บทที่ 3

หลักวิธีการทางสถิติและผลของการวิเคราะห์ของต่างประเทศ

การวิจัยการวัดผลผลิตของการทำโปรแกรมยังอยู่ในระยะที่เริ่มศึกษาค้นคว้า ดังนั้น การวิจัยในระยะแรกเริ่มนี้จึงมุ่งที่การวิเคราะห์สหสัมพันธ์และการถดถอยของตัวแปรผันต่าง ๆ ว่าตัวแปรใดมีความสัมพันธ์ใกล้ชิดกันเพียงใด และการเปลี่ยนแปลงที่เกิดขึ้นกับตัวแปรผันตัวหนึ่ง หนึ่งหน่วยจะทำให้ตัวผันแปรอีกตัวหนึ่งเปลี่ยนแปลงไปเป็นจำนวนเท่าใด วิธีการที่ใช้วัดความสัมพันธ์ต่าง ๆ เหล่านี้ได้แก่ "การถดถอยแบบเส้นตรง" (Linear regression) เมื่อได้ความสัมพันธ์ระหว่างตัวแปรผัน 2 ตัว เราจะใช้ตัวแปรผันตัวหนึ่งในการคาดคะเนล่วงหน้า หรือควบคุมตัวแปรผันอีกตัวหนึ่ง เช่น ระหว่างตัวแปรผัน "จำนวนบรรทัดของโปรแกรม" และตัวแปรผัน "จำนวนคน-เดือน" (man - month) หากเราทราบความสามารถของผู้ทำว่าสามารถผลิตได้กี่บรรทัดในหนึ่งเดือน เราก็สามารถคาดคะเนระยะเวลาและจำนวนผู้ทำโปรแกรมได้ว่าต้องใช้เวลาเท่าไร แต่ในการนี้เราต้องรู้เทคนิคในการคาดคะเนล่วงหน้า และการวัดความผิดพลาดที่อาจเกิดขึ้นกับการคาดคะเนล่วงหน้าของเรา เทคนิคเหล่านี้รวมเรียกว่า "การวิเคราะห์การถดถอย" (Regression analysis) จากความสัมพันธ์ที่ได้จากการวิเคราะห์การถดถอย เราจะนำไปใช้อธิบายเกี่ยวกับการคาดคะเนได้ดีเพียงใด เรามีมาตรการที่ใช้วัดความใช้ได้ของความสัมพันธ์ดังกล่าว เรียกว่า "สัมประสิทธิ์สหสัมพันธ์" หรือ "r" (correlation coefficient) ค่า (r) จะอยู่ระหว่างลบหนึ่งถึงบวกหนึ่ง. ถ้า r = 0 แสดงว่าตัวแปรผันทั้งสองไม่มีสหสัมพันธ์เลย แต่ถ้าค่าของ (r) เข้าใกล้บวกหนึ่งหรือลบหนึ่งมากเพียงใด แสดงถึงสหสัมพันธ์ของตัวแปรผันมีมากจนเกือบจะเรียกว่า มีสหสัมพันธ์ที่สมบูรณ์ เครื่องหมายของ "r" เหมือนของ "b" ในสมการถดถอย ถ้า r = -1 จุดทุกจุดจะอยู่บนเส้นถดถอยที่ลาดจากซ้ายไปขวา ในการปฏิบัติใช้ (r**2) (coefficient of determination) เป็นตัวพิจารณาซึ่งใช้ได้ดีกว่าสัมประสิทธิ์สหสัมพันธ์ เพราะเป็นตัวเลขที่แสดงให้เห็นสัดส่วนของค่าแปรปรวนของ Y ที่ขึ้นอยู่กับ X ซึ่งเป็นตัวแปรผันอิสระ เนื่องจาก

สัมประสิทธิ์สหสัมพันธ์มักจะชี้ให้เห็นสหสัมพันธ์ที่มีค่าสูงกว่าความเป็นจริง เช่น ถ้าค่าแปรปรวนของ Y ร้อยละ 50 อธิบายได้โดย X คือ $r^2 = 0.50$ แต่สัมประสิทธิ์สหสัมพันธ์ คือ $r = 0.71$ ซึ่งมีค่าสูงกว่า ดังนั้น จึงนิยมใช้ (r**2) เป็นตัวพิจารณาที่ดีกว่า

เนื่องจากข้อมูลที่รวบรวมมีช่วงกว้างมาก เช่น จำนวนบรรทัดของโปรแกรมมีช่วงระหว่าง 1,000 – 1,000,000 บรรทัดของโปรแกรม หรือจำนวนแรงงานที่ใช้ในการพัฒนาโปรแกรมมีช่วงระหว่าง 10 – 10,000 (man - month) ดังนั้น ถ้าหากว่าเราจะนำเอาข้อมูลเหล่านี้มาแทนค่าลงในกราฟ เพื่อต้องการหาความสัมพันธ์ก็คงจะไม่สดวก จึงต้องใช้ เซมิลอกกราฟ (semilog graph) ในการแทนข้อมูลเหล่านี้

จากการวิจัยของ IBM และ SEL การหาความสัมพันธ์ของตัวแปรผันจะอยู่ในรูปของสมการเส้นตรง (y = a + bx) หรือสมการเอกซ์โปเนนท์เชียล (y = a * (x**b)) ซึ่งก็มีความสัมพันธ์แบบเส้นตรงเช่นกัน คือ

$$y = ax^b$$

take log $\log y = \log a + b \log x$

ซึ่งสามารถเขียนได้ ดังนี้ $Y = A + BX$ ซึ่งเป็นสูตรเดียวกับสมการเส้นตรงโดย

$$\log y = Y \ , \ \log a = A \ , \ \log x = X$$

ในการคำนวณหาความสัมพันธ์ระหว่างตัวแปรผัน กรณีที่มีความสัมพันธ์แบบเอ็กซ์-โปเนนท์เชียล ขั้นแรก ข้อมูลจะถูกทำให้มีลักษณะเส้นตรงโดยการ take log แล้วใช้วิธีกำลังสองน้อยที่สุดในการหาสมการถดถอยแบบเส้นตรง และสัมประสิทธิ์ของเส้นตรงที่ได้ก็จะมีความสัมพันธ์แบบเส้นตรง เมื่อแปลงกลับอีกที

3.1 ผลการวิจัยของ IBM

3.1.1 <u>ความสัมพันธ์ระหว่างแรงงาน–เวลาที่ใช้ไปในการพัฒนาโปรแกรม และจำนวนบรรทัดของโปรแกรม</u> ดังรูป 3.1.1 ความสัมพันธ์ที่ได้จากข้อมูลที่รวบรวม คือ E = 5.2 *(L ** 0.91)

ความสัมพันธ์ที่ได้มีค่าใกล้เคียงกำลังหนึ่งนั้น คือความสัมพันธ์นี้มีค่าเป็นเส้นตรง ส่วนเส้นประเป็นความคลาดเคลื่อนของการกะประมาณ ณ ที่จำนวนบรรทัดของโปรแกรม

มีขนาด 10,000 บรรทัด จะพบว่าแรงงาน–เวลา ที่ใช้ไปในการผลิตอาจแปรผันได้ตั้งแต่ 17 ถึง 106 คน–เดือน คือความคลาดเคลื่อนของการกะประมาณแตกต่างกันถึง 6.3 เท่า ความแตกต่างนี้อาจเนื่องมาจาก

1. เครื่องคอมพิวเตอร์ที่ใช้ต่างกันถึง 66 เครื่อง
2. ภาษาที่ใช้ต่างกันถึง 28 ภาษา
3. ความยากง่ายของโปรแกรมที่ต่างกันมาก

