่ริม จังหวัดเชียงใหม่
2. โรงเรียนยุพราชวิทยาลัย อำเภอเมือง จังหวัดเชียงใหม่
3. ศาลหลักเมือง อำเภอเมือง จังหวัดลำปาง
4. สถานีอนามัยสบป้าด อำเภอแม่เมาะ จังหวัดลำปาง
5. สถานีอนามัยท่าสี อำเภอแม่เมาะ จังหวัดลำปาง
6. สำนักงานการประปาส่วนภูมิภาคแม่เมาะ อำเภอแม่เมาะ จังหวัดลำปาง
7. วิทยาลัยอาชีวศึกษานครสวรรค์ อำเภอเมือง จังหวัดนครสวรรค์
8. สำนักงานทรัพยากรธรรมชาติและสิ่งแวดล้อมจังหวัดเชียงราย อำเภอเมือง จังหวัดเชียงราย
9. สำนักงานทรัพยากรธรรมชาติและสิ่งแวดล้อมจังหวัดแม่ฮ่องสอน อำเภอเมือง จังหวัดแม่ฮ่องสอน
10. บ้านพักปลัดอำเภอ อำเภอเมือง จังหวัดขอนแก่น
11. บ้านพักทหารมณฑลทหารบกที่ 21 อำเภอเมือง จังหวัดนครราชสีมา
12. โรงเรียนอยุธยาวิทยาลัย อำเภอพระนครศรีอยุธยา จังหวัดพระนครศรีอยุธยา
13. สถานีตำรวจภูธรตำบลหน้าพระลาน (เดิมเป็นโรงเรียนหน้าพระลาน) อำเภอเฉลิมพระเกียรติ จังหวัดสระบุรี
14. สถานีดับเพลิงเขาน้อย อำเภอเมือง จังหวัดสระบุรี
15. ศูนย์วิศวกรรมการแพทย์ที่ 1 (เดิมเป็นศูนย์ช่างบำรุงที่ 1) อำเภอเมือง จังหวัดราชบุรี
16. องค์การบริหารส่วนตำบลตาสิทธิ์ อำเภอปลวกแดง จังหวัดระยอง
17. สถานีอนามัยมาบตาพุด อำเภอเมือง จังหวัดระยอง
18. ชุมสายโทรศัพท์ อำเภอเมือง จังหวัดระยอง
19. ศูนย์วิจัยพืชไร่ อำเภอเมือง จังหวัดระยอง
20. สนามกีฬาเทศบาลแหลมฉบัง อำเภอศรีราชา จังหวัดชลบุรี
21. ศูนย์เยาวชนเทศบาล อำเภอศรีราชา จังหวัดชลบุรี

22. สำนักงานสามัญศึกษา อำเภอเมือง จังหวัดชลบุรี

23. องค์การบริหารส่วนตำบลวังน้ำเย็น อำเภอแปลงยาว จังหวัดฉะเชิงเทรา

24. ที่ว่าการอำเภอเมือง จังหวัดสุราษฎร์ธานี

25. ศูนย์บริการสาธารณสุข เทศบาลนครภูเก็ต อำเภอเมือง จังหวัดภูเก็ต

26. ศูนย์ส่งเสริมสิ่งแวดล้อมเทศบาลนครหาดใหญ่ อำเภอหาดใหญ่ จังหวัดสงขลา

27. ศาลากลาง อำเภอเมือง จังหวัดนราธิวาส

28. สนามโรงพิธีช้างเผือก อำเภอเมือง จังหวัดนราธิวาส

งานวิจัยนี้จะเปรียบเทียบคุณภาพอากาศโดยใช้สารมลพิษสำคัญ 5 ชนิด เป็นเกณฑ์ และเนื่องจากสถานีตรวจวัดคุณภาพอากาศในจังหวัดเชียงราย แม่ฮ่องสอน และนราธิวาสไม่มีการตรวจวัดซัลเฟอร์ไดออกไซด์ และไนโตรเจนไดออกไซด์ จึงเปรียบเทียบเฉพาะสถานีตรวจวัดคุณภาพอากาศเพียง 24 สถานี ใน 14 จังหวัด ตั้งแต่ปี พ.ศ. 2548 – 2551 (ดูรายละเอียดเพิ่มเติมได้จาก http://www.aqnis.pcd.go.th/station/allstation.htm)

