บทที่ 4

ผลการวิจัย

ผลการวิจัยครั้งนี้เป็นผลการวิจัยการศึกษาเปรียบเทียบวิธีพยากรณ์ คือ เทคนิคการทำให้เรียบแบบเอกซโพเนนเชียล เทคนิคการทำให้เรียบแบบเอกซโพเนนเชียลเมื่อประมาณค่าด้วยวิธีค่าสัมบูรณ์ต่ำสุด และเทคนิคการทำให้เรียบแบบเอกซโพเนนเชียลเมื่อประมาณค่าด้วยวิธีการของฮูเบอร์ ซึ่งสองเทคนิคหลังเป็นวิธีการที่ปรับแก้มาจากเทคนิคการทำให้เรียบแบบเอกซโพเนนเชียล จากวิธีพยากรณ์ดังกล่าวเมื่อนำมาใช้ในการวิเคราะห์ข้อมูลอนุกรมเวลาที่มีข้อมูลผิดปกติ โดยทำการศึกษาค่าความคลาดเคลื่อนที่เกิดขึ้นในการพยากรณ์ล่วงหน้า 12 คาบเวลา เมื่อคำนึงถึงการแจกแจงของความคลาดเคลื่อนในตัวแบบของข้อมูลอนุกรมเวลา ขนาดตัวอย่าง เปอร์เซ็นต์ของการปลอมปน และวิธีการพยากรณ์แบบต่าง ๆ ที่ใช้ในการวิจัย เพื่อหาข้อสรุปว่า วิธีการพยากรณ์ใดเหมาะสมที่จะนำไปใช้ในการพยากรณ์ในแต่ละสถานการณ์ที่เกิดขึ้นในการทดลอง ซึ่งค่าของความคลาดเคลื่อนของการศึกษาที่ปรากฏในบทนี้ จะนำเสนอผลในรูปของตาราง และกราฟของรากที่สองของความคลาดเคลื่อนกำลังสองระหว่างค่าจริงและค่าพยากรณ์เฉลี่ย 12 คาบเวลาพยากรณ์ สำหรับตารางและกราฟของค่า RMSE ของ 12 คาบเวลาพยากรณ์จะแสดงไว้ในภาคผนวก ข โดยสัญลักษณ์ที่ใช้ในการแทนความหมายต่าง ๆ ที่จะปรากฏในตารางและกราฟมีดังนี้

SINGLE หมายถึง วิธีพยากรณ์โดยใช้การทำให้เรียบแบบเอกซโพเนนเชียลครั้งเดียว

S-LAV หมายถึง วิธีพยากรณ์โดยใช้การทำให้เรียบแบบเอกซโพเนนเชียลครั้งเดียว เมื่อประมาณค่าด้วยวิธีค่าสัมบูรณ์ต่ำสุด

S-HUBER หมายถึง วิธีพยากรณ์โดยใช้การทำให้เรียบแบบเอกซโพเนนเชียลครั้งเดียว เมื่อประมาณค่าด้วยวิธีการของฮูเบอร์

DOUBLE หมายถึง วิธีพยากรณ์โดยใช้การทำให้เรียบแบบเอกซโพเนนเชียลซ้ำสองครั้ง

D-LAV หมายถึง วิธีพยากรณ์โดยใช้การทำให้เรียบแบบเอกซโพเนนเชียลซ้ำสองครั้ง เมื่อประมาณค่าด้วยวิธีค่าสัมบูรณ์ต่ำสุด

D-HUBER หมายถึง วิธีพยากรณ์โดยใช้การทำให้เรียบแบบเอกซโพเนนเชียลซ้ำสองครั้ง เมื่อประมาณค่าด้วยวิธีการของฮูเบอร์

n หมายถึง ขนาดตัวอย่าง

p หมายถึง เปอร์เซ็นต์การปลอมปน

N(0,1) หมายถึง การแจกแจงปกติมาตรฐาน

SCN(0,c*c) หมายถึง การแจกแจงแบบสเกลคอนทามิเนตนอร์มอล โดยมีค่าสเกลแฟกเตอร์เท่ากับ c (scale factor)

CNL(0,β) หมายถึง การแจกแจงปลอมปนระหว่างการแจกแจงปกติมาตรฐาน N(0,1) กับการแจกแจงลาปลาซที่มีค่าเฉลี่ยเท่ากับ 0 และความแปรปรวนเท่ากับ $2\beta^2$

CNU(a,b) หมายถึง การแจกแจงปลอมปนระหว่างการแจกแจงปกติมาตรฐาน N(0,1) กับการแจกแจงสม่ำเสมอ ที่มีค่าเฉลี่ยเท่ากับ (a+b)/2 และความแปรปรวนเท่ากับ $(b-a)^2/12$

ผลการศึกษาเปรียบเทียบวิธีพยากรณ์ทั้ง 3 วิธี

การเปรียบเทียบความสามารถในการพยากรณ์ทั้ง 3 วิธีในการวิเคราะห์อนุกรมเวลาที่มีข้อมูลผิดปกติ เกณฑ์ที่ใช้ในการพิจารณาเปรียบเทียบคือ ค่าความคลาดเคลื่อนระหว่างค่าพยากรณ์กับค่าจริง ในรูปของรากที่สองของค่าเฉลี่ยของความคลาดเคลื่อนกำลังสอง (RMSE) เฉลี่ยของ 12 คาบเวลาพยากรณ์ วิธีพยากรณ์ที่ค่า RMSE เฉลี่ย 12 คาบเวลาพยากรณ์ต่ำกว่าจะเป็นวิธีพยากรณ์ที่ใช้ได้ดีกว่า สำหรับค่า RMSE ของ 12 คาบเวลาพยากรณ์จะแสดงไว้ในภาคผนวก ข

เนื่องจากข้อมูลอนุกรมเวลาที่ใช้ในการศึกษาครั้งนี้มี 2 รูปแบบ คือ ข้อมูลอนุกรมเวลาที่มีลักษณะเคลื่อนไหวอยู่ในระดับค่าเฉลี่ยและมีลักษณะแนวโน้มเชิงเส้น ดังนั้นจึงขอเสนอผลการเปรียบเทียบวิธีพยากรณ์ทั้ง 3 วิธีแยกเป็น 2 กรณี ดังต่อไปนี้

1. กรณีข้อมูลที่มีลักษณะเคลื่อนไหวอยู่ในระดับค่าเฉลี่ย

การวิเคราะห์ผลสำหรับข้อมูลที่มีลักษณะเคลื่อนไหวอยู่ในระดับค่าเฉลี่ย เมื่อความคลาดเคลื่อนมีการแจกแจงแบบต่าง ๆ สามารถแสดงได้ดังนี้

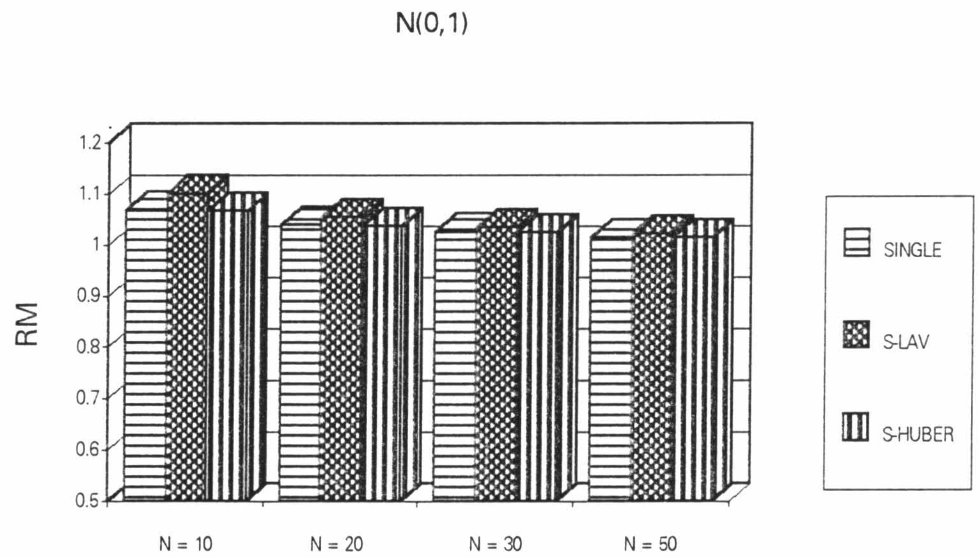
1.1 เมื่อความคลาดเคลื่อนมีการแจกแจงปกติมาตรฐาน

การพยากรณ์สำหรับข้อมูลที่มีลักษณะเคลื่อนไหวอยู่ในระดับค่าเฉลี่ย ความคลาดเคลื่อนมีการแจกแจงปกติมาตรฐาน โดยขนาดตัวอย่างที่ใช้ในการศึกษาคือ 10, 20, 30 และ 50 ทำการพยากรณ์ด้วยวิธีการทำให้เรียบแบบเอกซโพเนนเชียลครั้งเดียว (SINGLE) การทำให้เรียบแบบเอกซโพเนนเชียลครั้งเดียวเมื่อประมาณค่าด้วยวิธีค่าสัมบูรณ์ต่ำสุด (S--LAV) และการทำให้เรียบแบบเอกซโพเนนเชียลครั้งเดียวเมื่อประมาณค่าด้วยวิธีการของฮูเบอร์ (S-HUBER) ทำการคำนวณค่ารากที่สองของค่าเฉลี่ยของความคลาดเคลื่อนกำลังสอง (RMSE) เฉลี่ย 12 คาบเวลา ได้ดังตารางที่ 4.1 และรูปที่ 4.1 และค่า RMSE ของ 12 คาบเวลา ได้ในภาคผนวก ข ดังตารางที่ 4.17 และรูปที่ 4.23

ตารางที่ 4.1 แสดงค่า RMSE เฉลี่ย 12 คาบเวลาของวิธีพยากรณ์ทั้ง 3 วิธี ในการวิเคราะห์ข้อมูลอนุกรมเวลาที่มีลักษณะเคลื่อนไหวอยู่ในระดับค่าเฉลี่ย เมื่อความคลาดเคลื่อนมีการแจกแจงปกติมาตรฐาน

n	วิธีพยากรณ์		
	SINGLE	S-LAV	S-HUBER
10	1.0677	1.1009	1.0670
20	1.0387	1.0539	1.0364
30	1.0245	1.0339	1.0236
50	1.0141	1.0214	1.0138

รูปที่ 4.1 แสดงค่า RMSE เฉลี่ย 12 คาบเวลาของวิธีพยากรณ์ทั้ง 3 วิธี ในการวิเคราะห์ข้อมูลอนุกรมเวลาที่มีลักษณะเคลื่อนไหวอยู่ในระดับค่าเฉลี่ย เมื่อความคลาดเคลื่อนมีการแจกแจงปกติมาตรฐาน



จากตารางที่ 4.1 และตารางที่ 4.17 ได้ผลดังนี้

ในการวิเคราะห์อนุกรมเวลาที่มีลักษณะเคลื่อนไหวอยู่ในระดับค่าเฉลี่ย เมื่อความคลาดเคลื่อนมีการแจกแจงปกติมาตรฐานซึ่งไม่มีค่าผิดปกติปรากฏในข้อมูลอนุกรมเวลา เปรียบเทียบค่า RMSE พบว่า

วิธี S-HUBER ให้ค่า RMSE โดยเฉลี่ยต่ำสุดในทุกขนาดตัวอย่าง ซึ่งใกล้เคียงกับวิธี SINGLE ซึ่งจากการพิจารณา 12 คาบเวลาพยากรณ์ วิธี S-HUBER ให้ค่า RMSE ต่ำสุดในทุกคาบเวลา เมื่อขนาดตัวอย่างเท่ากับ 10, 20 และ 30 และเกือบทุกคาบเวลา เมื่อขนาดตัวอย่างเท่ากับ 50 ยกเว้นคาบเวลาที่ 8, 9 และ 12 วิธี SINGLE ให้ค่า RMSE ต่ำสุด

1.2 เมื่อความคลาดเคลื่อนมีฟังก์ชันการแจกแจง $f(x) = (1-p)N(0,1) + pN(0,c^2\sigma^2)$ เมื่อสเกลแฟกเตอร์ (c) เท่ากับ 3 และ σ เท่ากับ 1

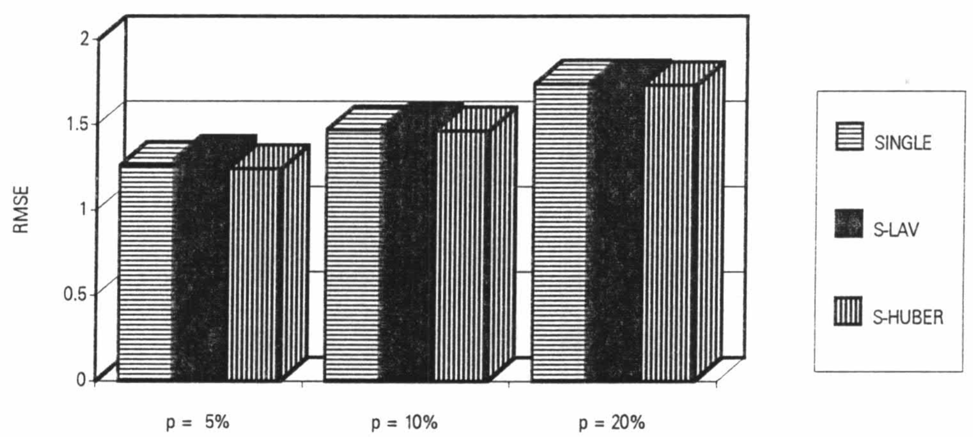
การพยากรณ์สำหรับข้อมูลที่มีลักษณะเคลื่อนไหวอยู่ในระดับค่าเฉลี่ย ความคลาดเคลื่อนมีการแจกแจงปลอมปนที่มีฟังก์ชันการแจกแจง $f(x) = (1-p)N(0,1) + pN(0,c^2\sigma^2)$ เมื่อสเกลแฟกเตอร์ (c) เท่ากับ 3 และ σ เท่ากับ 1 เปอร์เซ็นต์การปลอม (p) เท่ากับ 5, 10 และ 20 ขนาดตัวอย่างที่ใช้ในการศึกษาคือ 10, 20, 30 และ 50 ทำการพยากรณ์ด้วยวิธีการทำให้เรียบแบบเอกซโพเนนเชียลครั้งเดียว (SINGLE) การทำให้เรียบแบบเอกซโพเนนเชียลครั้งเดียวเมื่อประมาณค่าด้วยวิธีค่าสัมบูรณ์ต่ำสุด (S--LAV) และการทำให้เรียบแบบเอกซโพเนนเชียลครั้งเดียวเมื่อประมาณค่าด้วยวิธีการของฮูเบอร์ (S-HUBER) ทำการคำนวณค่ารากที่สองของค่าเฉลี่ยของความคลาดเคลื่อนกำลังสอง (RMSE) เฉลี่ย 12 คาบเวลา ได้ดังตารางที่ 4.2 และรูปที่ 4.2 และค่า RMSE ของ 12 คาบเวลา ได้ในภาคผนวก ข ดังตารางที่ 4.18 - 4.21 และรูปที่ 4.24 - 4.27

ตารางที่ 4.2 แสดงค่า RMSE เฉลี่ย 12 คาบเวลาของวิธีพยากรณ์ทั้ง 3 วิธี ในการวิเคราะห์ข้อมูลอนุกรมเวลาที่มีลักษณะเคลื่อนไหวอยู่ในระดับค่าเฉลี่ย เมื่อความคลาดเคลื่อนมีฟังก์ชันการแจกแจงอยู่ในรูป $f(x) = (1-p)\ N(0,1) + p\ N(0,c^2\sigma^2)$ เมื่อ $c = 3$ และ $\sigma = 1$ จำแนกตามขนาดตัวอย่าง (n) และเปอร์เซ็นต์การปลอมปน (p)

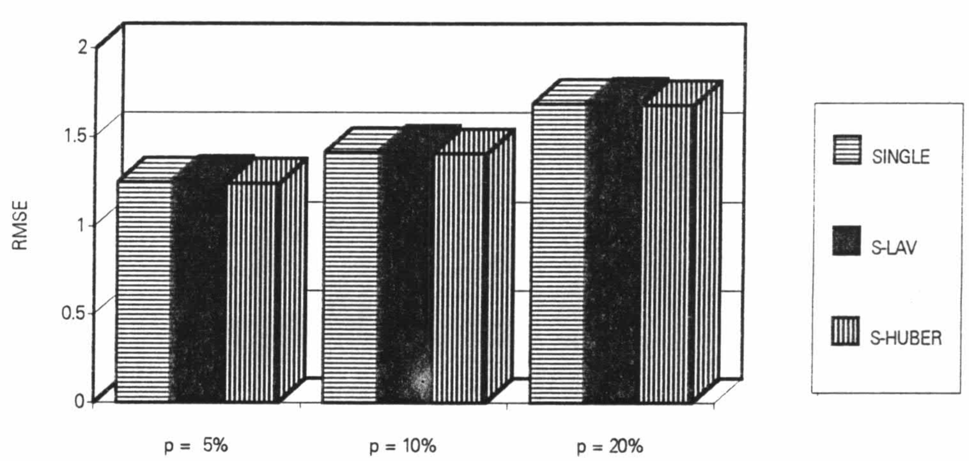
n	p	วิธีพยากรณ์		
		SINGLE	S-LAV	S-HUBER
10	5	1.2623	1.2975	1.2444
	10	1.4725	1.4848	1.4652
	20	1.7383	1.7400	1.7300
20	5	1.2478	1.2583	1.2150
	10	1.4153	1.4281	1.4081
	20	1.6885	1.6936	1.6807
30	5	1.1860	1.1874	1.1826
	10	1.3776	1.3741	1.3684
	20	1.6738	1.6660	1.6666
50	5	1.1618	1.1631	1.1588
	10	1.3525	1.3510	1.3497
	20	1.6426	1.6393	1.6387