3.1.2 <u>ความสัมพันธ์ระหว่างจำนวนเอกสารและจำนวนบรรทัดของโปรแกรม</u> ดังรูป 3.1.2 ความสัมพันธ์ที่ได้จากข้อมูลที่รวบรวม DØC = 49* L** 1.01

การจัดทำเอกสารเป็นผลผลิตที่ต้องใช้ความระมัดระวังของทุก ๆ โครงการด้านซอฟท์แวร์และค่าใช้จ่ายในการจัดทำเอกสารก็เป็นส่วนสำคัญอันหนึ่งของการคาดคะเนกระบวนการผลิตซอฟท์แวร์ การวัดจำนวนเอกสารเราใช้วิธีนับเป็นแผ่น เอกสารต่าง ๆ เหล่านี้ ได้แก่ ข้อมูลจำเพาะของหน้าที่ของแต่ละโปรแกรมและรายละเอียดต่าง ๆ คำแนะนำสำหรับผู้ใช้ การทดสอบข้อมูลจำเพาะและผลลัพธ์ ผังการทำงาน และ source listing

3.1.3 <u>ความสัมพันธ์ระหว่างระยะเวลาของการทำโครงการและจำนวนบรรทัดของโปรแกรม</u> ดังรูป 3.1.3

ความสัมพันธ์ที่ได้จากข้อมูลที่รวบรวม D = 4.1* L** 0.36

3.1.4 <u>ความสัมพันธ์ระหว่างระยะเวลาของโครงการและจำนวนแรงงาน–เวลาที่ใช้ไปทั้งหมด</u> ดังรูป 3.1.4

ความสัมพันธ์ที่ได้จากข้อมูลที่รวบรวม D = 2.47 * (E** 0.35)

การคาดคะเนระยะเวลาที่ใช้ไปในการพัฒนาซอฟท์แวร์เป็นสิ่งที่กำหนดได้ยาก จากความสัมพันธ์ในหัวข้อ 3.1.3 และ 3.1.4 แสดงให้เห็นว่า ระยะเวลาของโครงการแปรผันเป็นรากที่สามของจำนวนบรรทัดของโปรแกรมและจำนวนแรงงาน–เวลาในการพัฒนาซอฟท์แวร์

3.1.5 <u>ความสัมพันธ์ระหว่างค่าใช้จ่ายทางคอมพิวเตอร์และจำนวนบรรทัดของโปรแกรม</u> ดังรูป 3.1.5

ความสัมพันธ์ที่ได้จากข้อมูลที่รวบรวม C = 1.84* (L ** 0.96)

การกะประมาณค่าใช้จ่ายทางคอมพิวเตอร์เป็นสิ่งที่ทำได้ยากมาก แต่ก็เป็นส่วนสำคัญมากเช่นกัน มีเพียง 18 โครงการเท่านั้นที่มีการทำรายงานเหล่านี้ และเก็บไว้ในฐานข้อมูล มี 2 โครงการที่อยู่นอกขอบเขตของความเบี่ยงเบนมาตรฐาน. ข้อมูลซึ่งถูกตัดออกไปไม่นำมาใช้ในการคำนวณวิธีกำลังสองน้อยที่สุด

3.1.6 <u>ความสัมพันธ์ระหว่างค่าใช้จ่ายทางคอมพิวเตอร์ และ แรงงาน-เวลาทั้งหมด</u> ดังรูป 3.1.6

ความสัมพันธ์ที่ได้จากข้อมูลที่รวบรวม C = 1.1 * (E** 0.81)

จากข้อ 3.1.5 มี 2 โครงการซึ่งออกนอกขอบเขตของความเบี่ยงเบนมาตรฐาน (วงกลมโปร่งมีจุดตรงกลางในรูป)

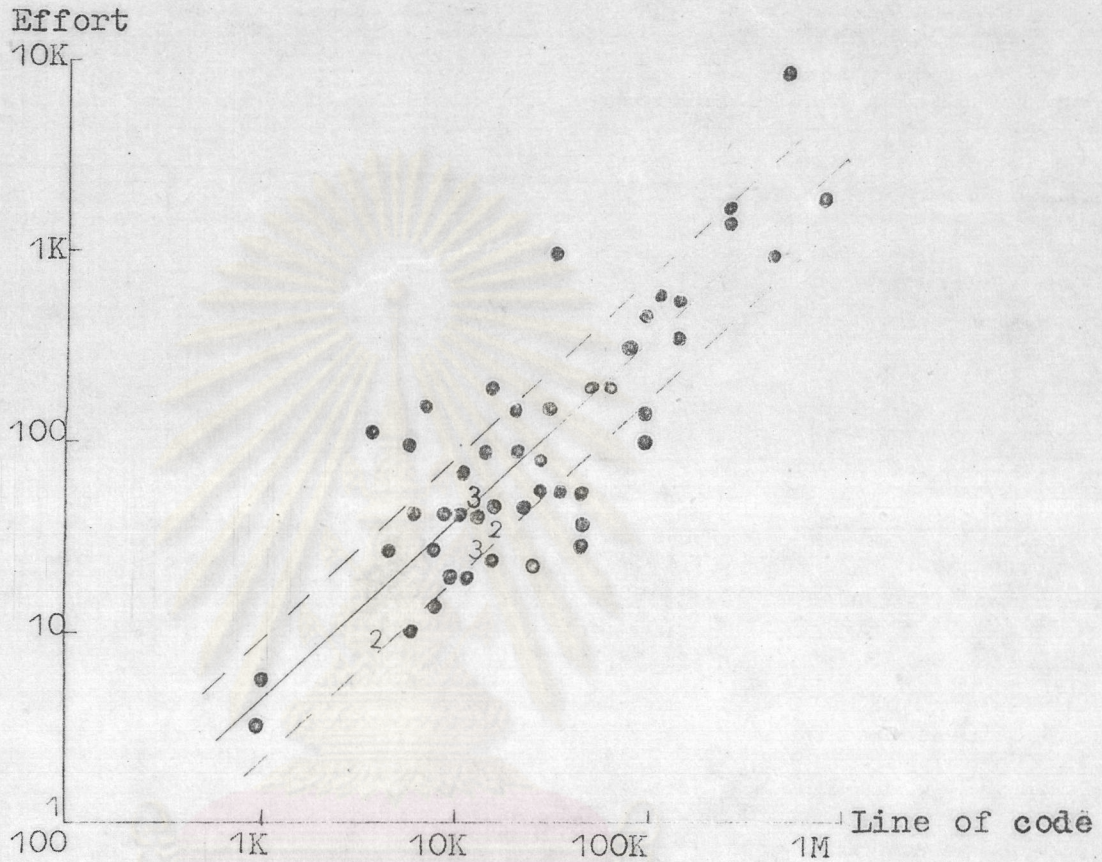
เมื่อเรานำมาหาความสัมพันธ์ระหว่างค่าใช้จ่ายทางคอมพิวเตอร์และแรงงาน-เวลา กลับปรากฏว่าค่าใช้จ่ายทางคอมพิวเตอร์มีความสัมพันธ์กับจำนวนแรงงาน-เวลาที่ใช้ไปในโครงการมากกว่า

<u>สรุป</u> ผลการวิจัยของ IBM (ตารางที่ 3.1)

Estimated Variable	Equation	SE(a)	r**2(b)
Total effort	E = 5.2 *(L ** 0.91)	2.51	0.64
Documentation	DOC = 49*(L ** 1.01)	2.68	0.62
Project Duration	D = 4.1*(L ** 0.36)	1.72	0.41
	D = 2.47*(E** 0.35)	1.52	0.79

(a) = ความเบี่ยงเบนมาตรฐานของการกะประมาณ (Standard error of estimate)