การแจกแจงแบบเลขชี้กำลัง

ตัวแปรสุ่ม X มีการแจกแจงแบบเลขชี้กำลัง และ θ เป็นพารามิเตอร์ โดย $\theta > 0$ มีฟังก์ชันความหนาแน่นของความน่าจะเป็น (Probability Density Function) ดังนี้

$$f(x) = \begin{cases} (1/\theta)e^{-(x/\theta)} & ; x > 0 \\ 0 & ; x \leq 0 \end{cases}$$

ค่าคาดหมาย $E(X) = \theta$ และความแปรปรวน $V(X) = \theta^2$

กรณีที่เราทราบว่าตัวแปรสุ่ม X มีค่ามากกว่า γ โดยที่ γ เป็นจำนวนจริงใด ๆ แล้ว X จะมีการแจกแจงแบบเลขชี้กำลังชนิดสองพารามิเตอร์ คือ θ เป็นพารามิเตอร์แสดงระยะห่าง (Scale Parameter) และ γ เป็นพารามิเตอร์แสดงตำแหน่ง (Location Parameter) มีฟังก์ชันความหนาแน่นของความน่าจะเป็น ดังนี้

$$f(x) = \begin{cases} (1/\theta)e^{-(x-\gamma)/\theta} & ; x > \gamma \\ 0 & ; x \leq \gamma \end{cases}$$

ค่าคาดหมาย $E(X) = \mu = \gamma + \theta$ และความแปรปรวน $V(X) = \theta^2$

1. การประมาณค่า θ และ μ ในการแจกแจงแบบแบบเลขชี้กำลังชนิดสองพารามิเตอร์

ให้ $x_1, x_2, \ldots, x_n$ เป็นค่าสังเกตที่ได้จากการสุ่มตัวอย่างขนาด n จากประชากรที่มีการแจกแจงแบบเลขชี้กำลัง ตัวประมาณค่าภาวะความน่าจะเป็นสูงสุดของ θ และ μ คือ

$$\hat{\theta}_{MLE} = \sum_{i=1}^{n}(x_i - x_{(1)})/n \text{ เมื่อ } x_{(1)} = \min\{x_i / i = 1, 2, ..., n\}$$

$$MSE(\hat{\theta}_{MLE}) = \theta^2 / n$$

$$\hat{\mu}_{MLE} = \overline{X} \quad \text{เมื่อ } \overline{X} = \sum_{i=1}^{n} x_i$$

$$MSE(\hat{\mu}_{MLE}) = \theta^2 / n$$

ตัวประมาณค่าแบบไม่เอนเอียงของ θ และ μ คือ

$$\hat{\theta}_{UE} = \sum_{i=1}^{n}(x_i - x_{(1)})/(n-1)$$

$$MSE(\hat{\theta}_{UE}) = \theta^2 / n$$

$$\hat{\mu}_{UE} = \overline{X}$$

$$MSE(\hat{\mu}_{UE}) = \theta^2 / n$$

ตัวประมาณค่าที่มีความคลาดเคลื่อนกำลังสองเฉลี่ยต่ำสุดของ θ ที่อยู่ในคลาสของตัวประมาณค่าที่ไม่แปรเปลี่ยน (Class of Invariant Estimator) $c\cdot(\overline{x} - x_{(1)})$ คือ $\hat{\theta}_{MLE}$