รูปที่ 4.2 แสดงการเปรียบเทียบค่า RMSE เฉลี่ย 12 คาบเวลาของวิธีพยากรณ์ทั้ง 3 วิธี ในการวิเคราะห์ข้อมูลอนุกรมเวลาที่มีลักษณะการเคลื่อนไหวระดับค่าเฉลี่ย เมื่อความคลาดเคลื่อนมีฟังก์ชันการแจกแจงอยู่ในรูปของ $f(x) = (1-p)\,N(0,1) + p\,N(0,c^2\sigma^2)$ เมื่อสเกลแฟกเตอร์ (c) เท่ากับ 3 และ σ เท่ากับ 1 จำแนกตามขนาดตัวอย่าง (n) และเปอร์เซ็นต์การปลอมปน (p)

SCN(0,3*3) , n = 10

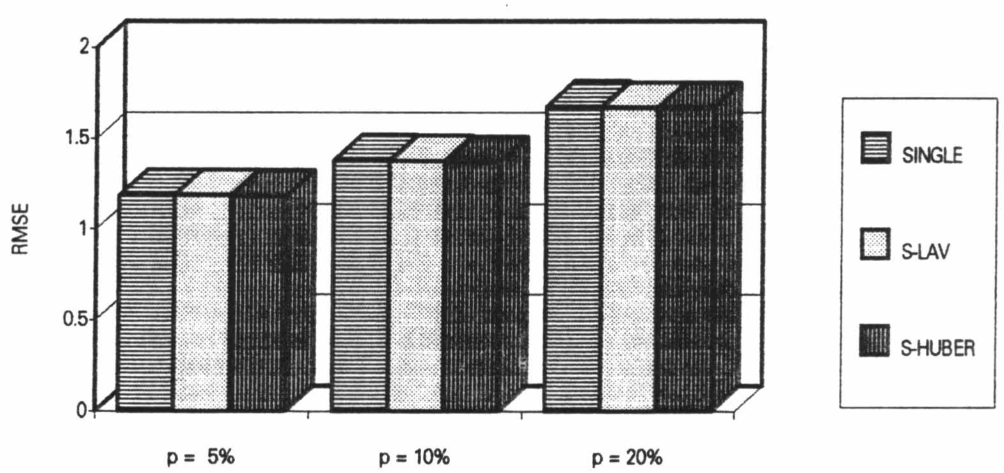


SCN(0,3*3) , n = 20

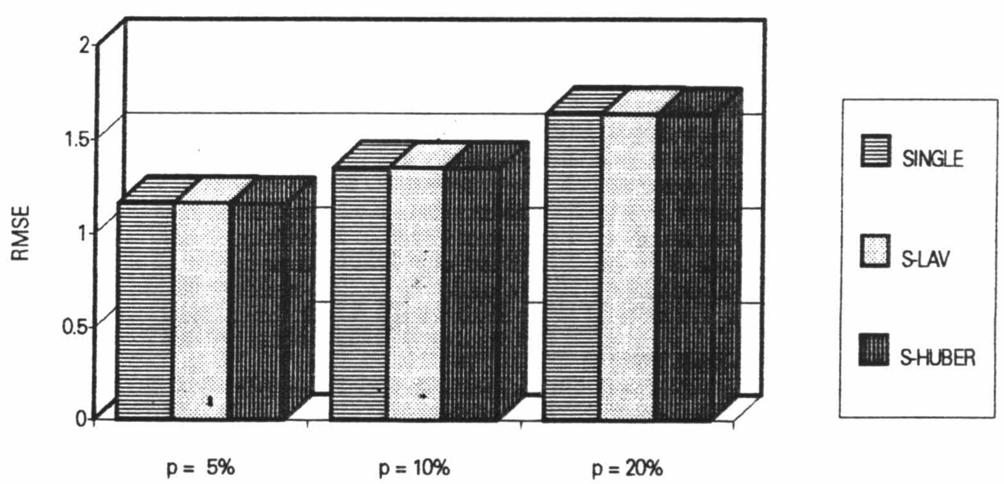


รูปที่ 4.2 (ต่อ)

SCN(0,3*3) , n = 30



SCN(0,3*3) , n = 50



จากตารางที่ 4.2 และตารางที่ 4.18 - 4.21 ได้ผลดังนี้

ในการวิเคราะห์อนุกรมเวลาที่มีลักษณะเคลื่อนไหวอยู่ในระดับค่าเฉลี่ย เมื่อความคลาดเคลื่อนมีการแจกแจงอยู่ในรูปปลอมปนที่มีฟังก์ชันการแจกแจง $f(x) = (1-p)\,N(0,1) + p\,N(0,c^2\sigma^2)$ เมื่อสเกลแฟกเตอร์ (c) เท่ากับ 3 และ σ เท่ากับ 1 เปรียบเทียบค่า RMSE พบว่า

เมื่อขนาดตัวอย่าง 10 วิธี S-HUBER ให้ค่า RMSE โดยเฉลี่ยต่ำสุดในทุกระดับของเปอร์เซ็นต์การปลอมปน ซึ่งจากการพิจารณาในทุกคาบเวลาพยากรณ์ พบว่า วิธี S-HUBER ให้ค่า RMSE ต่ำสุดในทุกคาบเวลาพยากรณ์ ที่ระดับเปอร์เซ็นต์การปลอมปนเท่ากับ 5 และ 10 สำหรับที่ระดับเปอร์เซ็นต์การปลอมปนเท่ากับ 20 พบว่า โดยทั่วไป วิธี S-HUBER ให้ค่า RMSE โดยเฉลี่ยต่ำสุด

เมื่อขนาดตัวอย่างเท่ากับ 20 วิธี S-HUBER ให้ค่า RMSE ของทุกคาบเวลา และโดยเฉลี่ยต่ำสุดในทุกระดับของเปอร์เซ็นต์การปลอมปน

เมื่อขนาดตัวอย่างเท่ากับ 30 วิธี S-HUBER ให้ค่า RMSE โดยเฉลี่ยต่ำสุดที่เปอร์เซ็นต์การปลอมปนเท่ากับ 5 และ 10 และวิธี S-LAV ให้ค่า RMSE โดยเฉลี่ยต่ำสุดเมื่อเปอร์เซ็นต์การปลอมปนเท่ากับ 20 เมื่อพิจารณาในทุกคาบเวลาพยากรณ์ พบว่า ที่ระดับเปอร์เซ็นต์การปลอมปนเท่ากับ 5 วิธี S-HUBER ให้ค่าต่ำสุดในทุกคาบเวลา และเกือบทุกคาบเวลา ยกเว้นคาบเวลาที่ 1, 2 และ 10 ที่วิธี S-LAV ให้ค่า RMSE ต่ำสุด สำหรับเมื่อเปอร์เซ็นต์การปลอมปนเท่ากับ 20 จำนวนคาบเวลาที่ให้ค่า RMSE ของวิธี S-HUBER และ S-LAV ต่ำสุดมีจำนวนใกล้เคียงกัน

เมื่อขนาดตัวอย่างเท่ากับ 50 วิธี S-HUBER ให้ค่า RMSE โดยเฉลี่ยต่ำสุดในทุกระดับของเปอร์เซ็นต์การปลอมปน ซึ่งเมื่อพิจารณาทุกคาบเวลาพยากรณ์พบว่า โดยทั่วไปวิธี S-HUBER ให้ค่า RMSE ต่ำสุด

1.3 เมื่อความคลาดเคลื่อนมีฟังก์ชันการแจกแจง $f(x) = (1-p)N(0,1) + pN(0,c^2\sigma^2)$ เมื่อสเกลแฟกเตอร์ (c) เท่ากับ 10 และ σ เท่ากับ 1

การพยากรณ์สำหรับข้อมูลที่มีลักษณะเคลื่อนไหวอยู่ในระดับค่าเฉลี่ย โดยความคลาดเคลื่อนมีการแจกแจงปลอมปนที่มีฟังก์ชันการแจกแจง $f(x) = (1-p)\,N(0,1) + pN(0,c^2\sigma^2)$ เมื่อสเกลแฟกเตอร์ (c) เท่ากับ 10 และ σ เท่ากับ 1 เปอร์เซ็นต์การปลอม (p) เท่ากับ 5, 10 และ 20 ขนาดตัวอย่างที่ใช้ในการศึกษา คือ 10, 20, 30 และ 50 ทำการ

พยากรณ์ด้วยวิธีการทำให้เรียบแบบเอกซโพเนนเชียลครั้งเดียว (SINGLE) การทำให้เรียบแบบเอกซโพเนนเชียลครั้งเดียวเมื่อประมาณค่าด้วยวิธีค่าสัมบูรณ์ต่ำสุด (S--LAV) และการทำให้เรียบแบบเอกซโพเนนเชียลครั้งเดียวเมื่อประมาณค่าด้วยวิธีการของฮูเบอร์ (S-HUBER) ทำการคำนวณค่ารากที่สองของค่าเฉลี่ยของความคลาดเคลื่อนกำลังสอง (RMSE) เฉลี่ย 12 คาบเวลา ได้ดังตารางที่ 4.3 และรูปที่ 4.3 และค่า RMSE ของ 12 คาบเวลาได้ในภาคผนวก ข ดังตารางที่ 4.22 - 4.25 และรูปที่ 4.28 - 4.31

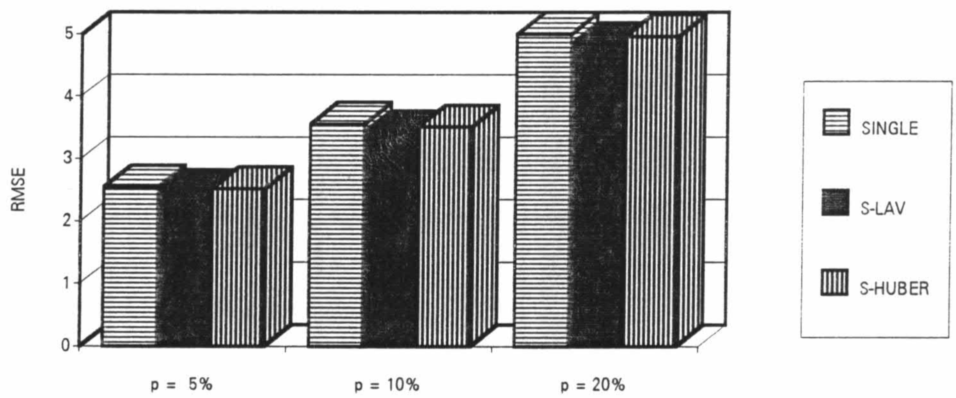
ตารางที่ 4.3 แสดงค่า RMSE เฉลี่ย 12 คาบเวลาของวิธีพยากรณ์ทั้ง 3 วิธี ในการวิเคราะห์ข้อมูลอนุกรมเวลาที่มีลักษณะเคลื่อนไหวอยู่ในระดับค่าเฉลี่ย เมื่อความคลาดเคลื่อนมีฟังก์ชันการแจกแจงอยู่ในรูป $f(x) = (1-p)\ N(0,1) + p\ N(0,c^2\sigma^2)$ เมื่อ $c = 10$ และ $\sigma = 1$ จำแนกตามขนาดตัวอย่าง (n) และเปอร์เซ็นต์การปลอมปน (p)

n	p	วิธีพยากรณ์		
		SINGLE	S-LAV	S-UBER
10	5	2.5490	2.4754	2.5225
	10	3.5535	3.4238	3.5119
	20	4.9976	4.8321	4.9600
20	5	2.5476	2.5102	2.5233
	10	3.2504	3.1825	3.2123
	20	4.7712	4.6517	4.7270
30	5	2.4653	2.4288	2.4398
	10	3.2082	3.1584	3.1684
	20	4.6531	4.5505	4.6160
50	5	2.4302	2.4042	2.4159
	10	3.1265	3.0730	3.0931
	20	4.4363	4.3570	4.3929

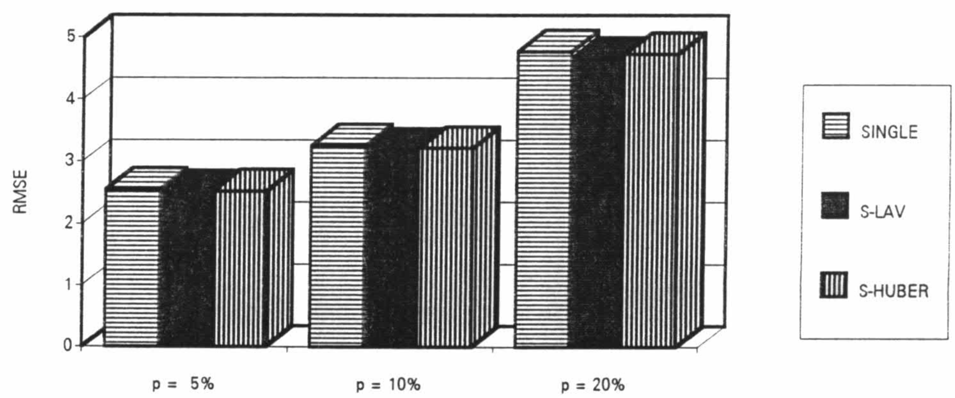
จากตารางที่ 4.3 และตารางที่ 4.22 - 4.25 ได้ผลดังนี้

รูปที่ 4.3 แสดงการเปรียบเทียบค่า RMSE เฉลี่ย 12 คาบเวลาของวิธีพยากรณ์ทั้ง 3 วิธี ในการวิเคราะห์ข้อมูลอนุกรมเวลาที่มีลักษณะการเคลื่อนไหวระดับค่าเฉลี่ย เมื่อความคลาดเคลื่อนมีฟังก์ชันการแจกแจงอยู่ในรูปของ $f(x) = (1-p)\ N(0,1) + p\ N(0,c^2\sigma^2)$ เมื่อสเกลแฟกเตอร์ (c) เท่ากับ 10 และ σ เท่ากับ 1 จำแนกตามขนาดตัวอย่าง (n) และเปอร์เซ็นต์การปลอมปน (p)

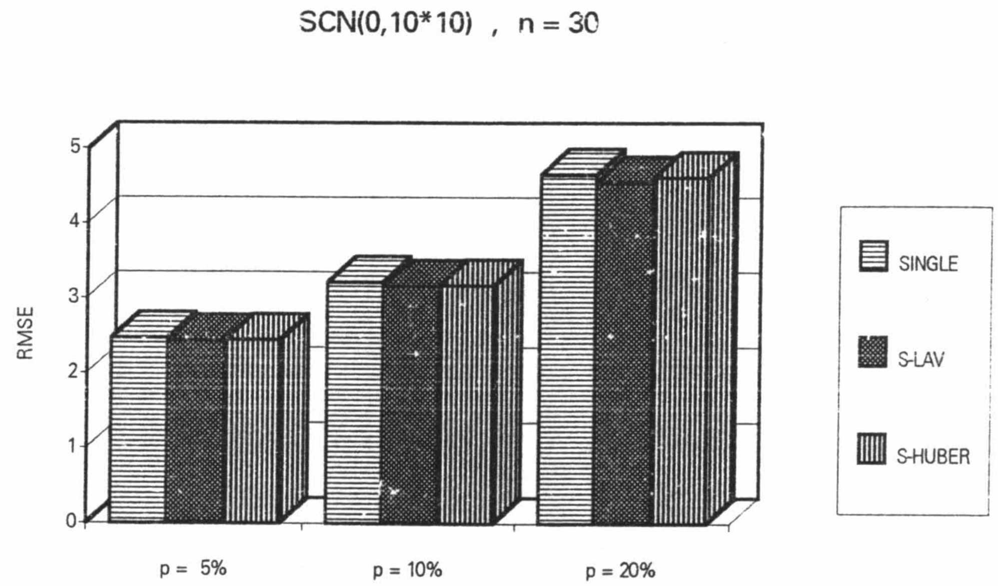
SCN(0,10*10) , n = 10

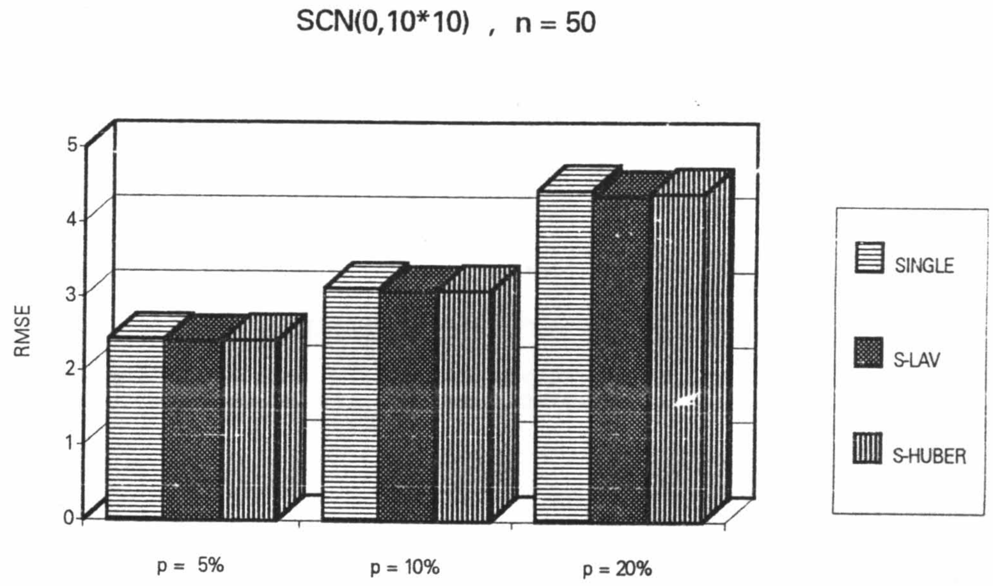


SCN(0,10*10) , n = 20



รูปที่ 4.3 (ต่อ)