(b) = สัมประสิทธิ์ของการตัดสินใจ (Coefficient of determination)

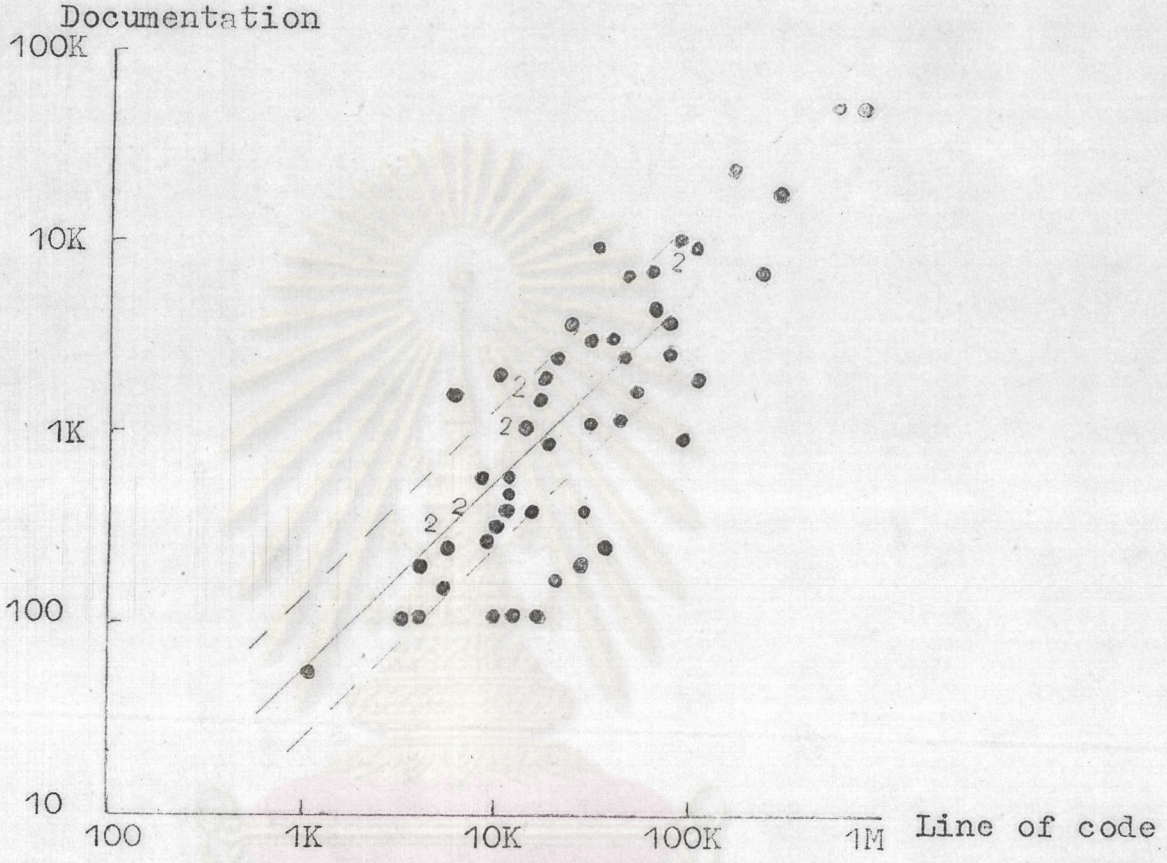


$$E = 5.2 * (L ** 0.91)$$

E = แรงงานทั้งหมดที่ใช้ไปในการพัฒนาโปรแกรม มีหน่วยเป็นผลคูณของจำนวนคน และเวลาที่ใช้ไป (man - month)

L = จำนวนบรรทัดของโปรแกรม มีหน่วยเป็นพันบรรทัดของโปรแกรม

รูปที่ 3.1.1 ความสัมพันธ์ระหว่างแรงงาน–เวลาที่ใช้ไปในการพัฒนาโปรแกรม และจำนวนบรรทัดของโปรแกรม