ตัวประมาณค่าที่มีความคลาดเคลื่อนกำลังสองเฉลี่ยต่ำสุดของ μ ที่อยู่ในคลาสของตัวประมาณค่าที่ไม่แปรเปลี่ยน $x_{(1)} + c\cdot(\overline{x} - x_{(1)})$ คือ

$$\hat{\mu}_{MMSE} = x_{(1)} + (n-1)(\overline{x} - x_{(1)})/n$$

$$MSE(\hat{\mu}_{MMSE}) = \theta^2 / n^2 + (n-1)^2 \theta^2 / n^3$$

ตัวประมาณค่าแบบย่อ (Shrinkage estimators) กรณีที่ทราบค่าของ θ เบื้องต้นเป็น θ_0 Kourouklis (1994) ได้เสนอคลาสของตัวประมาณแบบย่อสำหรับ θ และ μ ดังนี้

$$\hat{\theta}_{(p)} = \theta_0 + \alpha(p)\cdot(\hat{\theta}_{UE} - \theta_0)$$

$$MSE(\hat{\theta}_{(p)}) = [(1-\alpha(p))^2 r^2 + \alpha^2(p)/(n-1)]\theta^2$$

$$\hat{\mu}_{(p)} = x_{(1)} + (n-1)\hat{\theta}_{(p)}/n$$

$$MSE(\hat{\mu}_{(p)}) = \theta^2 / n^2 + [(n-1)^2 / n^2]\cdot MSE(\hat{\theta}_{(p)})$$

เมื่อ p เป็นจำนวนจริงที่ไม่เท่ากับศูนย์ และ $p \in (-\infty, (n-1)/2)$ ซึ่ง Kourouklis (1994) ได้พิจารณาที่ $p =$ -2,-1,1,2 $\qquad \alpha(p) = \Gamma(n-1-p)/[\Gamma(n-1-2p)\cdot(n-1)^p]$ และ $r = (\theta_0 / \theta) - 1$

2. การประมาณค่าความแปรปรวน θ^2

2.1 การแจกแจงแบบเลขชี้กำลังชนิดหนึ่งพารามิเตอร์

ตัวประมาณค่าที่ไม่เอนเอียงสำหรับ θ^2 คือ

$$\hat{\theta}^2_{1UE} = [n/(n+1)]\overline{x}^2$$

$$MSE(\hat{\theta}^2_{1UE}) = 2(2n+3)\theta^4 / [n(n+1)]$$

ในบรรดาตัวประมาณค่าที่อยู่ในรูป $M\overline{X}^2$ เมื่อ M เป็นค่าคงตัว Pandy and Singh (1977) พบว่า ที่ $M = n^2 / [(n+2)(n+3)]$ จะได้ตัวประมาณค่าที่มีความคลาดเคลื่อนกำลังสองเฉลี่ยต่ำสุด ดังนี้

$$\hat{\theta}^2_{1MMSE} = n^2\overline{x}^2 / [(n+2)(n+3)]$$

$$MSE(\hat{\theta}^2_{1MMSE}) = 2(2n+3)\theta^4 / [(n+2)(n+3)]$$

นอกจากนี้ Pandy and Singh (1977) ยังได้เสนอตัวประมาณที่มีความคลาดเคลื่อนกำลังสองเฉลี่ยต่ำสุด ที่อยู่ในรูปของ $M'S^2$ เมื่อ M' เป็นค่าคงตัว โดยที่ $M' = [n(n-1)] / (n^2 + 7n - 6)$ จะได้ตัวประมาณค่าดังกล่าว ดังนี้

$$\hat{\theta}^2_{1MMSE*} = [n(n-1)s^2] / (n^2 + 7n - 6)$$

เมื่อ $s^2 = \sum_{i=1}^{n} (x_i - \overline{x})^2 / (n-1)$

$$MSE(\hat{\theta}^2_{1MMSE*}) = (8n-6)\theta^4 / (n^2 + 7n - 6)$$

ซึ่ง $\hat{\theta}^2_{1MMSE}$ เป็นตัวประมาณที่มีประสิทธิภาพมากที่สุดในบรรดา $\hat{\theta}^2_{1UE}$, $\hat{\theta}^2_{1MMSE}$ และ s^2