SCN(0,10*10) , n = 30
RMSE
0
1
2
3
4
5
p = 5%
p = 10%
p = 20%
SINGLE
S-LAV
S-HUBER


SCN(0,10*10) , n = 50
RMSE
0
1
2
3
4
5
p = 5%
p = 10%
p = 20%
SINGLE
S-LAV
S-HUBER

ในการวิเคราะห์อนุกรมเวลาที่มีลักษณะเคลื่อนไหวอยู่ในระดับค่าเฉลี่ย เมื่อความคลาดเคลื่อนมีการแจกแจงอยู่ในรูปปลอมปนที่มีฟังก์ชันการแจกแจง $f(x) = (1-p)\ N(0,1) + p\ N(0,c^2\sigma^2)$ เมื่อสเกลแฟกเตอร์ (c) เท่ากับ 10 และ σ เท่ากับ 1 เปรียบเทียบค่า RMSE พบว่า

เมื่อขนาดตัวอย่าง 10 วิธี S-LAV ให้ค่า RMSE โดยเฉลี่ยต่ำสุดในทุกระดับของเปอร์เซ็นต์การปลอมปน ซึ่งจากการพิจารณาในทุกคาบเวลาพยากรณ์พบว่า วิธี S-LAV ให้ค่า RMSE ต่ำสุดในทุกคาบเวลาพยากรณ์ของทุกระดับของเปอร์เซ็นต์การปลอมปน ยกเว้นคาบเวลาที่ 5 วิธี S-HUBER ให้ค่า RMSE ต่ำสุด

เมื่อขนาดตัวอย่างเท่ากับ 20 วิธี S-LAV ให้ค่า RMSE โดยเฉลี่ยต่ำสุดในทุกระดับของเปอร์เซ็นต์การปลอมปน ซึ่งจากการพิจารณาทุกคาบเวลาพยากรณ์พบว่า วิธี S-LAV ให้ค่า RMSE ต่ำสุดในทุกคาบเวลา ยกเว้นคาบเวลาที่ 6 ที่ระดับการปลอมปนเท่ากับ 5 วิธี S-HUBER ให้ค่า RMSE ต่ำสุด

เมื่อขนาดตัวอย่างเท่ากับ 30 วิธี S-LAV ให้ค่า RMSE ของทุกคาบเวลาและโดยเฉลี่ยต่ำสุดในทุกระดับของเปอร์เซ็นต์การปลอมปน

เมื่อขนาดตัวอย่างเท่ากับ 50 วิธี S-LAV ให้ค่า RMSE โดยเฉลี่ยต่ำสุดในทุกระดับของเปอร์เซ็นต์การปลอมปน ซึ่งเมื่อพิจารณาทุกคาบเวลาพยากรณ์พบว่า คาบเวลาโดยทั่วไป วิธี S-LAV ให้ค่า RMSE ต่ำสุด

1.4 เมื่อความคลาดเคลื่อนมีฟังก์ชันการแจกแจง $f(x) = (1-p)N(0,1) + pL(0,\beta)$ เมื่อ β เท่ากับ 1

การพยากรณ์สำหรับข้อมูลที่มีลักษณะเคลื่อนไหวอยู่ในระดับค่าเฉลี่ย ความคลาดเคลื่อนมีการแจกแจงปลอมปนที่มีฟังก์ชันการแจกแจง $f(x) = (1-p)\ N(0,1) + p\ L(0,\beta)$ เมื่อ β เท่ากับ 1 ระดับเปอร์เซ็นต์การปลอม (p) เท่ากับ 5, 10 และ 20 ขนาดตัวอย่างที่ใช้ในการศึกษาคือ 10, 20, 30 และ 50 ทำการพยากรณ์ด้วยวิธีการทำให้เรียบแบบเอกซโพเนนเชียลครั้งเดียว (SINGLE) การทำให้เรียบแบบเอกซโพเนนเชียลครั้งเดียวเมื่อประมาณค่าด้วยวิธีค่าสัมบูรณ์ต่ำสุด (S--LAV) และการทำให้เรียบแบบเอกซโพเนนเชียลครั้งเดียวเมื่อประมาณค่าด้วยวิธีการของฮูเบอร์ (S-HUBER) ทำการคำนวณค่ารากที่สองของค่าเฉลี่ยของความคลาดเคลื่อนกำลังสอง (RMSE) เฉลี่ย 12 คาบเวลา ได้ดังตารางที่ 4.4 และรูปที่ 4.44 และค่า

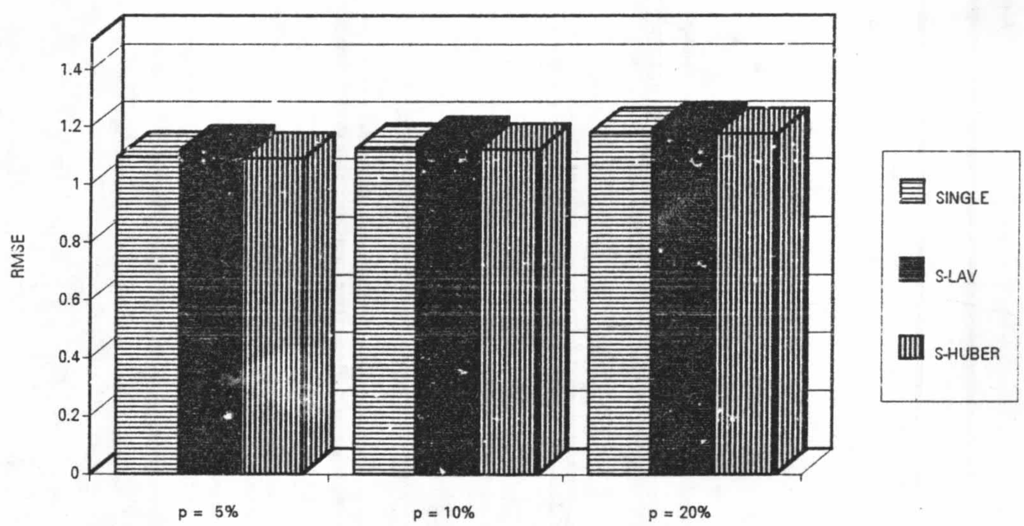
RMSE ของ 12 คาบเวลาได้ในภาคผนวก ข ดังตารางที่ 4.26 - 4.29 และรูปที่ 4.32 - 4.35

ตารางที่ 4.4 แสดงค่า RMSE เฉลี่ย 12 คาบเวลาของวิธีพยากรณ์ทั้ง 3 วิธี ในการวิเคราะห์ข้อมูลอนุกรมเวลาที่มีลักษณะเคลื่อนไหวอยู่ในระดับค่าเฉลี่ย เมื่อความคลาดเคลื่อนมีฟังก์ชันการแจกแจงอยู่ในรูป $f(x) = (1-p)\, N(0,1) + p\, L(0,\beta)$ เมื่อ $\beta = 1$ จำแนกตามขนาดตัวอย่าง (n) และเปอร์เซ็นต์การปลอมปน (p)

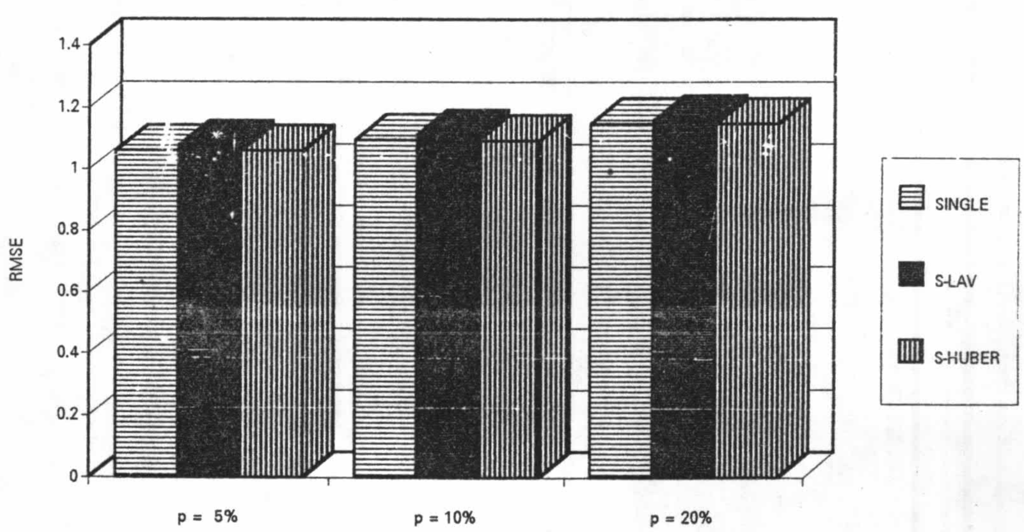
n	p	วิธีพยากรณ์		
		SINGLE	S-LAV	S-HUBER
10	5	1.0918	1.1142	1.0882
	10	1.1211	1.1452	1.1206
	20	1.1775	1.1890	1.1734
20	5	1.0602	1.0755	1.0590
	10	1.0955	1.1111	1.0932
	20	1.1477	1.1566	1.1465
30	5	1.0497	1.0615	1.0491
	10	1.0863	1.0908	1.0845
	20	1.1231	1.1291	1.1219
50	5	1.0239	1.0274	1.0227
	10	1.0582	1.0651	1.0574
	20	1.1205	1.1278	1.1189

รูปที่ 4.4 แสดงการเปรียบเทียบค่า RMSE เฉลี่ย 12 คาบเวลาของวิธีพยากรณ์ทั้ง 3 วิธี ในการวิเคราะห์ข้อมูลอนุกรมเวลาที่มีลักษณะการเคลื่อนไหวระดับค่าเฉลี่ย เมื่อความคลาดเคลื่อนมีฟังก์ชันการแจกแจงอยู่ในรูปของ f(x) = (1-p) N(0,1) + p L(0,β) เมื่อ β เท่ากับ 1 จำแนกตามขนาดตัวอย่าง (n) และเปอร์เซ็นต์การปลอมปน (p)

CNL(0,1) , n = 10

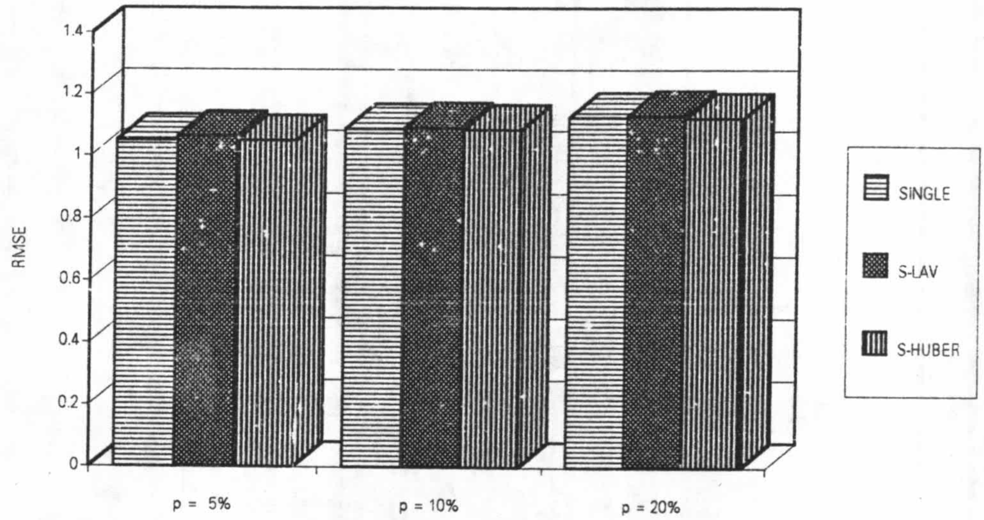


CNL(0,1) , n = 20

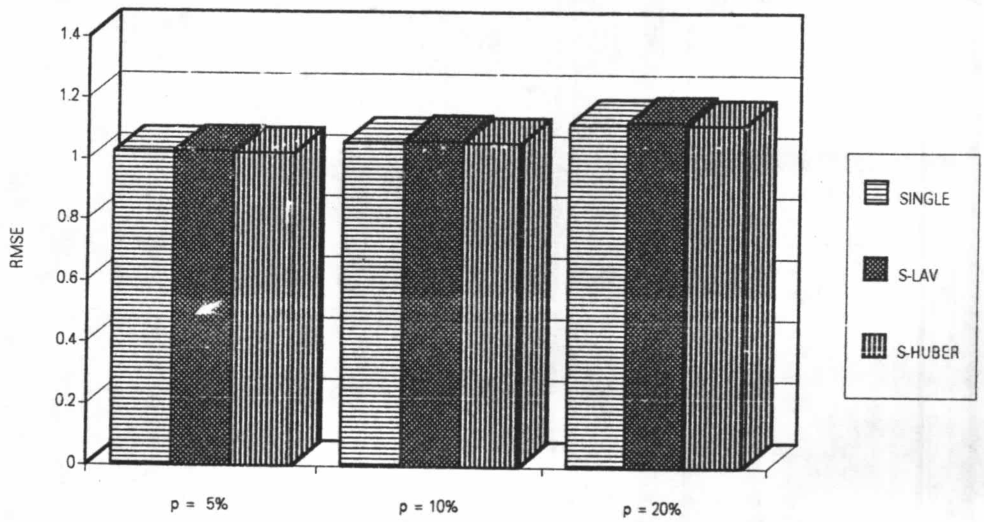


รูปที่ 4.4 (ต่อ)

CNL(0,1) , n = 30



CNL(0,1) , n = 50



จากตารางที่ 4.4 และตารางที่ 4.26 - 4.29 ได้ผลดังนี้

ในการวิเคราะห์อนุกรมเวลาที่มีลักษณะเคลื่อนไหวอยู่ในระดับค่าเฉลี่ย เมื่อความคลาดเคลื่อนมีการแจกแจงอยู่ในรูปปลอมปนที่มีฟังก์ชันการแจกแจง $f(x) = (1-p)\ N(0,1) + p\ L(0,\beta)$ เมื่อ β เท่ากับ 1 เปรียบเทียบค่า RMSE พบว่า

เมื่อขนาดตัวอย่างเท่ากับ 10 วิธี S-HUBER ให้ค่า RMSE โดยเฉลี่ยต่ำสุดในทุกระดับของเปอร์เซ็นต์การปลอมปน ซึ่งจากการพิจารณาในทุกคาบเวลาพยากรณ์พบว่าคาบเวลาโดยทั่วไปวิธี S-HUBER ให้ค่า RMSE ต่ำสุด ในทุกระดับของเปอร์เซ็นต์การปลอมปน

เมื่อขนาดตัวอย่างเท่ากับ 20 วิธี S-HUBER ให้ค่า RMSE โดยเฉลี่ยต่ำสุดในทุกระดับของเปอร์เซ็นต์การปลอมปน และเมื่อพิจารณาทุกคาบเวลาของทุกระดับเปอร์เซ็นต์การปลอมปนพบว่า วิธี S-HUBER ให้ค่า RMSE ต่ำสุดโดยส่วนใหญ่

เมื่อขนาดตัวอย่างเท่ากับ 30 วิธี S-HUBER ให้ค่า RMSE โดยเฉลี่ยต่ำสุดในทุกระดับของเปอร์เซ็นต์การปลอมปน เมื่อพิจารณาในทุกคาบเวลาพยากรณ์พบว่า วิธี S-HUBER ให้ค่าต่ำสุดในทุกคาบเวลาเมื่อเปอร์เซ็นต์การปลอมปนเท่ากับ 5 และเกือบทุกคาบเวลายกเว้นคาบเวลาที่ 2 ที่เปอร์เซ็นต์การปลอมปนเท่ากับ 10 และคาบเวลาที่ 1 และ 4 เมื่อเปอร์เซ็นต์การปลอมปนเท่ากับ 20 วิธี S-LAV ให้ค่า RMSE ต่ำสุด