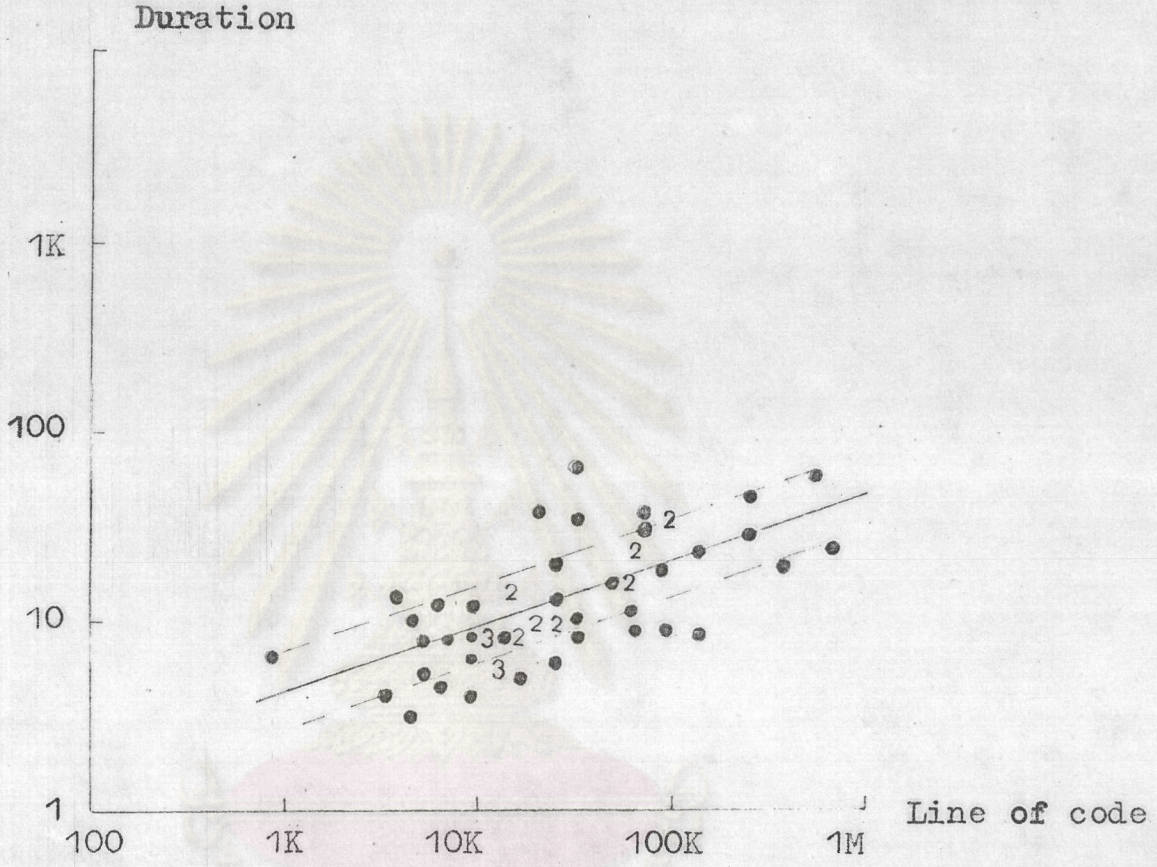


DØC = 49 * L ** 1.01

DØC = จำนวนแผนของเอกสาร

L = จำนวนบรรทัดของโปรแกรม

รูปที่ 3.1.2 ความสัมพันธ์ระหว่างจำนวนเอกสารและจำนวนบรรทัดของโปรแกรม

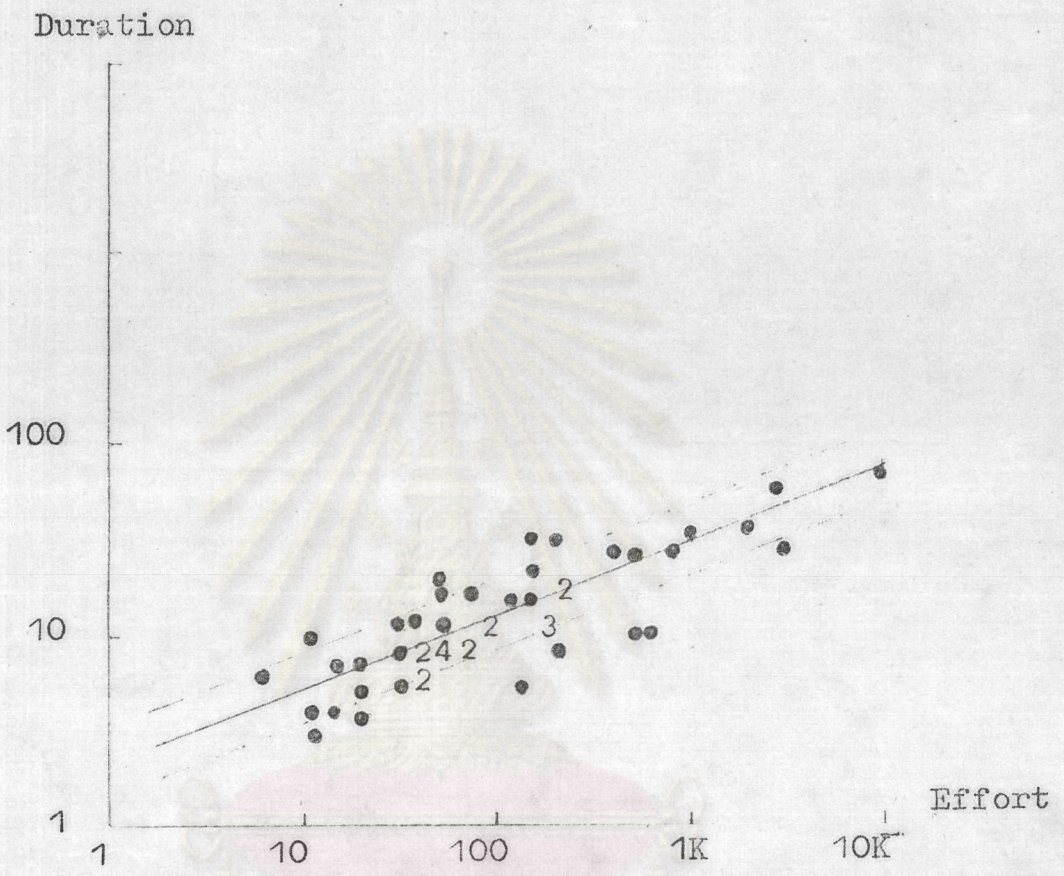


$D = 4.1 * L ** 0.36$

D = ระยะเวลาเป็นเดือน

L = 1,000 บรรทัดของโปรแกรม

รูปที่ 3.1.3 ความสัมพันธ์ระหว่างระยะเวลาของการทำโครงการและจำนวนบรรทัดของโปรแกรม

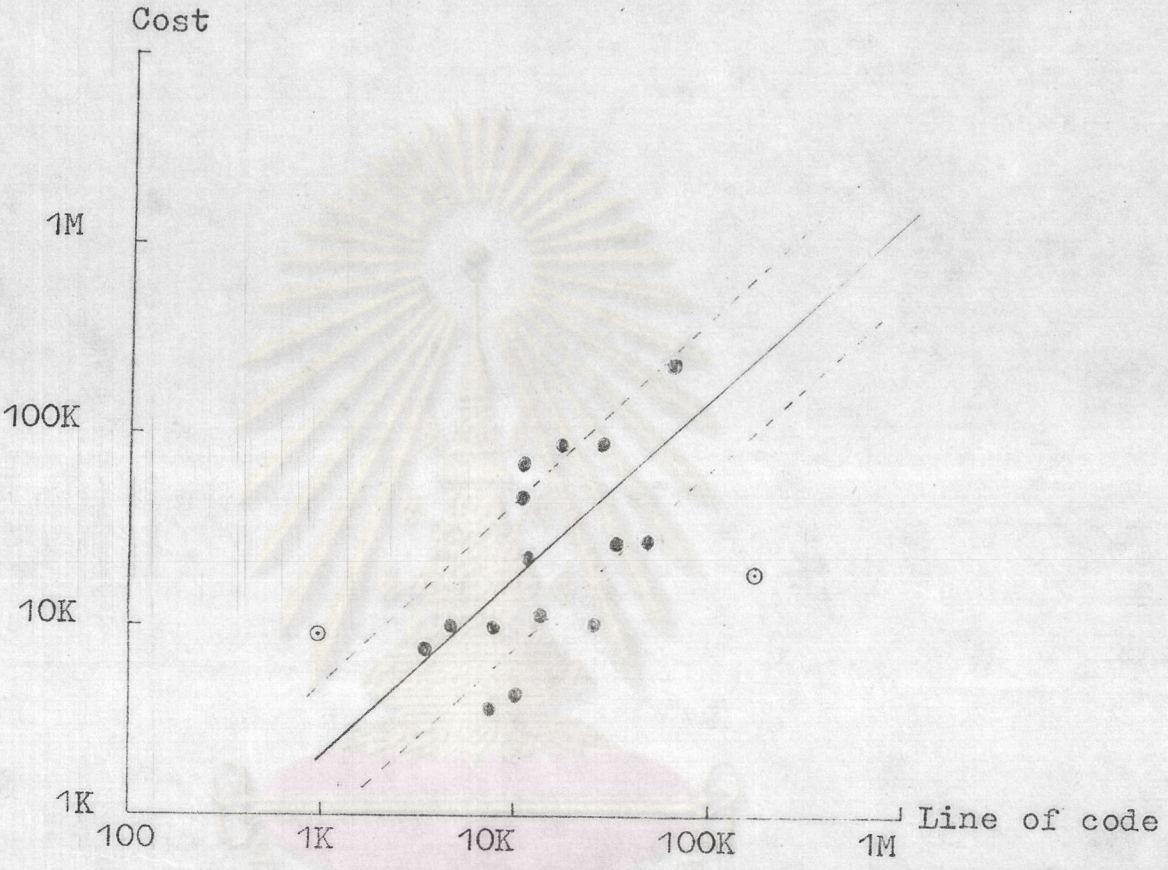


$$D = 2.47 * (E^{**}0.35)$$

D = ระยะเวลาเป็นเดือน

E = แรงงานที่ใช้ไปคิดเป็นหน่วย คน–เดือน

รูปที่ 3.1.4 ความสัมพันธ์ระหว่างระยะเวลาของโครงการและจำนวนแรงงาน–เวลาที่ใช้ไปทั้งหมด