กรณีที่ทราบค่าของ θ^2 เบื้องต้นว่าเป็น θ_0^2 Tracy, Singh and Raghuvarishi (1996) ได้เสนอคลาสของตัวประมาณแบบย่อสำหรับ θ^2 ดังนี้

$$\hat{\theta}^2_{1(p)} = \theta_0^2 + \alpha(p)(\overline{x}^2 - \theta_0^2)$$

$$\begin{aligned} MSE(\hat{\theta}^2_{1(p)}) = \theta^4 \{ & \lambda^2 (1-\alpha(p))^2 \\ & - 2\lambda[1 + (n+1)\alpha(p)^2 / n - (2n+1)\alpha(p)/n] \\ & + 1 - 2(n+1)\alpha(p)/n + (n+1)(n+2)(n+3)\alpha(p)^2 / n^3 \} \end{aligned}$$

เมื่อ p เป็นจำนวนจริงที่ไม่เท่ากับศูนย์, $\alpha(p) = n^{2p}\Gamma(n+2p)/\Gamma(n+4p)$ และ $\lambda = \theta_0^2 / \theta^2$

นอกจากนี้ Tracy et al. (1996) ยังได้เสนอตัวประมาณค่าแบบย่ออีกหนึ่งตัวคือ

$$\hat{\theta}^2_s = \theta_0^2 + \{[n(n-1)]/[n^2 + 7n - 6]\}(s^2 - \theta_0^2)$$

$$MSE(\hat{\theta}^2_s) = MSE(\hat{\theta}^2_{1MMSE*}) + \lambda(\lambda - 2)(8n-6)^2\theta^4 / (n^2 + 7n - 6)^2$$

ซึ่ง $MSE(\hat{\theta}^2_s) < MSE(\hat{\theta}^2_{1MMSE*})$ เมื่อ $0 < \lambda < 2$ หรือ $\theta_0^2 / 2 \leq \theta^2 < \infty$

ดังนั้น หากพิจารณาช่วง $0 < \lambda < 2$ แล้ว ตัวประมาณค่าที่จะนำมาเปรียบเทียบกัน จึงมีเพียง $\hat{\theta}^2_{1MMSE}$, $\hat{\theta}^2_{1(p)}$ และ $\hat{\theta}^2_{s}$

2.2 การแจกแจงแบบเลขชี้กำลังชนิดสองพารามิเตอร์

ตัวประมาณค่าที่ไม่เอนเอียงสำหรับ θ^2 ได้แก่

$$\hat{\theta}^2_{2UE} = n(\overline{x} - x_{(1)})^2 / (n-1)$$

$$MSE(\hat{\theta}^2_{2UE}) = 2(2n+1)\theta^4 / [n(n-1)]$$

ในบรรดาตัวประมาณที่อยู่ในรูป $M^*(\overline{x} - x_{(1)})^2$ เมื่อ M^* เป็นค่าคงตัว Pandy and Singh (1977) พบว่าที่ $M^* = n^2 / [(n+1)(n+2)]$ จะได้ตัวประมาณค่าที่มีความคลาดเคลื่อนกำลังสองเฉลี่ยต่ำสุด ดังนี้

$$\hat{\theta}^2_{2MMSE} = [n^2(\overline{x} - x_{(1)})^2] / [(n+1)(n+2)]$$

$$MSE(\hat{\theta}^2_{2MMSE}) = 2(2n+1)\theta^4 / [(n+1)(n+2)]$$

กรณีที่ทราบค่าของ θ^2 เบื้องต้นว่าคือ θ^2_0 Tracy et al. ได้เสนอตัวประมาณค่าแบบย่อสำหรับ θ^2_0 ดังนี้

$$\hat{\theta}^2_{2(p)} = \theta^2_0 + \beta(p)(Z^2 - \theta^2_0)$$

$$\begin{aligned} MSE(\hat{\theta}^2_{2(p)}) = \theta^4\{&\lambda^2(1-\beta(p))^2 \\ &-2\lambda \times [1+(n-1)\beta(p)^2 / n - (2n-1)\beta(p)/n] \\ &+[1+(n-1)(n+1)(n+2)\beta(p)^2 / n^3 - 2(n-1)\beta(p)/n]\} \end{aligned}$$

โดยที่ $Z = (\overline{x} - x_{(1)})$ และ $\beta(p) = n^{2q}\Gamma(n+2p-1)/\Gamma(n+4p-1)$ เมื่อ p เป็นจำนวนจริงที่ไม่เท่ากับศูนย์