เมื่อขนาดตัวอย่างเท่ากับ 50 วิธี S-HUBER ให้ค่า RMSE โดยเฉลี่ยต่ำสุดในทุกระดับของเปอร์เซ็นต์การปลอมปน ซึ่งเมื่อพิจารณาทุกคาบเวลาพยากรณ์พบว่า โดยทั่วไปวิธี S-HUBER ให้ค่า RMSE ต่ำสุด

1.5 เมื่อความคลาดเคลื่อนมีฟังก์ชันการแจกแจง $f(x) = (1-p)N(0,1) + pL(0,\beta)$ เมื่อ β เท่ากับ 10

การพยากรณ์สำหรับข้อมูลที่มีลักษณะเคลื่อนไหวอยู่ในระดับค่าเฉลี่ย ความคลาดเคลื่อนมีการแจกแจงปลอมปนที่มีฟังก์ชันการแจกแจง $f(x) = (1-p)\ N(0,1) + p\ L(0,\beta)$ เมื่อ β เท่ากับ 10 เปอร์เซ็นต์การปลอมเท่ากับ 5, 10 และ 20 ขนาดตัวอย่างที่ใช้ในการศึกษาคือ 10, 20, 30 และ 50 ทำการพยากรณ์ด้วยวิธีการทำให้เรียบแบบเอกซโพเนนเชียลครั้งเดียว (SINGLE) การทำให้เรียบแบบเอกซโพเนนเชียลครั้งเดียวเมื่อประมาณค่าด้วยวิธีค่าสัมบูรณ์ต่ำสุด (S--LAV) และการทำให้เรียบแบบเอกซโพเนนเชียลครั้งเดียวเมื่อประมาณค่าด้วยวิธีการของฮูเบอร์ (S-HUBER) ทำการคำนวณค่ารากที่สองของค่าเฉลี่ยของความคลาดเคลื่อนกำลังสอง

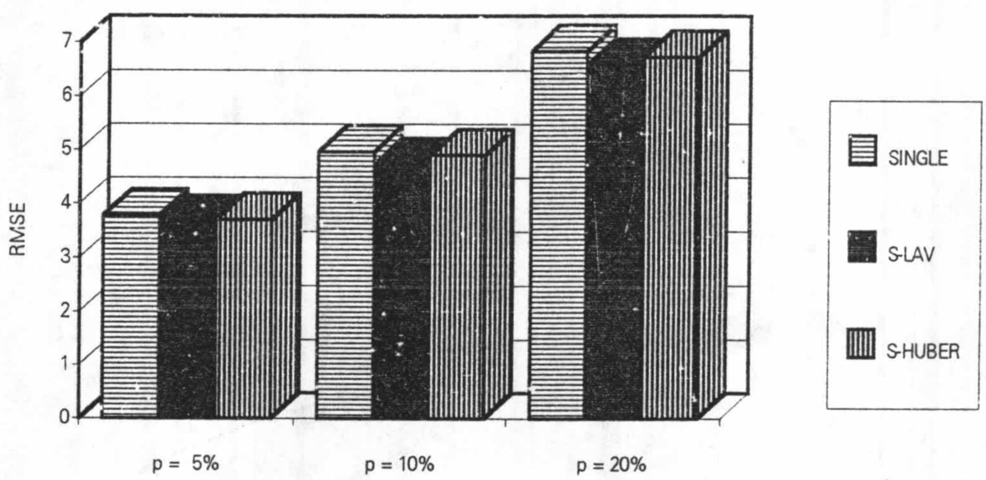
(RMSE) เฉลี่ย 12 คาบเวลาได้ดังตารางที่ 4.5 และรูปที่ 4.5 และค่า RMSE ของ 12 คาบเวลาได้ในภาคผนวก ข ดังตารางที่ 4.30 - 4.33 และรูปที่ 4.36 - 4.39

ตารางที่ 4.5 แสดงค่า RMSE เฉลี่ย 12 คาบเวลาของวิธีพยากรณ์ทั้ง 3 วิธี ในการวิเคราะห์ข้อมูลอนุกรมเวลาที่มีลักษณะเคลื่อนไหวอยู่ในระดับค่าเฉลี่ย เมื่อความคลาดเคลื่อนมีฟังก์ชันการแจกแจงอยู่ในรูป $f(x) = (1-p)\,N(0,1) + p\,L(0,\beta)$ เมื่อ $\beta = 10$ จำแนกตามขนาดตัวอย่าง (n) และ เปอร์เซ็นต์การปลอมปน (p)

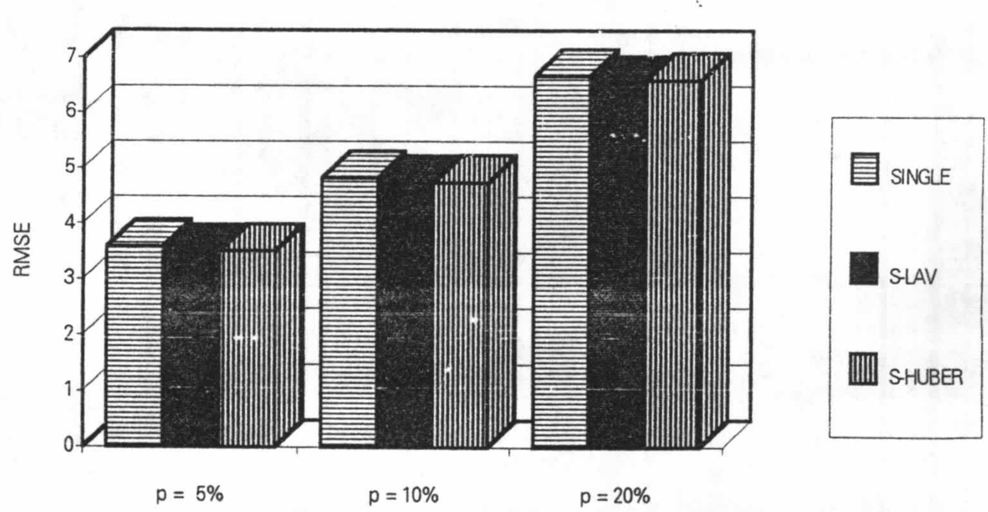
n	p	วิธีพยากรณ์		
		SINGLE	S-LAV	S-HUBER
10	5	3.7835	3.6000	3.7076
	10	4.9731	4.7285	4.9036
	20	6.8321	6.5824	6.7281
20	5	3.5845	3.4887	3.5066
	10	4.8243	4.7408	4.7472
	20	6.6672	6.5451	6.5943
30	5	3.3998	3.3625	3.3769
	10	4.7780	4.6728	4.7173
	20	6.4567	6.3305	6.4070
50	5	3.2610	3.2410	3.2439
	10	4.5206	4.4834	4.4969
	20	6.3721	6.2831	6.3094

รูปที่ 4.5 แสดงการเปรียบเทียบค่า RMSE เฉลี่ย 12 คาบเวลาของวิธีพยากรณ์ทั้ง 3 วิธี ในการวิเคราะห์ข้อมูลอนุกรมเวลาที่มีลักษณะการเคลื่อนไหวระดับค่าเฉลี่ย เมื่อความคลาดเคลื่อนมีฟังก์ชันการแจกแจงอยู่ในรูปของ $f(x) = (1-p)\ N(0,1) + p\ L(0,\beta)$ เมื่อ β เท่ากับ 10 จำแนกตามขนาดตัวอย่าง (n) และเปอร์เซ็นต์การปลอมปน (p)

CNL(0,10) , n = 10

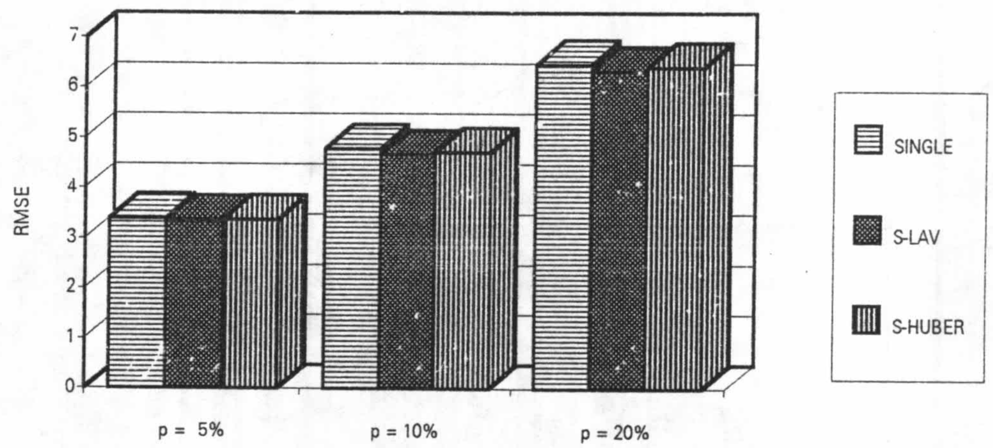


CNL(0,10) , n = 20

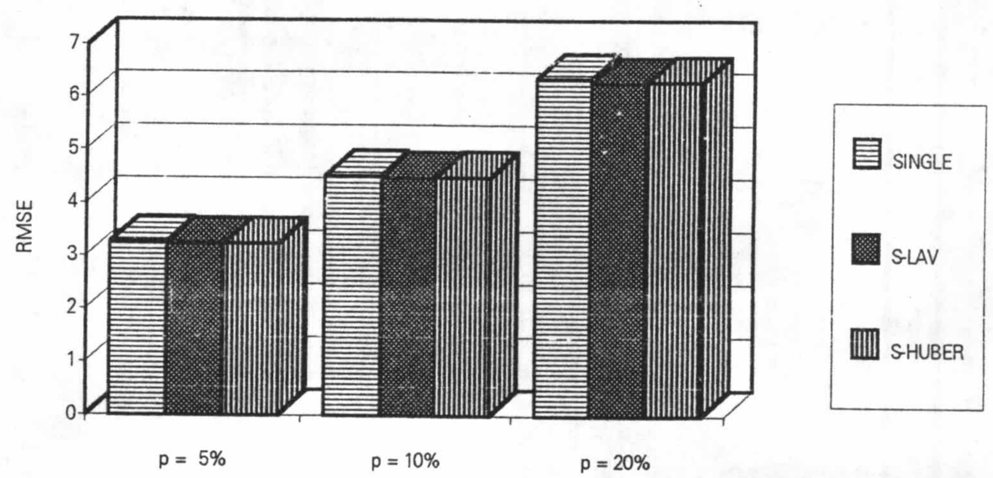


รูปที่ 4.5 (ต่อ)

CNL(0,10) , n = 30



CNL(0,10) , n = 50



จากตารางที่ 4.5 และตารางที่ 4.30 - 4.33 ได้ผลดังนี้

ในการวิเคราะห์อนุกรมเวลาที่มีลักษณะเคลื่อนไหวอยู่ในระดับค่าเฉลี่ย เมื่อความคลาดเคลื่อนมีการแจกแจงอยู่ในรูปปลอมปนที่มีฟังก์ชันการแจกแจง $f(x) = (1-p)\ N(0,1) + p\ L(0,\beta)$ เมื่อ β เท่ากับ 10 เปรียบ-เทียบค่า RMSE พบว่า

เมื่อขนาดตัวอย่างเท่ากับ 10 และ 20 วิธี S-LAV ให้ค่า RMSE ของทุกคาบเวลาและโดยเฉลี่ยต่ำสุดในทุกระดับของเปอร์เซ็นต์การปลอมปน

เมื่อขนาดตัวอย่างเท่ากับ 30 วิธี S-LAV ให้ค่า RMSE โดยเฉลี่ยต่ำสุดในทุกระดับของเปอร์เซ็นต์การปลอมปน เมื่อพิจารณาคาบเวลาของการพยากรณ์ วิธี S-LAV ให้ค่า RMSE ต่ำสุดในเกือบคาบเวลา ของทุกระดับเปอร์เซ็นต์การปลอมปน ยกเว้นคาบเวลาที่ 12 เมื่อเปอร์-เซ็นต์การปลอมปนเท่ากับ 10 วิธี S-HUBER ให้ค่า RMSE ต่ำสุด

เมื่อขนาดตัวอย่างเท่ากับ 50 วิธี S-LAV ให้ค่า RMSE โดยเฉลี่ยต่ำสุดในทุกระดับของเปอร์เซ็นต์การปลอมปน ซึ่งเมื่อพิจารณาทุกคาบเวลาพยากรณ์พบว่า คาบเวลาโดยทั่วไปวิธี S-LAV ให้ค่า RMSE ต่ำสุด ในทุกระดับของเปอร์เซ็นต์การปลอมปน

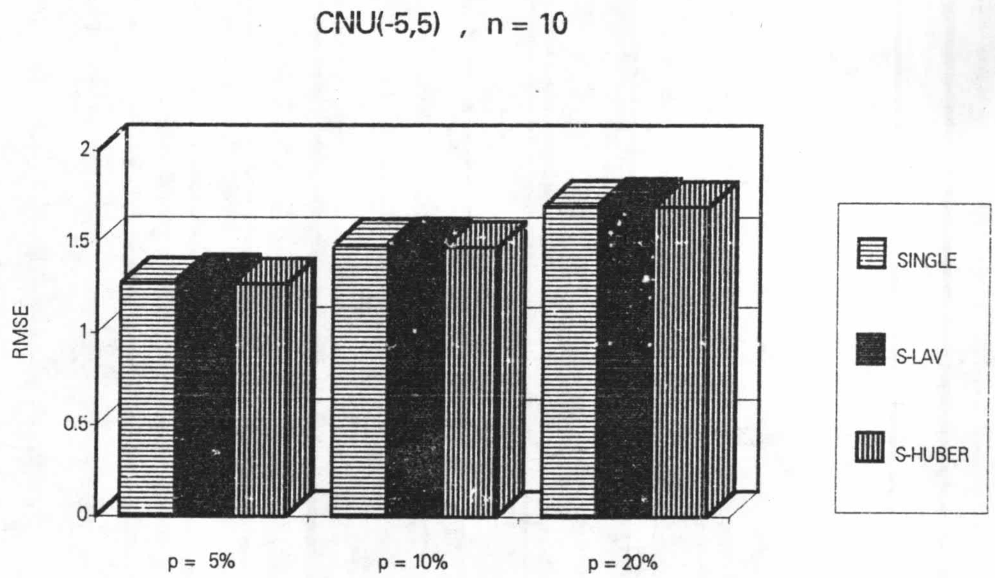
1.6 เมื่อความคลาดเคลื่อนมีฟังก์ชันการแจกแจง $f(x) = (1-p)N(0,1) + pU(-5,5)$

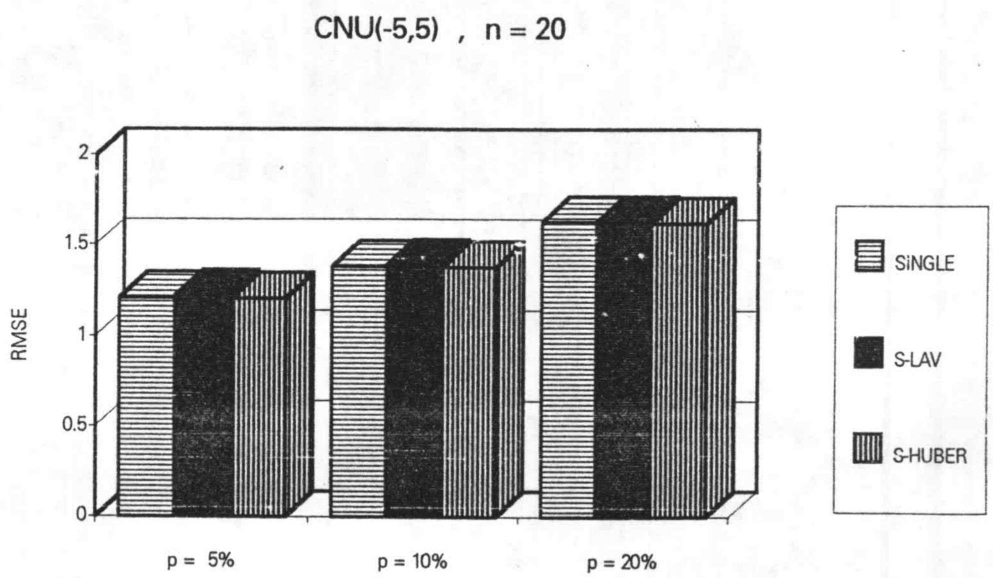
การพยากรณ์สำหรับข้อมูลที่มีลักษณะเคลื่อนไหวอยู่ในระดับค่าเฉลี่ย ความคลาดเคลื่อนมีการแจกแจงปลอมปนที่มีฟังก์ชันการแจกแจง $f(x) = (1-p)\ N(0,1) + p\ U(-5,5)$ เปอร์เซ็นต์การปลอม (p) เท่ากับ 5, 10 และ 20 ขนาดตัวอย่างที่ใช้ในการศึกษา คือ 10, 20, 30 และ 50 ทำการพยากรณ์ด้วยวิธีการทำให้เรียบแบบเอกซโพเนนเชียลครั้งเดียว (SINGLE) การทำให้เรียบแบบเอกซโพเนนเชียลครั้งเดียวเมื่อประมาณค่าด้วยวิธีค่าสัมบูรณ์ต่ำสุด (S--LAV) และการทำให้เรียบแบบเอกซโพเนนเชียลครั้งเดียวเมื่อประมาณค่าด้วยวิธีการของฮูเบอร์ (S-HUBER) การคำนวณค่ารากที่สองของค่าเฉลี่ยของความคลาดเคลื่อนกำลังสอง (RMSE) เฉลี่ย 12 คาบเวลาได้ดังตารางที่ 4.6 และรูปที่ 4.6 และค่า RMSE ของ 12 คาบเวลาได้ในภาคผนวก ข ดังตารางที่ 4.34 - 4.37 และรูปที่ 4.40 - 4.43