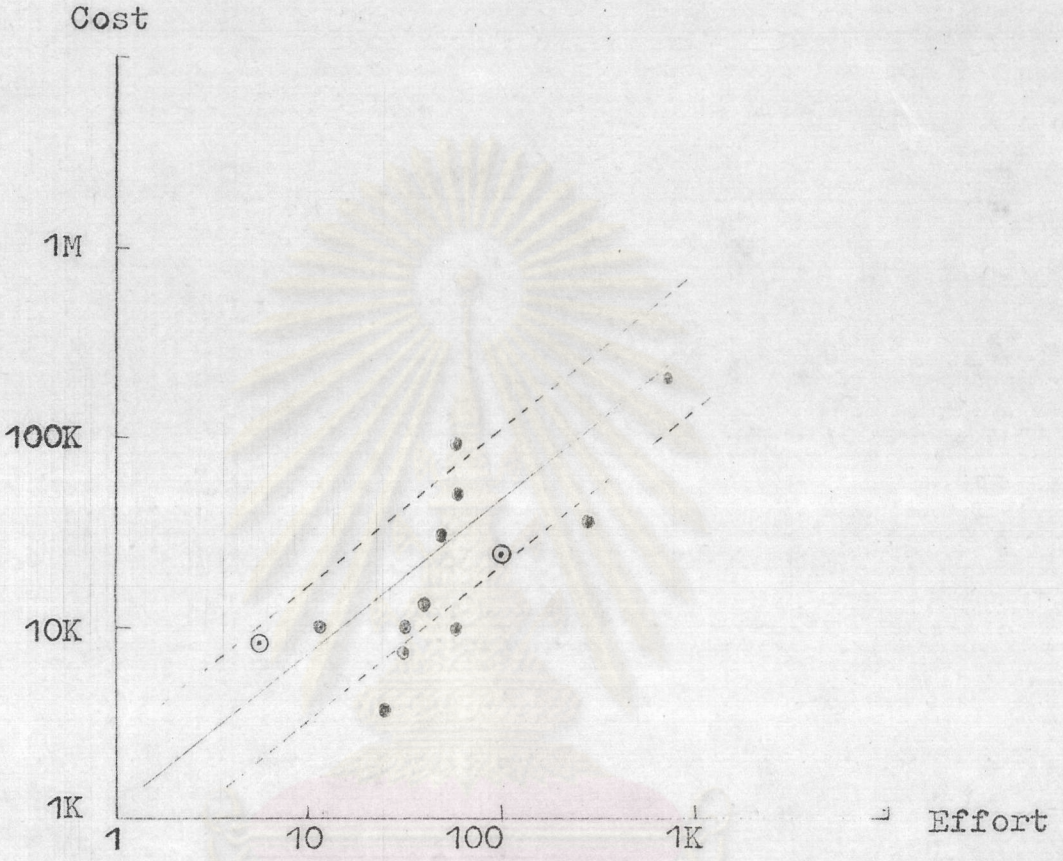


$$C = 1.84^{*} (L ^{**} 0.96)$$

C = ค่าใช้จ่ายทางคอมพิวเตอร์ หน่วย พันดอลล่าร์

L = จำนวนบรรทัดของโปรแกรม หน่วย พันบรรทัด

รูปที่ 3.1.5 ความสัมพันธ์ระหว่างค่าใช้จ่ายทางคอมพิวเตอร์และจำนวนบรรทัดของโปรแกรม



$$C = 1.1 * (E^{**} 0.81)$$

C = ค่าใช้จ่ายทางคอมพิวเตอร์ หน่วยเป็นดอลล่าร์

E = แรงงานทั้งหมด หน่วย คน-เดือน

รูปที่ 3.1.6 ความสัมพันธ์ระหว่างค่าใช้จ่ายทางคอมพิวเตอร์ และแรงงาน-เวลาทั้งหมด

3.2 ผลการวิจัยของ SEL

3.2.1 ความสัมพันธ์ระหว่างแรงงาน–เวลาและจำนวนบรรทัดของโปรแกรม

(ดังรูป 3.2.1)

สมการของความสัมพันธ์ คือ E = 1.38 * (L**0.93)

ความเบี่ยงเบนมาตรฐานของการคาดคะเน สามารถที่จะกำหนดขอบเขตในสมการข้างต้น ในที่นี้ความเบี่ยงเบนมาตรฐานมีค่า 1.41 ซึ่งจะทำให้ได้สมการของขีดจำกัดบนและขีดจำกัดล่าง ดังนี้

E = 1.95 * (L**0.93)

E = 0.98 * (L**0.93)

สัมประสิทธิ์ของการตัดสินใจ (r**2 = 0.93) มีนัยสำคัญสูง ซึ่งแสดงให้เห็นว่าสำหรับโครงการนี้ความสัมพันธ์ระหว่างแรงงานและจำนวนบรรทัดของโปรแกรมค่อนข้างจะเป็นเส้นกรง

ความสัมพันธ์ระหว่างแรงงาน–เวลา และจำนวนบรรทัดของโปรแกรมที่พัฒนาขึ้นใหม่กับจำนวนบรรทัดของโปรแกรมที่ได้รับการพัฒนา

E = 1.58 * (NL**0.99)

แทนค่าจำนวนบรรทัดของโปรแกรมที่ได้รับการพัฒนาขึ้นใหม่ (NL) ด้วยจำนวนบรรทัดของโปรแกรมที่ได้รับการพัฒนา (DL) ลงในสมการข้างต้น จะได้

E = 1.48* (DL**0.98)

ความสัมพันธ์ระหว่างแรงงาน–เวลา และจำนวนบรรทัดของโปรแกรมที่พัฒนาขึ้นใหม่ (NL) และจำนวนบรรทัดของโปรแกรมที่ได้รับการพัฒนา (DL) มีสัมประสิทธิ์ของการตัดสินใจสูง แสดงให้เห็นว่าเราสามารถที่จะใช้พยากรณ์แรงงาน–เวลาทั้งหมด ถ้าหากว่าเราทราบจำนวนบรรทัดของจำนวนโปรแกรมล่วงหน้า

3.2.2 ความสัมพันธ์ระหว่างแรงงาน–เวลา และหน่วยย่อยของโปรแกรม (ดังรูป 3.2.2)

ความสัมพันธ์ที่ได้จากข้อมูลที่รวบรวม $E = 0.65 * (M**1.19)$

ความสัมพันธ์ไม่เป็นเส้นตรงแต่สัมประสิทธิ์ของการตัดสินใจชี้ให้เห็นว่า มีสหสัมพันธ์สูงระหว่างตัวแปรผันทั้งสองนี้

ความสัมพันธ์ระหว่างแรงงาน–เวลา และหน่วยย่อยของโปรแกรมที่ทำขึ้นใหม่ คือ

$$E = 0.183 * (NM**1.05)$$

หน่วยย่อยของโปรแกรมที่ทำขึ้นใหม่ คือ หน่วยย่อยของโปรแกรมที่ทำขึ้นใหม่ทั้งหมดอย่างสมบูรณ์ หรือเป็นหน่วยย่อยของโปรแกรมเก่าซึ่งได้รับการพัฒนาแก้ไขมากกว่าร้อยละ 20

จากความสัมพันธ์ต่าง ๆ ข้างต้น ระหว่างแรงงานกับขนาดของงานปรากฏว่าเป็นความสัมพันธ์ในรูปแบบเส้นตรงโดยไม่ขึ้นกับขนาดของโครงการ ซึ่งหมายความว่าผลผลิตยังคงที่ขณะที่ขนาดของโครงการเปลี่ยนแปลง

3.2.3 ความสัมพันธ์ระหว่างจำนวนบรรทัดของโปรแกรม (L) และจำนวนเอกสาร (ดังรูป 3.2.3)

เอกสารเป็นส่วนสำคัญของโครงการทางด้านซอฟท์แวร์ และค่าใช้จ่ายของการทำเอกสารเป็นองค์ประกอบหนึ่งของการคาดคะเนขบวนการทางด้านซอฟท์แวร์ เอกสารต่าง ๆ เหล่านี้ ได้แก่การออกแบบโปรแกรม การวางแผนเพื่อการทดสอบโปรแกรม คู่มือการใช้ ฯลฯ ความสัมพันธ์ของสองตัวแปรนี้คือ

$$DOC = 30.4 * (L**0.90)$$

นอกจากนี้ยังมีความสัมพันธ์ระหว่างจำนวนเอกสารกับจำนวนบรรทัดของโปรแกรมที่พัฒนาขึ้นใหม่ ซึ่งมีสมการ ดังนี้

$$DOC = 38.1 * (NL ** 0.93)$$

และความสัมพันธ์ระหว่างจำนวนเอกสารและจำนวนบรรทัดของโปรแกรมที่ได้รับการพัฒนา (DL) ดังมีสมการ ดังนี้

$$DOC = 34.7 * (DL ** 0.93)$$

3.2.4 <u>ความสัมพันธ์ระหว่างจำนวนเอกสารและจำนวนหน่วยย่อยของโปรแกรม</u> (ดังรูป 3.2.4)

สมการของความสัมพันธ์ คือ $DOC = 1.54* (M** 1.12)$

นอกจากนี้ SEL ยังได้ทำการหาความสัมพันธ์ระหว่างจำนวนเอกสารและจำนวนหน่วยย่อยของโปรแกรมเฉพาะที่ผลิตขึ้นใหม่ (NM)

$$DOC = 4.82* (NM** 0.99)$$

3.2.5 <u>ความสัมพันธ์ระหว่างระยะเวลาเวลาและจำนวนบรรทัดของโปรแกรม</u> (ดังรูป 3.2.5)

สมการของความสัมพันธ์ คือ $D = 4.55* (L** 0.26)$

นอกจากนี้ SEL ยังได้ทำการวิเคราะห์ความสัมพันธ์ระหว่างระยะเวลากับจำนวนบรรทัดใหม่และจำนวนบรรทัดที่ได้รับการพัฒนาแล้ว ซึ่งมีสมการของความสัมพันธ์ดังนี้

$$D = 4.62* (NL** 0.28)$$

$$D = 4.58* (DL** 0.28)$$

3.2.6 <u>ความสัมพันธ์ระหว่างระยะเวลาและจำนวนหน่วยย่อยของโปรแกรม</u> (ดังรูป 3.2.6)

สมการของความสัมพันธ์ คือ $D = 1.96* (M** 0.33)$

และสมการความสัมพันธ์ระหว่างระยะเวลาของโปรแกรมและจำนวนหน่วยย่อยของโปรแกรมที่พัฒนาขึ้นใหม่

$$D = 2.5* (NM** 0.30)$$

3.2.7 <u>ความสัมพันธ์ระหว่างระยะเวลาของการทำโปรแกรมและแรงงานที่ใช้ไปในการพัฒนาโปรแกรม</u> (ดังรูป 3.2.7)

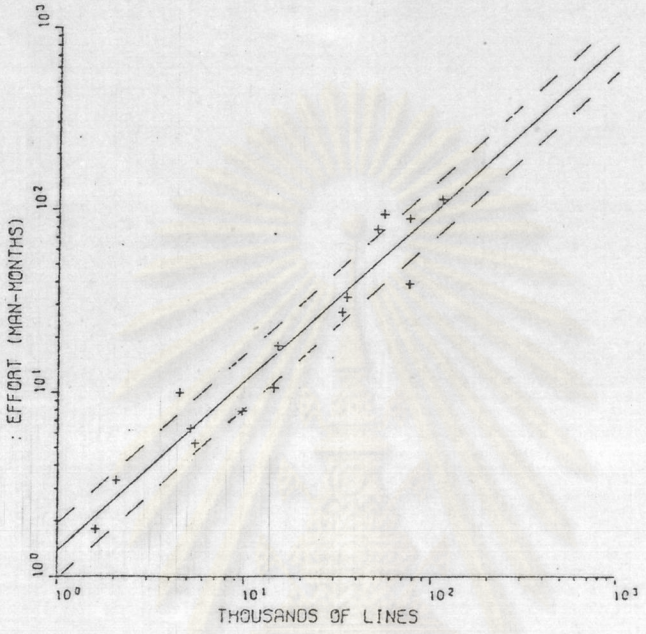
สมการของความสัมพันธ์ คือ $D = 4.39* (E** 0.26)$

3.2.8 สรุปผลการวิจัยของ SEL (ตารางที่ 3.2)

Estimated variable	SEL equation	SE	r**2
Total effort	E = 1.38* (L**0.93)	1.41	0.93
	E = 1.58* (NL**0.99)	1.31	0.96
	E = 1.48* (DL**0.98)	1.29	0.96
	E = 0.652* (M**1.19)	1.49	0.90
	E = 0.183* (NM**1.05)	1.57	0.87
Documentation	DOC = 30.4* (L**0.90)	1.41	0.92
	DOC = 38.1* (NL**0.93)	1.52	0.885
	DOC = 34.7* (DL**0.93)	1.45	0.91
	DOC = 1.54* (M**1.16)	1.45	0.91
	DOC = 4.82* (NM**0.99)	1.67	0.83
Duration	D = 4.55* (L**0.26)	1.36	0.55
	D = 1.96* (M** 0.33)	1.37	0.54
	D = 4.62* (NL**0.28)	1.33	0.61
	D = 4.58* (DL**0.28)	1.34	0.59
	D = 2.5* (NM**0.30)	1.38	0.55
	D = 4.39* (E**0.26)	1.37	0.52