ตารางที่ 4.6 แสดงค่า RMSE เฉลี่ย 12 คาบเวลาของวิธีพยากรณ์ทั้ง 3 วิธี ในการวิเคราะห์ข้อมูลอนุกรมเวลาที่มีลักษณะเคลื่อนไหวอยู่ในระดับค่าเฉลี่ย เมื่อความคลาดเคลื่อนมีฟังก์ชันการแจกแจงอยู่ในรูป f(x) = (1-p) N(0,1) + p U(-5,5) จำแนกตามขนาดตัวอย่าง (n) และเปอร์เซ็นต์การปลอมปน (p)

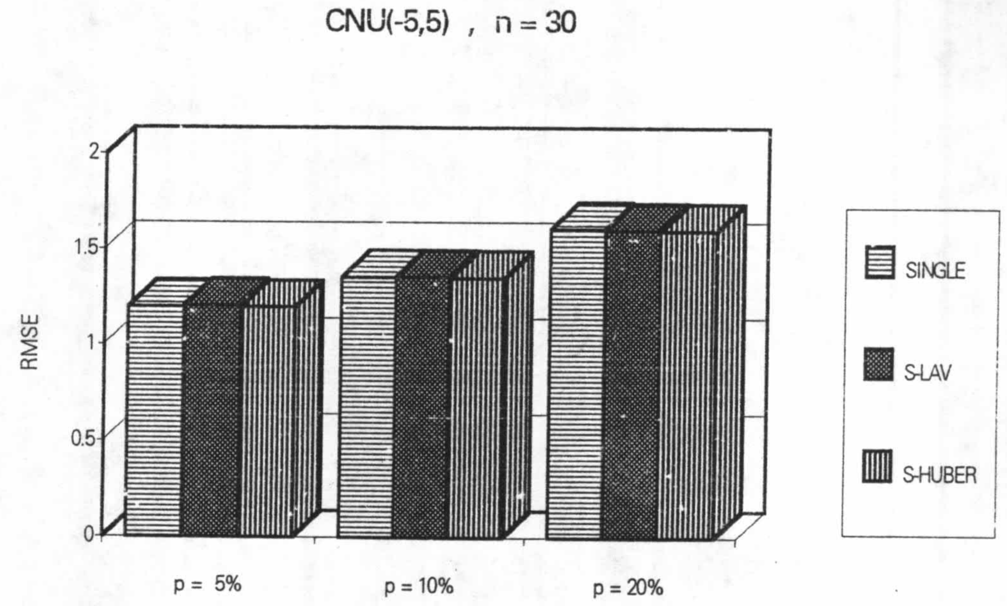
n	p	วิธีพยากรณ์		
		SINGLE	S-LAV	S-HUBER
10	5	1.2737	1.2918	1.2698
	10	1.4879	1.4968	1.4769
	20	1.7015	1.7105	1.6959
20	5	1.2110	1.2209	1.2071
	10	1.3824	1.3877	1.3772
	20	1.6272	1.6265	1.6200
30	5	1.1973	1.2047	1.1960
	10	1.3546	1.3560	1.3503
	20	1.6073	1.6016	1.5989
50	5	1.1502	1.1516	1.1479
	10	1.3408	1.3422	1.3385
	20	1.5872	1.5818	1.5830

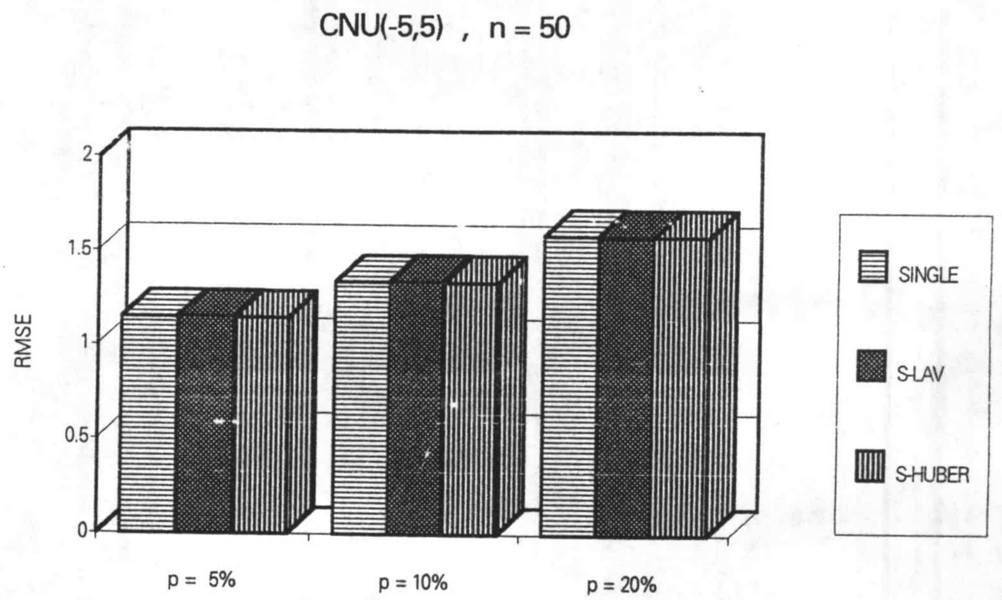
รูปที่ 4.6 แสดงการเปรียบเทียบค่า RMSE เฉลี่ย 12 คาบเวลาของวิธีพยากรณ์ทั้ง 3 วิธี ในการวิเคราะห์ข้อมูลอนุกรมเวลาที่มีลักษณะการเคลื่อนไหวระดับค่าเฉลี่ย เมื่อความคลาดเคลื่อนมีฟังก์ชันการแจกแจงอยู่ในรูปของ f(x) = 1-p) N(0,1) + p U(-5,5) จำแนกตามขนาดตัวอย่าง (n) และเปอร์เซ็นต์การปลอมปน (p)





รูปที่ 4.6 (ต่อ)


CNU(-5,5) , n = 30
RMSE
2
1.5
1
0.5
0
p = 5%
p = 10%
p = 20%
SINGLE
S-LAV
S-HUBER


CNU(-5,5) , n = 50
RMSE
2
1.5
1
0.5
0
p = 5%
p = 10%
p = 20%
SINGLE
S-LAV
S-HUBER

จากตารางที่ 4.6 และตารางที่ 4.34 - 4.37 ได้ผลดังนี้

ในการวิเคราะห์อนุกรมเวลาที่มีลักษณะเคลื่อนไหวอยู่ในระดับค่าเฉลี่ย เมื่อความคลาดเคลื่อนมีการแจกแจงอยู่ในรูปปลอมปนที่มีฟังก์ชันการแจกแจง f(x) = (1-p) N(0,1) + p U(-5,5) เปรียบเทียบค่า RMSE พบว่า

เมื่อขนาดตัวอย่าง 10 วิธี S-HUBER ให้ค่า RMSE โดยเฉลี่ยต่ำสุดในทุกระดับของเปอร์เซ็นต์การปลอมปน ซึ่งจากการพิจารณาในทุกคาบเวลาพยากรณ์พบว่า วิธี S-HUBER ให้ค่า RMSE ต่ำสุดในทุกคาบเวลาพยากรณ์ ทุกระดับเปอร์เซ็นต์การปลอมปน ยกเว้นคาบเวลาที่ 8 วิธี S-LAV ให้ค่า RMSE ต่ำสุด ที่ระดับการปลอมปนเท่ากับ 20

เมื่อขนาดตัวอย่างเท่ากับ 20, 30 และ 50 วิธี S-HUBER ให้ค่า RMSE โดยเฉลี่ยต่ำสุดในทุกระดับของเปอร์เซ็นต์การปลอมปน ซึ่งจากการพิจารณาทุกคาบเวลาพยากรณ์พบว่าโดยทั่วไปแล้ววิธี S-HUBER ให้ค่า RMSE ต่ำสุด

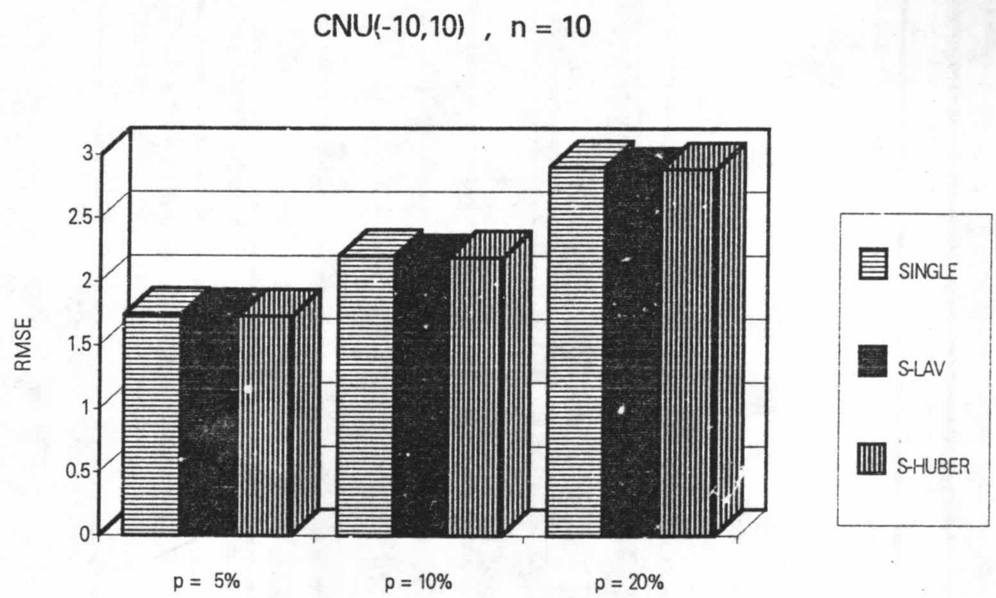
1.7 เมื่อความคลาดเคลื่อนมีฟังก์ชันการแจกแจง f(x) = (1-p)N(0,1) + p U(-10,10)

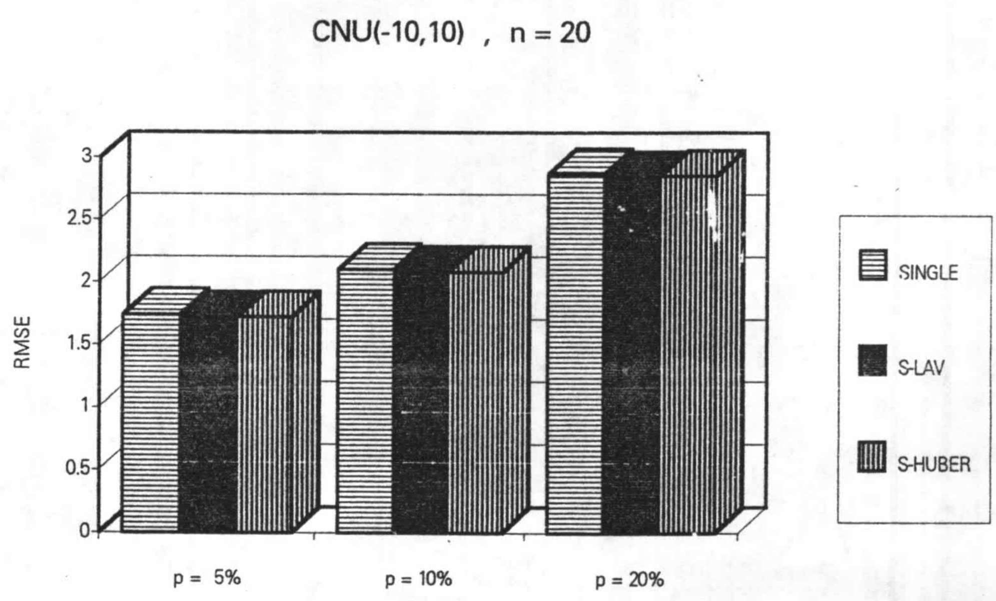
การพยากรณ์สำหรับข้อมูลที่มีลักษณะเคลื่อนไหวอยู่ในระดับค่าเฉลี่ย ความคลาดเคลื่อนมีการแจกแจงปลอมปนที่มีฟังก์ชันการแจกแจง f(x) = (1-p) N(0,1) + p U(-10,10) เปอร์เซ็นต์การปลอม (p) เท่ากับ 5, 10 และ 20 ขนาดตัวอย่างในการศึกษาคือ 10, 20, 30 และ 50 ทำการพยากรณ์ด้วยวิธีการทำให้เรียบแบบเอกซโพเนนเชียลครั้งเดียว (SINGLE) การทำให้เรียบแบบเอกซโพเนนเชียลครั้งเดียวเมื่อประมาณค่าด้วยวิธีค่าสัมบูรณ์ต่ำสุด (S--LAV) และการทำให้เรียบแบบเอกซโพเนนเชียลครั้งเดียวเมื่อประมาณค่าด้วยวิธีการของฮูเบอร์ (S-HUBER) ทำการคำนวณค่ารากที่สองของค่าเฉลี่ยของความคลาดเคลื่อนกำลังสอง (RMSE) เฉลี่ย 12 คาบเวลาได้ดังตารางที่ 4.7 และรูปที่ 4.7 และค่า RMSE ของ 12 คาบเวลาได้ในภาคผนวก ข ดังตารางที่4.38 - 4.41 และรูปที่ 4.44 - 4.47

ตารางที่ 4.7 แสดงค่า RMSE เฉลี่ย 12 คาบเวลาของวิธีพยากรณ์ทั้ง 3 วิธี ในการวิเคราะห์ข้อมูลอนุกรมเวลาที่มีลักษณะเคลื่อนไหวอยู่ในระดับค่าเฉลี่ย เมื่อความคลาดเคลื่อนมีฟังก์ชันการแจกแจงอยู่ในรูป $f(x) = (1-p)\ N(0,1) + p\ U(-10,10)$ จำแนกตามขนาดตัวอย่าง (n) และเปอร์เซ็นต์การปลอมปน (p)

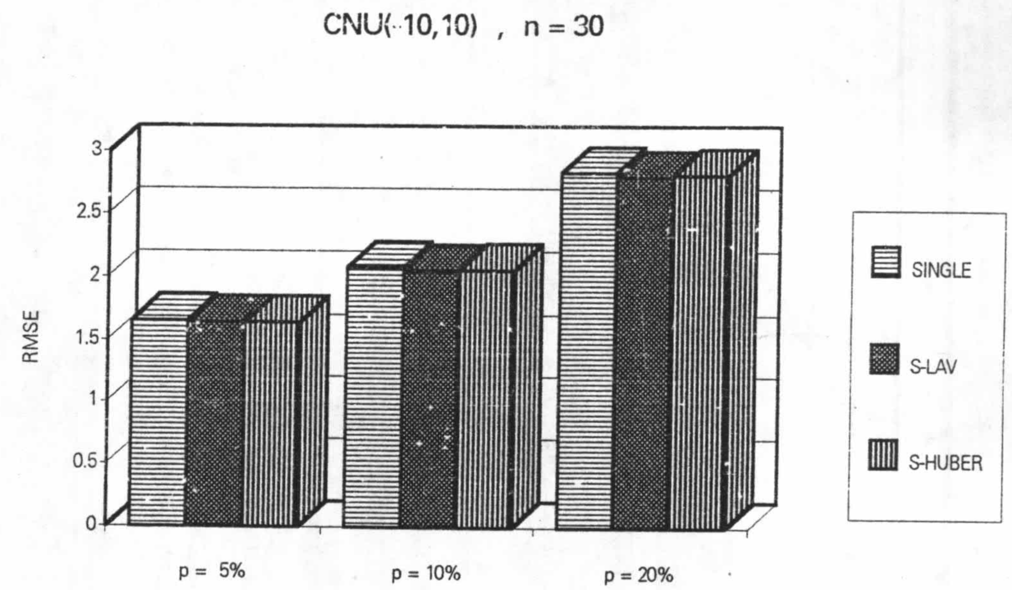
n	p	วิธีพยากรณ์		
		SINGLE	S-LAV	S-HUBER
10	5	1.7395	1.7296	1.7295
	10	2.2035	2.1607	2.1903
	20	2.9012	2.8477	2.8854
20	5	1.7391	1.7158	1.7208
	10	2.1024	2.0808	2.0811
	20	2.8711	2.8472	2.8563
30	5	1.6555	1.6434	1.6433
	10	2.0765	2.0511	2.0579
	20	2.8366	2.7984	2.8172
50	5	1.6348	1.6297	1.6295
	10	2.0060	1.9913	1.9949
	20	2.7773	2.7607	2.7689

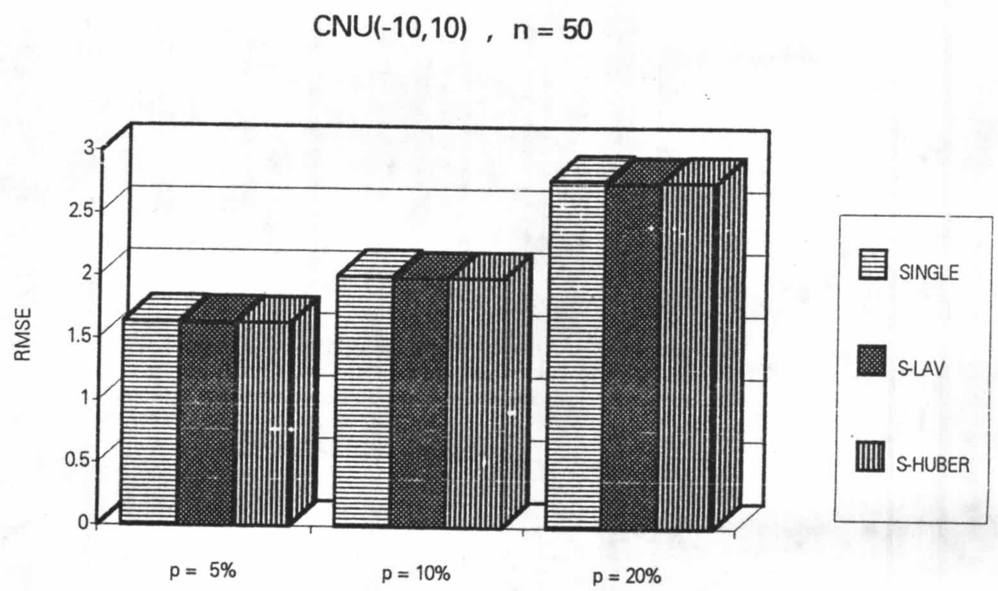
รูปที่ 4.7 แสดงการเปรียบเทียบค่า RMSE เฉลี่ย 12 คาบเวลาของวิธีพยากรณ์ทั้ง 3 วิธี ในการวิเคราะห์ข้อมูลอนุกรมเวลาที่มีลักษณะการเคลื่อนไหวระดับค่าเฉลี่ย เมื่อความคลาดเคลื่อนมีฟังก์ชันการแจกแจงอยู่ในรูปของ f(x) = 1-p) N(0,1) + p U(-10,10) จำแนกตามขนาดตัวอย่าง (n) และเปอร์เซ็นต์การปลอมปน (p)





รูปที่ 4.7 (ต่อ)


CNU(-10,10) , n = 30
3
2.5
2
1.5
1
0.5
0
RMSE
p = 5%
p = 10%
p = 20%
SINGLE
S-LAV
S-HUBER


CNU(-10,10) , n = 50
3
2.5
2
1.5
1
0.5
0
RMSE
p = 5%
p = 10%
p = 20%
SINGLE
S-LAV
S-HUBER

จากตารางที่ 4.7 และตารางที่ 4.38 - 4.41 ได้ผลดังนี้

ในการวิเคราะห์อนุกรมเวลาที่มีลักษณะเคลื่อนไหวอยู่ในระดับค่าเฉลี่ย เมื่อความคลาดเคลื่อนมีการแจกแจงอยู่ในรูปปลอมปนที่มีฟังก์ชันการแจกแจง f(x) = (1-p) N(0,1) + p U(-10,10) เปรียบเทียบค่า RMSE พบว่า

เมื่อขนาดตัวอย่าง 10 ที่ระดับเปอร์เซ็นต์การปลอมปนเท่ากับ 5 วิธี S-HUBER ให้ค่า RMSE โดยเฉลี่ยต่ำสุด ซึ่งใกล้เคียงกับวิธี S-LAV และจากการพิจารณาในทุกคาบเวลาพยากรณ์พบว่า คาบเวลาที่ให้วิธี S-LAV ให้ค่า RMSE ต่ำสุดมีจำนวนใกล้เคียงกับวิธี S-HUBER สำหรับที่เปอร์เซ็นต์การปลอมปนเท่ากับ 10 และ 20 วิธี S-LAV ให้ค่า RMSE ของทุกคาบเวลาและโดยเฉลี่ยต่ำสุด

เมื่อขนาดตัวอย่างเท่ากับ 20 วิธี S-LAV ให้ค่า RMSE โดยเฉลี่ยต่ำสุดในทุกระดับของเปอร์เซ็นต์การปลอมปน ซึ่งจากการพิจารณาทุกคาบเวลาพยากรณ์พบว่า คาบเวลาโดยทั่วไปวิธี S-LAV ให้ค่า RMSE ต่ำสุดใน ในทุกเปอร์เซ็นต์การปลอมปน

เมื่อขนาดตัวอย่างเท่ากับ 30 และ 50 วิธี S-HUBER ให้ค่า RMSE โดยเฉลี่ยต่ำสุด เมื่อเปอร์เซ็นต์การปลอมปนเท่ากับ 5 และ วิธี S-LAV .ให้ค่า RMSE โดยเฉลี่ยต่ำสุด เมื่อเปอร์-เซ็นต์การปลอมปนเท่ากับ 10 และ 20 เมื่อพิจารณาทุกคาบเวลาพยากรณ์พบว่า คาบเวลาโดยทั่วไปวิธี S-HUBER ให้ค่า RMSE ต่ำสุดเมื่อเปอร์เซ็นต์การปลอมปนเท่ากับ 5 และวิธี S-LAV ให้ค่า RMSE ต่ำสุดในทุกคาบเวลา เมื่อเปอร์เซ็นต์การปลอมปนเท่ากับ 10 และ 20

1.8 สรุปผลการวิจัยสำหรับข้อมูลที่มีลักษณะเคลื่อนไหวอยู่ในระดับค่าเฉลี่ยเมื่อความคลาดเคลื่อนมีฟังก์ชันการแจกแจงแบบต่าง ๆ

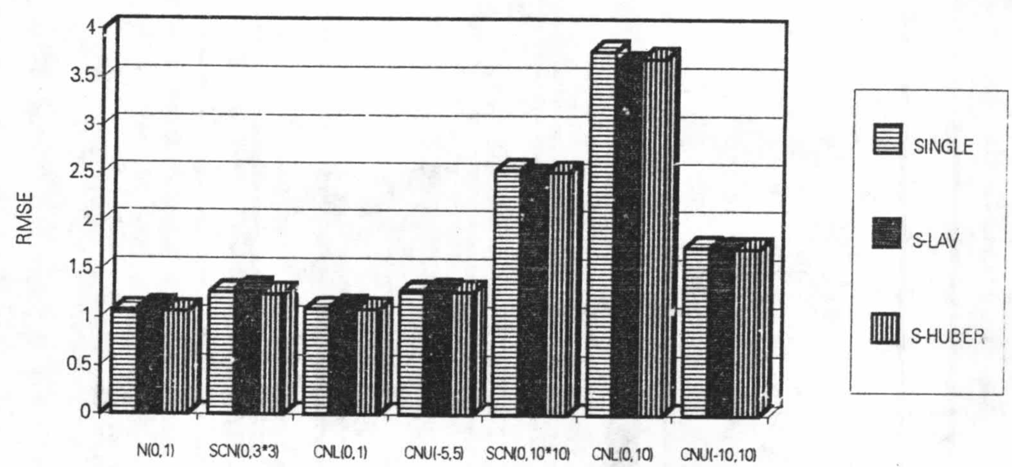
จากการวิเคราะห์ผลข้อมูลที่มีลักษณะเคลื่อนไหวอยู่ในระดับค่าเฉลี่ย เมื่อความคลาดเคลื่อนมีการแจกแจงแบบต่าง ๆ ซึ่งสามารถสรุปผลการเปรียบเทียบค่า RMSE โดยเฉลี่ย 12 คาบเวลาของวิธีพยากรณ์ทั้ง 3 วิธี เปอร์เซ็นต์การปลอมปนเท่ากับ 5, 10 และ 20 และขนาดตัวอย่างเท่ากับ 10, 20, 30 และ 50 ได้ดังตารางที่ 4.8 และรูปที่ 4.8 - 4.11

ตารางที่ 4.8 แสดงค่า RMSE เฉลี่ย 12 คาบเวลาพยากรณ์ของวิธีพยากรณ์ทั้ง 3 วิธี ในการวิเคราะห์อนุกรมเวลาที่มีลักษณะเคลื่อนไหวอยู่ในระดับค่าเฉลี่ย จำแนกตามขนาดตัวอย่าง (n) เปอร์เซ็นต์การปลอมปน (p) และการแจกแจงของความคลาดเคลื่อน

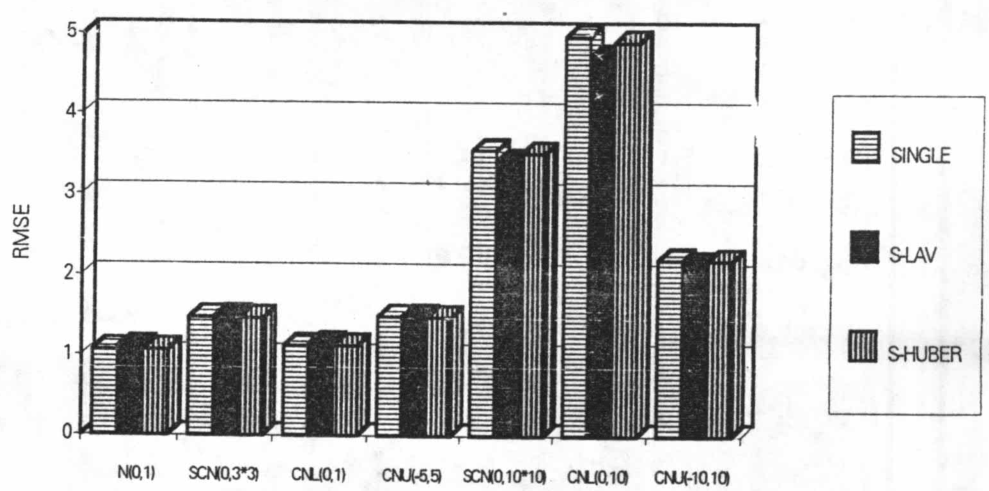
n	p	การแจกแจงของความคลาดเคลื่อน																	
		SCN(0,3*3)			CNL(0,1)			CNU(-5,5)			SCN(0,10*10)			CNL(0,10)			CNU(-10,10)		
		SINGLE	S-LAV	S-HUBER	SINGLE	S-LAV	S-UBER	SINGLE	S-LAV	S-HUBER	SINGLE	S-LAV	S-HUBER	SINGLE	S-LAV	S-HUBER	SINGLE	S-LAV	S-UBER
10	5	1.2623	1.2975	1.2444	1.0918	1.1142	1.0882	1.2737	1.2918	1.2698	2.5490	2.4754	2.5225	3.7835	3.6600	3.7076	1.7395	1.7296	1.7295
	10	1.4725	1.4848	1.4652	1.1211	1.1452	1.1206	1.4879	1.4968	1.4769	3.5535	3.4238	3.5119	4.9731	4.7285	4.9036	2.2035	2.1607	2.1903
	20	1.7383	1.7400	1.7300	1.1775	1.1890	1.1734	1.7015	1.7105	1.6959	4.9976	4.8321	4.9600	6.8321	6.5824	6.7281	2.9012	2.8477	2.8854
20	5	1.2478	1.2583	1.2150	1.0602	1.0755	1.0590	1.2110	1.2209	1.2071	2.5476	2.5102	2.5233	3.5845	3.4887	3.5066	1.7391	1.7158	1.7208
	10	1.4153	1.4281	1.4081	1.0955	1.1111	1.0932	1.3824	1.3877	1.3772	3.2504	3.1825	3.2123	4.8243	4.7408	4.7472	2.1024	2.0808	2.0811
	20	1.6885	1.6936	1.6807	1.1477	1.1566	1.1465	1.6272	1.6265	1.6200	4.7712	4.6517	4.7270	6.6672	6.5451	6.5943	2.8711	2.8472	2.8563
30	5	1.1860	1.1874	1.1826	1.0497	1.0615	1.0491	1.1973	1.2047	1.1960	2.4653	2.4288	2.4398	3.3998	3.3625	3.3769	1.6555	1.6434	1.6433
	10	1.3776	1.3741	1.3684	1.0863	1.0908	1.0845	1.3546	1.3560	1.3503	3.2082	3.1584	3.1684	4.7780	4.6728	4.7173	2.0765	2.0511	2.0579
	20	1.6738	1.6660	1.6666	1.1231	1.1291	1.1219	1.6073	1.6016	1.5989	4.6531	4.5505	4.6160	6.4567	6.3305	6.4070	2.8366	2.7984	2.8172
50	5	1.1618	1.1631	1.1588	1.0239	1.0274	1.0227	1.1502	1.1516	1.1479	2.4302	2.4042	2.4159	3.2610	3.2410	3.2439	1.6348	1.6297	1.6295
	10	1.3525	1.3510	1.3497	1.0582	1.0651	1.0574	1.3408	1.3422	1.3385	3.1265	3.0730	3.0931	4.5206	4.4834	4.4969	2.0060	1.9913	1.9949
	20	1.6426	1.6393	1.6387	1.1205	1.1278	1.1189	1.5872	1.5818	1.5830	4.4363	4.3570	4.3929	6.3721	6.2831	6.3094	2.7773	2.7607	2.7689

รูปที่ 4.8 แสดงค่า RMSE เฉลี่ย 12 คาบเวลาพยากรณ์ของวิธีทั้ง 3 วิธี ในการวิเคราะห์ข้อมูลอนุกรมเวลาที่มีลักษณะการเคลื่อนไหวระดับค่าเฉลี่ย เมื่อขนาดตัวอย่าง (n) เท่ากับ 10 จำแนกตามเปอร์เซ็นต์การปลอมปน (p) เมื่อความคลาดเคลื่อนมีการแจกแจงปลอมปนแบบต่าง ๆ โดยเปรียบเทียบกับเมื่อมีการแจกแจงปกติมาตรฐาน