ศูนย์วิทยทรัพยากร
จุฬาลงกรณ์มหาวิทยาลัย

Effort



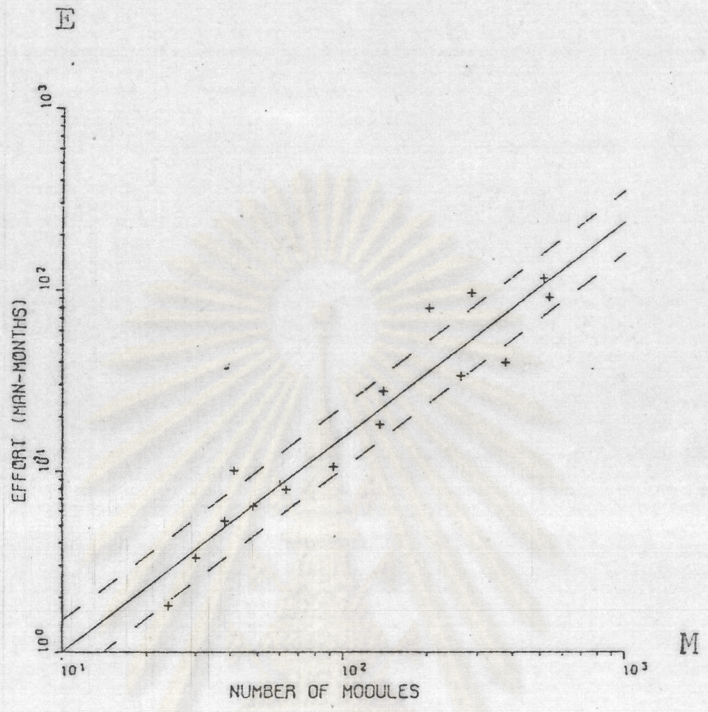
Line of code

$$E = 1.38 * (L**0.93)$$

E = แรงงาน–เวลา มีหน่วยเป็น คน–เดือน

L = จำนวนบรรทัดของโปรแกรม หน่วยพันบรรทัด

รูป 3.2.1 ความสัมพันธ์ระหว่างแรงงาน–เวลาและจำนวนบรรทัดของโปรแกรม



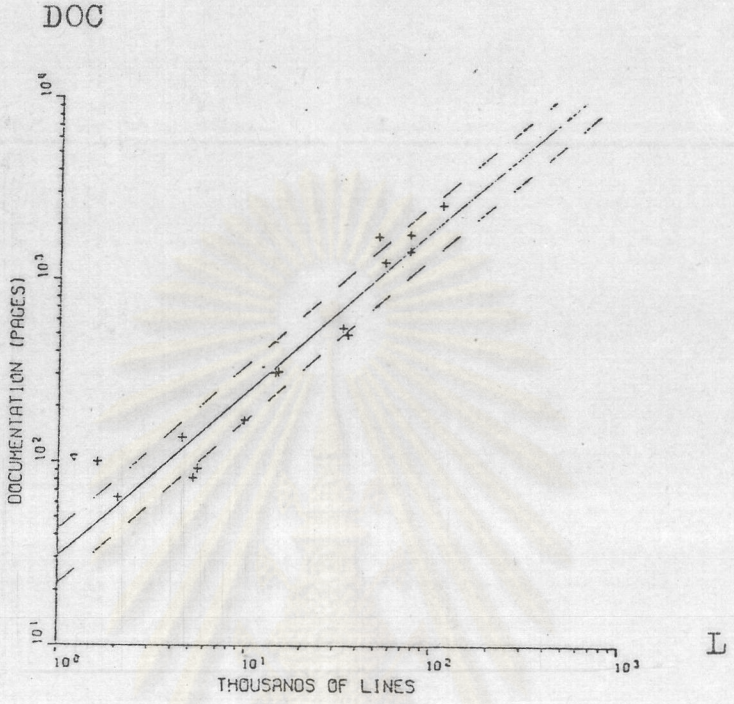
$$E = 0.65 * (M^{**}1.19)$$

E = แรงงาน–เวลา หน่วย คน–เดือน

M = จำนวนหน่วยย่อยของโปรแกรม

รูป 3.2.2 ความสัมพันธ์ระหว่างแรงงาน–เวลาและจำนวนหน่วยย่อยของโปรแกรม

ศูนย์วิทยทรัพยากร
จุฬาลงกรณ์มหาวิทยาลัย

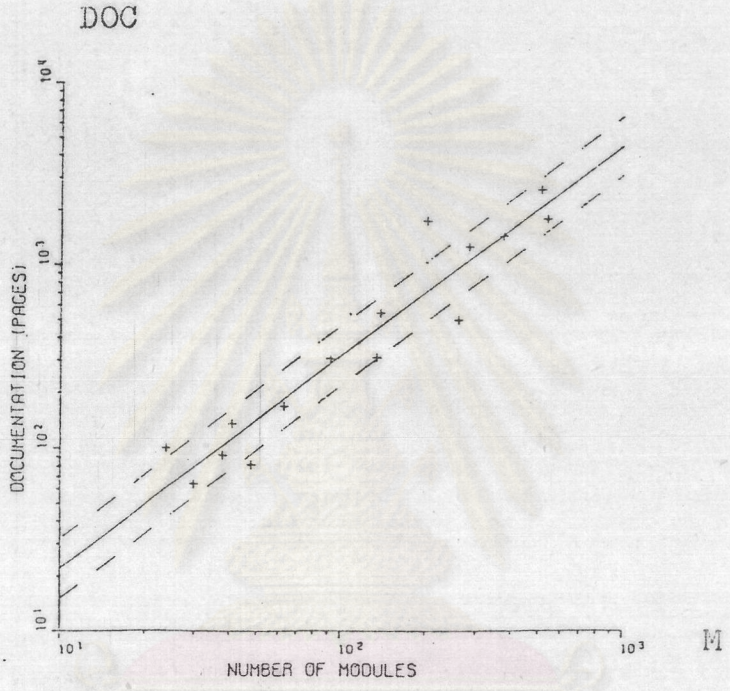


$$DOC = 30.4 * (L^{**} 0.90)$$

DOC = จำนวนเอกสาร

L = จำนวนบรรทัดของโปรแกรม หน่วยเป็นบรรทัด

รูป 3.2.3 ความสัมพันธ์ระหว่างจำนวนบรรทัดของโปรแกรม (L) และจำนวนเอกสาร

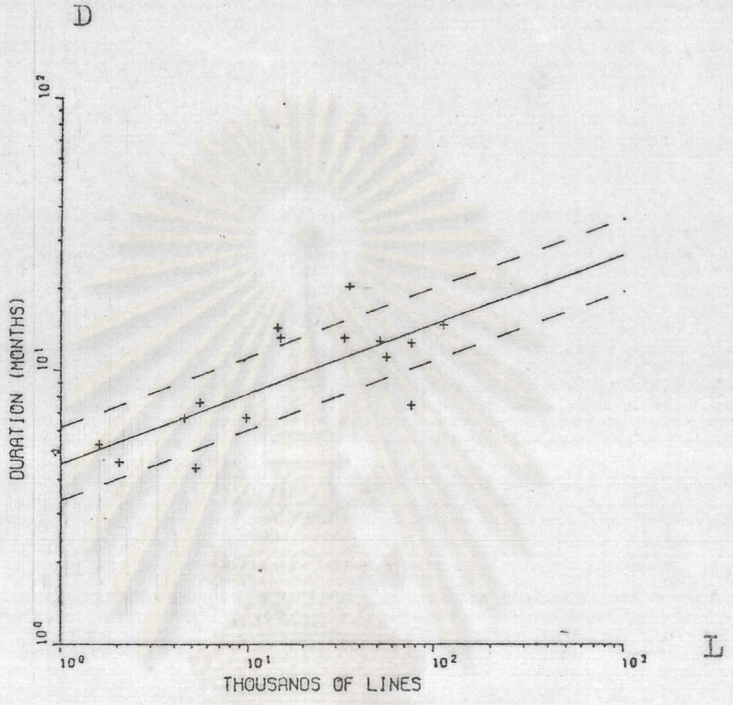


$$DOC = 1.54^{*} (M^{**}1.12)$$

DOC = จำนวนเอกสารซึ่งนับเป็นแผ่น

M = จำนวนของหน่วยย่อยของโปรแกรม

รูปที่ 3.2.4 ความสัมพันธ์ระหว่างจำนวนเอกสารและจำนวนหน่วยย่อยของโปรแกรม

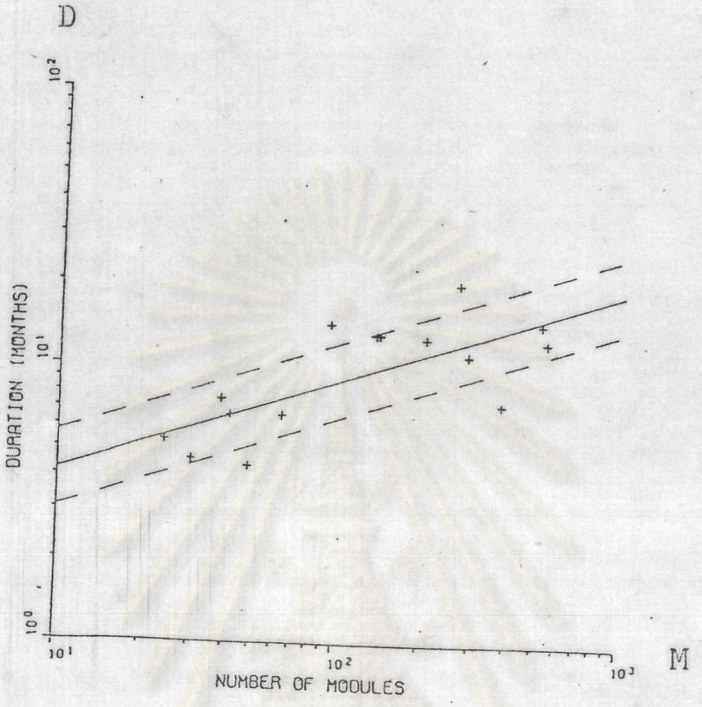


$$D = 4.55^{*} (L^{**}0.26)$$

D = ระยะเวลาของการทำโปรแกรม มีหน่วยเป็นเดือน

L = มีหน่วยเป็นจำนวนพันบรรทัดของโปรแกรม

รูปที่ 3.2.5 ความสัมพันธ์ระหว่างระยะเวลาและจำนวนบรรทัดของโปรแกรม

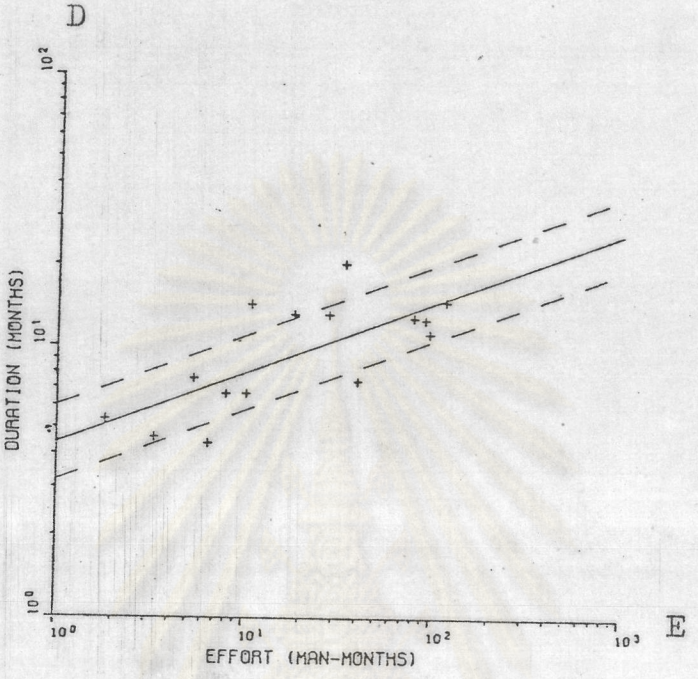


$$D = 1.96^{*} (M^{**}0.33)$$

D = เป็นระยะเวลาของการทำโปรแกรมมีหน่วยเป็นเดือน

M = เป็นจำนวนหน่วยย่อยของโปรแกรม

รูปที่ 3.2.6 ความสัมพันธ์ระหว่างเวลาและจำนวนหน่วยย่อยของโปรแกรม



$$D = 4.39^{*} (E^{**}0.26)$$

D = ระยะเวลาของการทำโปรแกรม

E = จำนวนแรงงานที่ใช้ไปในการพัฒนาโปรแกรม

ศูนย์วิทยทรัพยากร
จุฬาลงกรณ์มหาวิทยาลัย

รูปที่ 3.2.7 ความสัมพันธ์ระหว่างระยะเวลาของการทำโปรแกรมและแรงงานที่ใช้ไปในการพัฒนาโปรแกรม

3.3 เปรียบเทียบผลการวิจัยของ IBM และ SEL

ก่อนที่จะทำการเปรียบเทียบผลการวิจัยของ IBM และ SEL เราจะมาดูสิ่งแวดล้อมของทั้ง 2 แห่งก่อน เพื่อประกอบการพิจารณา

สิ่งแวดล้อม	IBM	SEL
เครื่องคอมพิวเตอร์	ใช้เครื่องคอมพิวเตอร์ขนาดต่าง ๆ กันทั้งเล็กและใหญ่ จำนวน 66 เครื่องในการพัฒนาโปรแกรม	ใช้เครื่องคอมพิวเตอร์ของ IBM-360 และ PDP = 11/70