n = 10, p = 5%

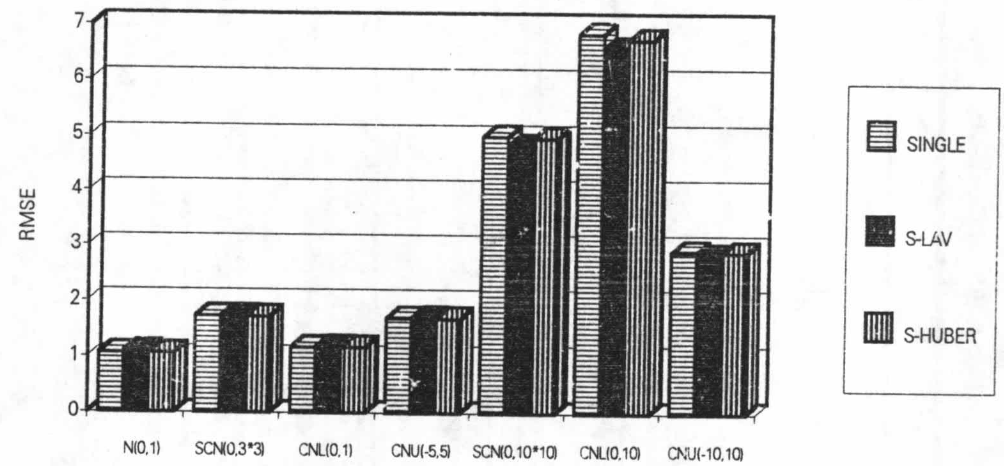


n = 10, p = 10%

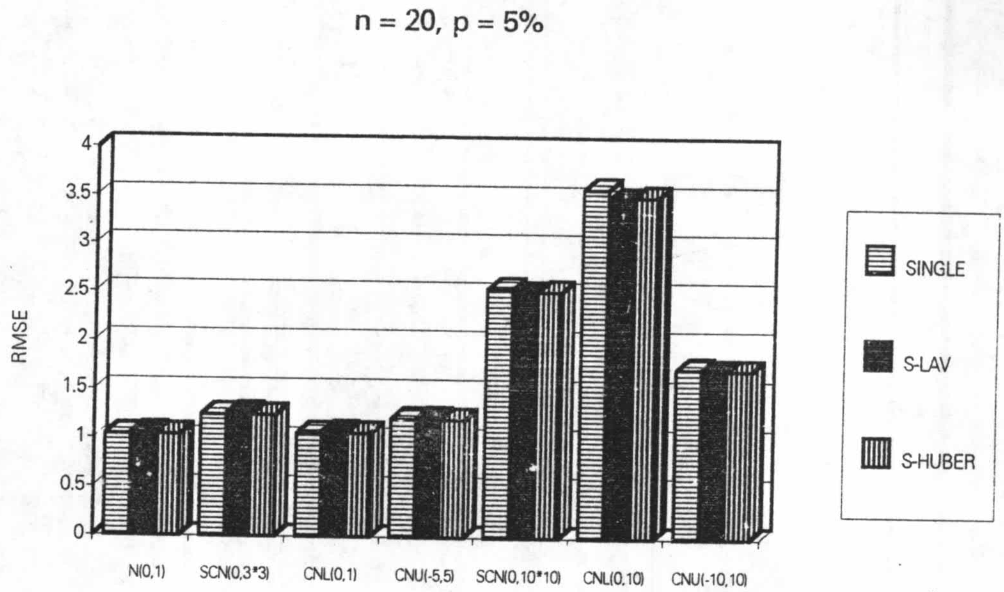


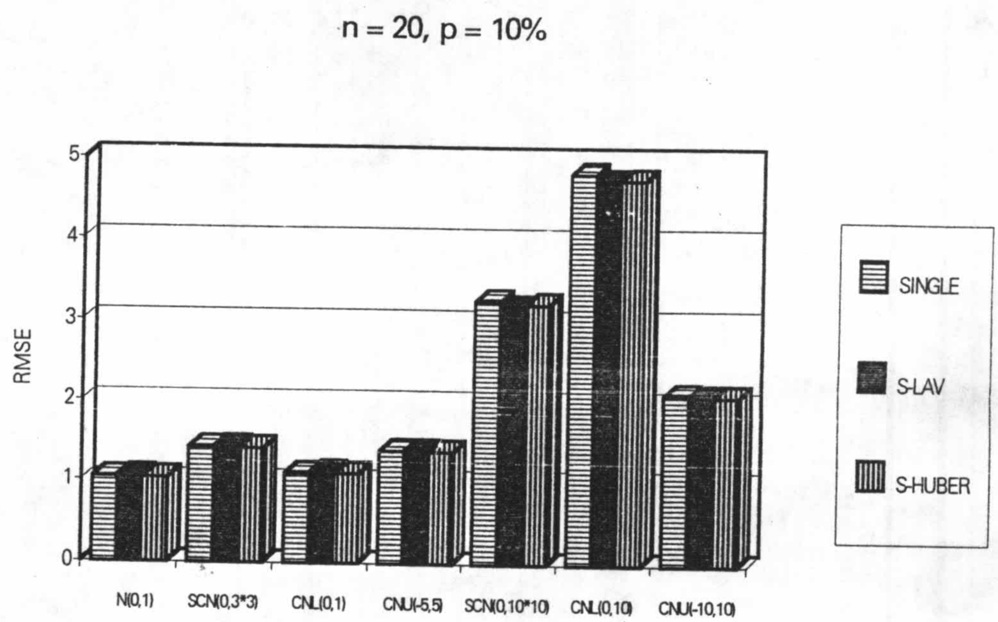
รูปที่ 4.8 (ต่อ)

n = 10, p = 20%



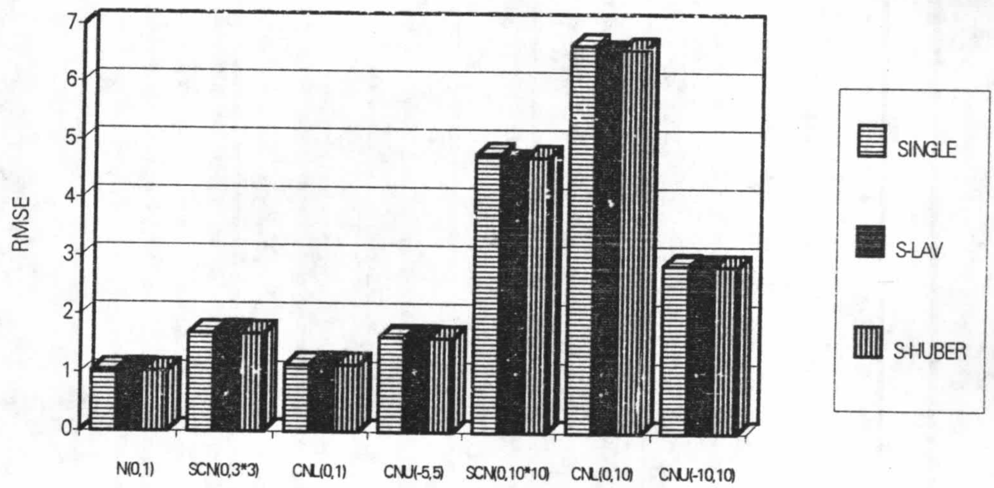
รูปที่ 4.9 แสดงค่า RMSE เฉลี่ย 12 คาบเวลาพยากรณ์ของวิธีทั้ง 3 วิธี ในการวิเคราะห์ข้อมูลอนุกรมเวลาที่มีลักษณะการเคลื่อนไหวระดับค่าเฉลี่ย เมื่อขนาดตัวอย่าง (n) เท่ากับ 20 จำแนกตามเปอร์เซ็นต์การปลอมปน (p) เมื่อความคลาดเคลื่อนมีการแจกแจงปลอมปนแบบต่าง ๆ โดยเปรียบเทียบกับเมื่อมีการแจกแจงปกติมาตรฐาน





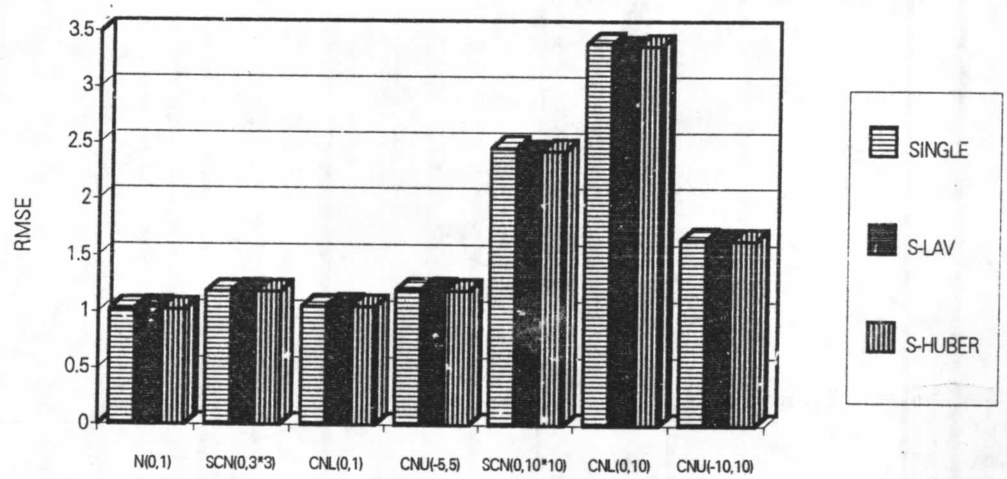
รูปที่ 4.9 (ต่อ)

n = 20, p = 20%

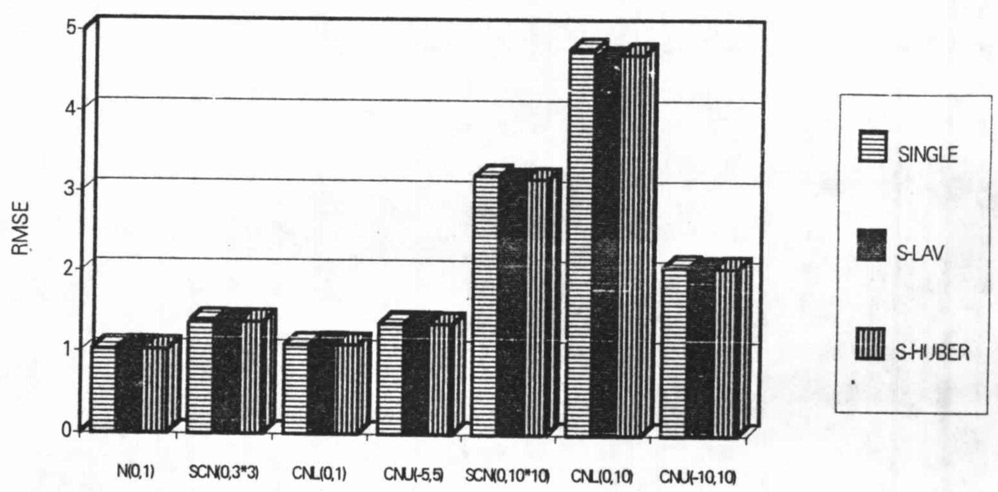


รูปที่ 4.10 แสดงค่า RMSE เฉลี่ย 12 คาบเวลาพยากรณ์ของวิธีทั้ง 3 วิธี ในการวิเคราะห์ข้อมูลอนุกรมเวลาที่มีลักษณะการเคลื่อนไหวระดับค่าเฉลี่ย เมื่อขนาดตัวอย่าง (n) เท่ากับ 30 จำแนกตามเปอร์เซ็นต์การปลอมปน (p) เมื่อความคลาดเคลื่อนมีการแจกแจงปลอมปนแบบต่าง ๆ โดยเปรียบเทียบกับเมื่อมีการแจกแจงปกติมาตรฐาน

n = 30, p = 5%

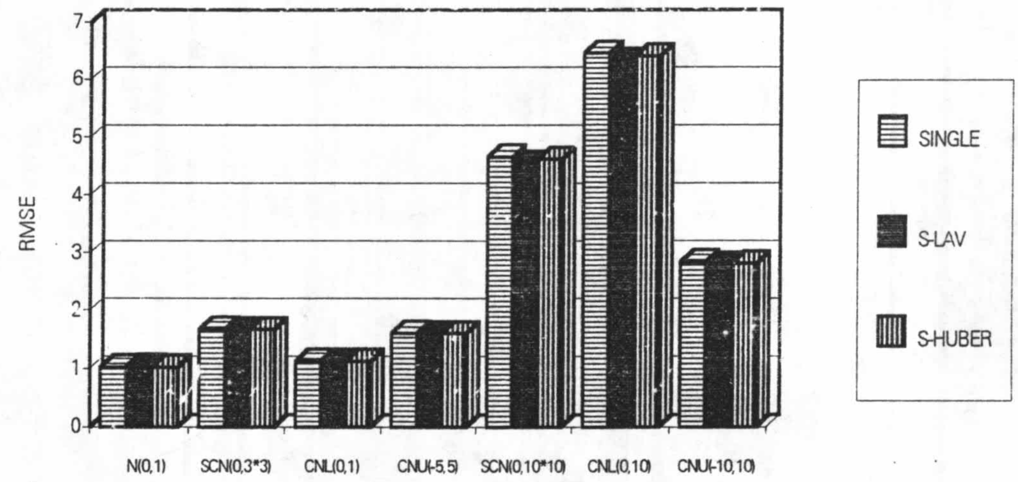


n = 30, p = 10%



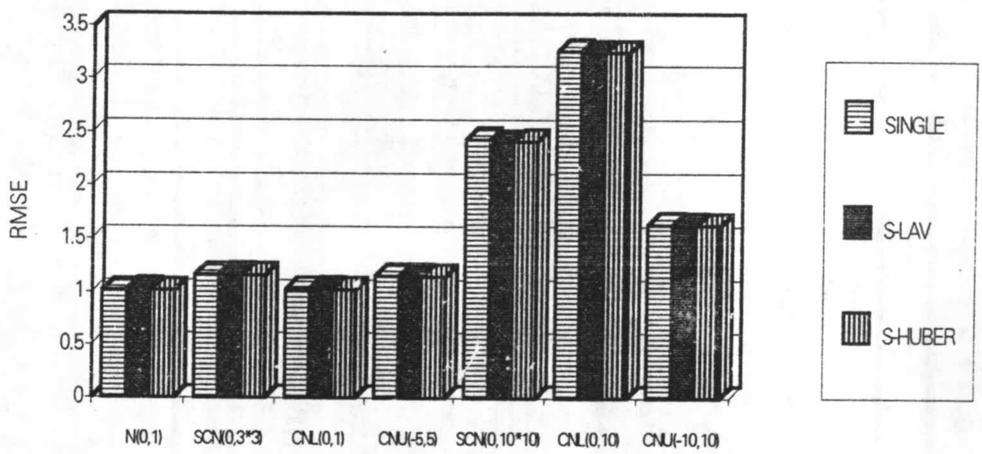
รูปที่ 4.10 (ต่อ)

n = 30, p = 20%

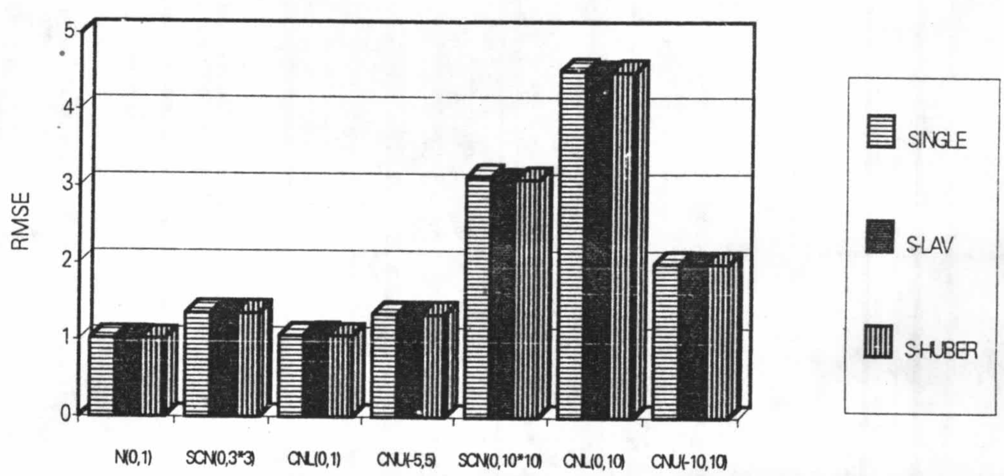


รูปที่ 4.11 แสดงค่า RMSE เฉลี่ย 12 คาบเวลาพยากรณ์ของวิธีทั้ง 3 วิธี ในการวิเคราะห์ข้อมูลอนุกรมเวลาที่มีลักษณะการเคลื่อนไหวระดับค่าเฉลี่ย เมื่อขนาดตัวอย่าง (n) เท่ากับ 50 จำแนกตามเปอร์เซ็นต์การปลอมปน (p) เมื่อความคลาดเคลื่อนมีการแจกแจงปลอมปนแบบต่าง ๆ โดยเปรียบเทียบกับเมื่อมีการแจกแจงปกติมาตรฐาน

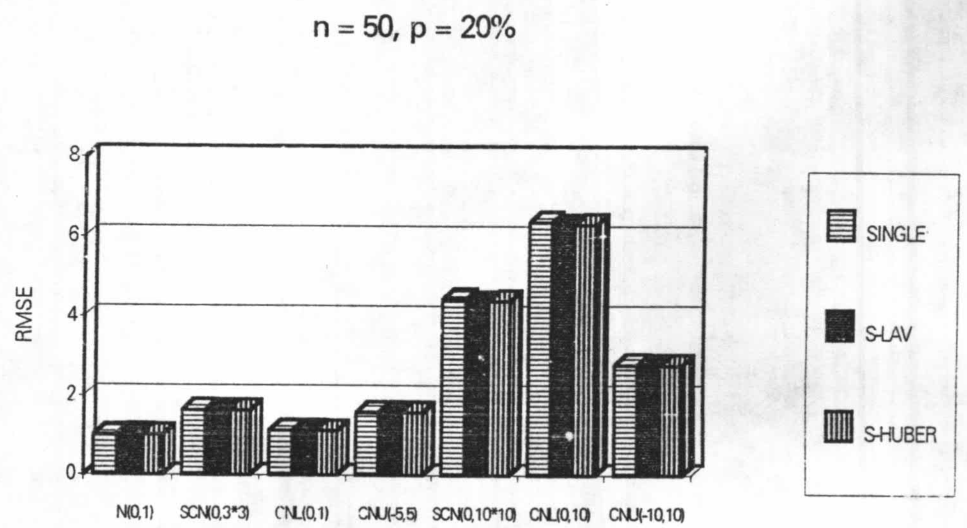
n = 50, p = 5%



n = 50, p = 10%



รูปที่ 4.11 (ต่อ)


n = 50, p = 20%
RMSE
8
6
4
2
0
N(0,1)
SCN(0,3*3)
CNL(0,1)
CNU(-5,5)
SCN(0,10*10)
CNL(0,10)
CNU(-10,10)
SINGLE
S-LAV
S-HUBER