ภาษาคอมพิวเตอร์ที่ใช้	ใช้ภาษาต่าง ๆ 28 ภาษาในการพัฒนาโปรแกรมทั้งหมด 60 โครงการ	ใช้ภาษาฟอร์แทรนเกือบทั้งหมด มี 5 โครงการที่ใช้ฟอร์แทรนผสมกับแอสเซมบลี
ลักษณะของงานที่ทำ	มีทั้งหมด 60 โครงการ ซึ่งมีขนาดตั้งแต่ 4,000–467,000 บรรทัดของโปรแกรม งานที่ทำมีตั้งแต่โปรแกรมอย่างง่ายจนถึงระบบซึ่งมีความสลับซับซ้อนมาก	มีทั้งหมด 15 โครงการ ขนาดตั้งแต่ 1,600 ถึง 112,000 บรรทัดของโปรแกรม

จะเห็นได้ว่าสิ่งแวดล้อมของการทำโปรแกรมของ IBM มีความแตกต่างกันกว้างมาก เมื่อเทียบกับของ SEL ซึ่งดูเกือบจะคงที่ สิ่งต่าง ๆ เหล่านี้ทำให้ความเบี่ยงเบนมาตรฐานของ IBM มีความแตกต่างกันมากถึงประมาณ 6.3 เท่าจากขีดจำกัดบนถึงขีดจำกัดล่าง ส่วนของ SEL แตกต่างกัน 2 เท่าเมื่อเทียบจากความสัมพันธ์เดียวกัน ในด้านของสมการความสัมพันธ์ แม้ว่าสัมประสิทธิ์ของสมการจะไม่เหมือนกัน แต่ด้วยกำลังของ เอกซ์โปเนนท์เชียลมีค่าใกล้เคียงกัน จึงแสดงให้เห็นว่า แม้ว่าวิธีการวัดผลด้วยวิธีนี้ แม้จะไม่ใช่วิธีที่ดีที่สุดก็ตาม แต่ก็ใช้ได้สำหรับ

หน่วยงานหรือองค์การต่าง ๆ ที่มีสิ่งแวดล้อมของการทำโปรแกรมที่แตกต่างกัน

ผลจากการวิจัยของ SEL และ IBM มีดังต่อไปนี้

1. จากการที่ SEL ได้เพิ่มการวัดข้อมูลของจำนวนบรรทัดของโปรแกรมที่พัฒนาขึ้นใหม่ (NL) และจำนวนบรรทัดของโปรแกรมที่พัฒนาแล้ว (DL) มาใช้ปรากฏว่าข้อมูล 2 ตัว นี้มีความสัมพันธ์กับแรงงานที่ใช้ไปในการผลิตโปรแกรมดีกว่าจำนวนบรรทัดของโปรแกรมทั้งหมด (L)

2. แม้ว่าการวัดผลผลิตที่ทำโดยวิธีนี้จะค่อนข้างหยาบ แต่เราก็เชื่อได้ว่าการนำเอาโปรแกรมเดิมซึ่งมีส่วนของโปรแกรมคล้ายกับงานใหม่ที่เรากำลังพัฒนาอยู่มาดัดแปลงแก้ไข จะมีส่วนช่วยในการลดค่าใช้จ่ายของการทำโปรแกรมลงได้

3. การใช้หน่วยย่อยของโปรแกรมมาช่วยในการพยากรณ์จำนวนแรงงานที่ต้องใช้ในการพัฒนาโปรแกรมจะเป็นสิ่งที่สามารถทำได้ง่ายกว่าการพยากรณ์โดยคาดคะเนจำนวนบรรทัดของโปรแกรม

4. ในการทำโปรแกรมขนาดใหญ่ เวลาปฏิทิน (Calendar time) เป็นข้อจำกัดหลักอันหนึ่งของการผลิต โดยเวลาจะเพิ่มขึ้นประมาณ 1 ใน 3 เท่าของขนาดของโปรแกรม ทั้งนี้เนื่องจากความสัมพันธ์ของระยะเวลาของการทำโปรแกรมกับจำนวนบรรทัดของโปรแกรมและจำนวนหน่วยย่อยของโปรแกรมมีค่ายกกำลังของเอกซ์โปเนนท์เฉลี่ยมีค่าระหว่าง 0.26 – 0.33 (ประมาณ $^1/_3$)

ศูนย์วิทยทรัพยากร
จุฬาลงกรณ์มหาวิทยาลัย