จากตารางที่ 4.8 ผลการศึกษาเปรียบเทียบวิธีพยากรณ์ 3 วิธี ได้แก่ การทำให้เรียบแบบเอกซโพเนนเชียลครั้งเดียว การทำให้เรียบแบบเอกซโพเนนเชียลครั้งเดียวเมื่อประมาณค่าด้วยวีค่าสัมบูรณ์ต่ำสุด และการทำให้เรียบแบบเอกซโพเนนเชียลครั้งเดียวเมื่อประมาณค่าด้วยวิธีการของฮูเบอร์ เมื่อความคลาดเคลื่อนมีการแจกแจงแบบต่าง ๆ ซึ่งสามารถสรุปผลได้ดังนี้

(1) เมื่อมีความคลาดเคลื่อนมีการแจกแจงมาตรฐาน ซึ่งเป็นกรณีที่ไม่ปรากฏข้อมูลผิดปกติในข้อมูลอนุกรมเวลา วิธีการทำให้เรียบแบบเอกซโพเนนเชียลครั้งเดียวเมื่อประมาณค่าด้วยวิธีการของฮูเบอร์ให้ค่าความคลาดเคลื่อนโดยเฉลี่ยต่ำสุด ซึ่งใกล้เคียงกับวิธีการทำให้เรียบแบบเอกซโพเนนเชียลครั้งเดียวในทุกขนาดตัวอย่าง (n = 10, 20, 30 และ 50) และเมื่อพิจารณาค่าความคลาดเคลื่อนของวิธีพยากรณ์ทั้ง 3 วิธี เมื่อขนาดตัวอย่างเพิ่มขึ้นค่าความคลาดเคลื่อนของ 3 วิธีพยากรณ์ลดลง

(2) เมื่อความคลาดเคลื่อนมีการแจกแจงปกติปลอมปน $f(x) = (1-p)\,N(0,1) + p\,N(0,c^2\sigma^2)$ เมื่อ c เท่ากับ 3 และ σ เท่ากับ 1 $f(x) = (1-p)\,N(0,1) + p\,L(0,\beta)$ เมื่อ β เท่ากับ 1 และ $f(x) = (1-p)\,N(0,1) + p\,U(-5,5)$ ซึ่งกำหนดค่าพารามิเตอร์ของการแจกแจงปลอมปนให้เกิดค่าของข้อมูลผิดปกติไม่เด่นชัดในอนุกรมเวลา พบว่า วิธีการทำให้เรียบแบบเอกซโพเนนเชียลครั้งเดียวเมื่อประมาณค่าด้วยวิธีการของฮูเบอร์ ให้ค่าความคลาดเคลื่อนในการพยากรณ์โดยทั่วไปต่ำสุด และเมื่อพิจารณาค่าความคลาดเคลื่อน ณ ขนาดตัวอย่างหนึ่ง ๆ พบว่าความคลาดเคลื่อนของวิธีพยากรณ์ทั้ง 3 วิธีเพิ่มขึ้นเมื่อเปอร์เซ็นต์การปลอมปนเพิ่มขึ้น และ ณ. ระดับการปลอมปนหนึ่ง ๆเมื่อขนาดตัวอย่างเพิ่มขึ้น ค่าความคลาดเคลื่อนของวิธีพยากรณ์ทั้ง 3 วิธีลดลง

(3) เมื่อความคลาดเคลื่อนมีการแจกแจงปกติปลอมปน $f(x) = (1-p)\,N(0,1) + p\,N(0,c^2\sigma^2)$ เมื่อ c เท่ากับ 10 และ σ เท่ากับ 1 $f(x) = (1-p)\,N(0,1) + p\,L(0,\beta)$ เมื่อ β เท่ากับ 10 และ $f(x) = (1-p)\,N(0,1) + p\,U(-10,10)$ ซึ่งกำหนดค่าพารามิเตอร์ของการแจกแจงปลอมปนให้เกิดค่าของข้อมูลผิดปกติเด่นชัด พบว่าวิธีการทำให้เรียบแบบเอกซโพเนนเชียลครั้งเดียวเมื่อประมาณค่าด้วยวิธีค่าสัมบูรณ์ต่ำสุด ให้ค่าความคลาดเคลื่อนในการพยากรณ์โดยทั่วไปต่ำสุดเมื่อระดับของเปอร์เซ็นต์การปลอมปน (p = 5%, 10% และ 20%) และทุกขนาดตัวอย่าง เมื่อพิจารณาค่าความคลาดเคลื่อน ณ. ขนาดตัวอย่างหนึ่ง ๆ แล้วพบว่าความคลาดเคลื่อนของวิธีพยากรณ์ทั้ง 3 วิธีเพิ่มขึ้น เมื่อเปอร์เซ็นต์การปลอมปนเพิ่มขึ้น และ ณ ระดับการปลอมปนหนึ่ง ๆ เมื่อขนาดตัวอย่างเพิ่มขึ้นค่าความคลาดเคลื่อนของวิธีพยากรณ์ทั้ง 3 วิธีลดลง

2. <u>กรณีข้อมูลมีลักษณะแนวโน้มเชิงเส้น</u>

การวิเคราะห์ผลสำหรับข้อมูลที่มีลักษณะแนวโน้มเชิงเส้น เมื่อความคลาดเคลื่อนมีการแจกแจงแบบต่าง ๆ สามารถแสดงได้ดังนี้

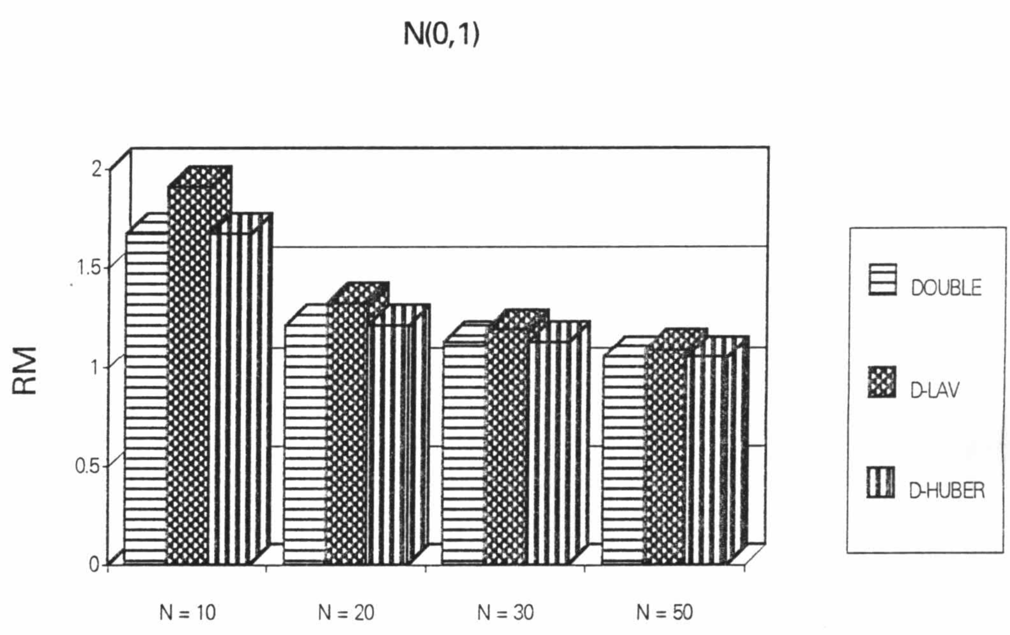
2.1 เมื่อความคลาดเคลื่อนมีการแจกแจงปกติมาตรฐาน

การพยากรณ์สำหรับข้อมูลที่มีลักษณะแนวโน้มเชิงเส้น ความคลาดเคลื่อนมีการแจกแจงปกติมาตรฐาน โดยขนาดตัวอย่างที่ใช้ในการศึกษา คือ 10, 20, 30 และ 50 ทำการพยากรณ์ด้วยวิธีการทำให้เรียบแบบเอกซโพเนนเชียลซ้ำสองครั้ง (DOUBLE) การทำให้เรียบแบบเอกซโพเนนเชียลซ้ำสองครั้งเมื่อประมาณค่าด้วยวิธีค่าสัมบูรณ์ต่ำสุด (D--LAV) และการทำให้เรียบแบบเอกซโพเนนเชียลซ้ำสองครั้งเมื่อประมาณค่าด้วยวิธีการของฮูเบอร์ (D-HUBER) และคำ-นวณค่ารากที่สองของค่าเฉลี่ยของความคลาดเคลื่อนกำลังสอง (RMSE) เฉลี่ย 12 คาบเวลา ได้ดังตารางที่ 4.9 และรูปที่ 4.12 และค่า RMSE ของ 12 คาบเวลา ได้ในภาคผนวก ข ดังตารางที่ 4.42 และรูปที่ 4.48

ตารางที่ 4.9 แสดงค่า RMSE เฉลี่ย 12 คาบเวลาของวิธีพยากรณ์ทั้ง 3 วิธี ในการวิเคราะห์ข้อมูลอนุกรมเวลาที่มีลักษณะแนวโน้มเชิงเส้น เมื่อความคลาดเคลื่อนการแจกแจงปกติมาตรฐาน

n	วิธีพยากรณ์		
	DOUBLE	D-LAV	D-HUBER
10	1.6703	1.9109	1.6681
20	1.2117	1.3196	1.2122
30	1.1255	1.1879	1.1252
50	1.0517	1.0858	1.0527

รูปที่ 4.12 แสดงค่า RMSE เฉลี่ย 12 คาบเวลาของวิธีพยากรณ์ทั้ง 3 วิธี ในการวิเคราะห์ข้อมูลอนุกรมเวลาที่มีลักษณะแนวโน้มเชิงเส้น เมื่อความคลาดเคลื่อนมีการแจกแจงปกติมาตรฐาน



จากตารางที่ 4.9 และตารางที่ 4.42 ได้ผลดังนี้

ในการวิเคราะห์อนุกรมเวลาที่มีลักษณะแนวโน้มเชิงเส้น เมื่อความคลาดเคลื่อนมีการแจกแจงปกติมาตรฐาน ซึ่งไม่มีค่าผิดปกติปรากฏในข้อมูลอนุกรมเวลา เปรียบเทียบค่า RMSE พบว่า

วิธี D-HUBER ให้ค่า RMSE โดยเฉลี่ยใกล้เคียงกับวิธี DOUBLE และใกล้เคียงมากขึ้น เมื่อขนาดตัวอย่างเพิ่มขึ้น ซึ่งจากการพิจารณา 12 คาบเวลาพยากรณ์ วิธี D-HUBER ให้ค่า RMSE ใกล้เคียงกับเมื่อใช้วิธี DOUBLE ในทุกขนาดตัวอย่าง

2.2 เมื่อความคลาดเคลื่อนมีฟังก์ชันการแจกแจง $f(x) = (1-p)N(0,1) + pN(0,c^2\sigma^2)$ เมื่อสเกลแฟกเตอร์ (c) เท่ากับ 3 และ σ เท่ากับ 1

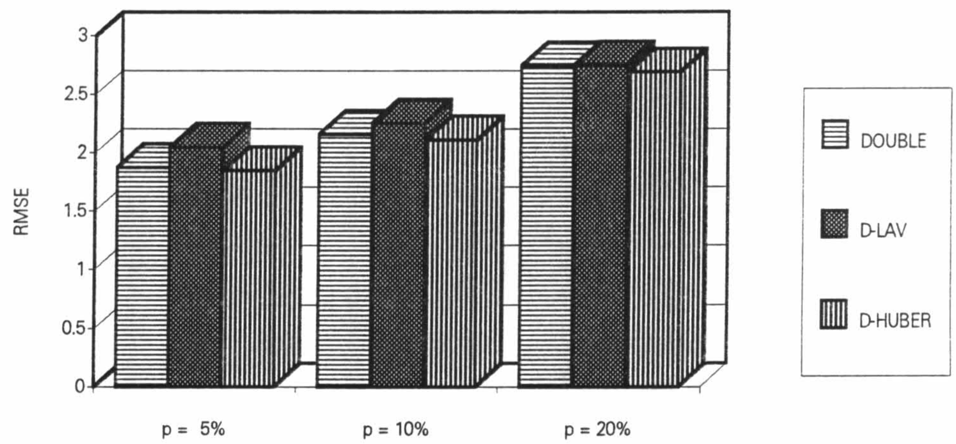
การพยากรณ์สำหรับข้อมูลที่มีลักษณะแนวโน้มเชิงเส้น ความคลาดเคลื่อนมีการแจกแจงปลอมปนที่มีฟังก์ชันการแจกแจง $f(x) = (1-p)\ N(0,1) + p\ N(0,c^2\sigma^2)$ เมื่อสเกลแฟกเตอร์ (c) เท่ากับ 3 และ σ เท่ากับ 1 เปอร์เซ็นต์การปลอม (p) เท่ากับ 5, 10 และ 20 โดยขนาดตัวอย่างที่ใช้ในการศึกษา คือ 10, 20, 30 และ 50 ทำการพยากรณ์ด้วยวิธีการทำให้เรียบแบบเอกซโพเนนเชียลซ้ำสองครั้ง (DOUBLE) การทำให้เรียบแบบเอกซโพเนนเชียลซ้ำสองครั้งเมื่อประมาณค่าด้วยวิธีค่าสัมบูรณ์ต่ำสุด (D--LAV) และการทำให้เรียบแบบเอกซโพเนนเชียลซ้ำสองครั้งเมื่อประมาณค่าด้วยวิธีการของฮูเบอร์ (D-HUBER) ทำการคำนวณค่ารากที่สองของค่าเฉลี่ยของความคลาดเคลื่อนกำลังสอง (RMSE) เฉลี่ย 12 คาบเวลา ได้ดังตารางที่ 4.10 และรูปที่ 4.13 และค่า RMSE ของ 12 คาบเวลาได้ในภาคผนวก ข ดังตารางที่ 4.43 - 4.46 และรูปที่ 4.49 - 4.52

ตารางที่ 4.10 แสดงค่า RMSE เฉลี่ย 12 คาบเวลาของวิธีพยากรณ์ทั้ง 3 วิธี ในการวิเคราะห์ข้อมูลอนุกรมเวลาที่มีลักษณะแนวโน้มเชิงเส้น เมื่อความคลาดเคลื่อนมีฟังก์ชันการแจกแจงอยู่ในรูป $f(x) = (1-p)\,N(0,1) + p\,N(0,c^2\sigma^2)$ เมื่อ $c = 3$ และ $\sigma = 1$ จำแนกตามขนาดตัวอย่าง (n) และ เปอร์เซ็นต์การปลอมปน (p)

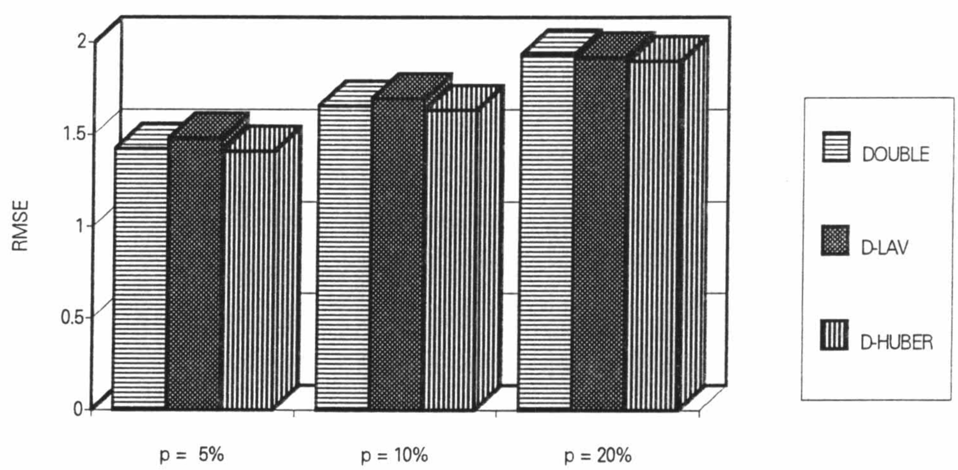
n	p	วิธีพยากรณ์		
		DOUBLE	D-LAV	D-HUBER
10	5	1.8765	2.0471	1.8472
	10	2.1596	2.2546	2.1137
	20	2.7447	2.7474	2.6884
20	5	1.4247	1.4799	1.4094
	10	1.6565	1.6943	1.6314
	20	1.9318	1.9150	1.8962
30	5	1.3184	1.3326	1.3042
	10	1.4819	1.4827	1.4616
	20	1.7709	1.7490	1.7398
50	5	1.2497	1.2668	1.2405
	10	1.4469	1.4347	1.4324
	20	1.6475	1.6386	1.6327

รูปที่ 4.13 การเปรียบเทียบค่า RMSE เฉลี่ย 12 คาบเวลาของวิธีพยากรณ์ทั้ง 3 วิธี ในการวิเคราะห์ข้อมูลอนุกรมเวลาที่มีลักษณะเชิงเส้น เมื่อความคลาดเคลื่อนมีฟังก์ชันการแจกแจงอยู่ในรูป $f(x) = (1-p)\ N(0,1) + p\ N(0,c^2.1)$ เมื่อสเกลแฟกเตอร์ (c) เท่ากับ 3 และ σ เท่ากับ 1 จำแนกตามขนาดตัวอย่าง(n) และเปอร์เซ็นต์การปลอมปน (p)

SCN(0,3*3) , n = 10

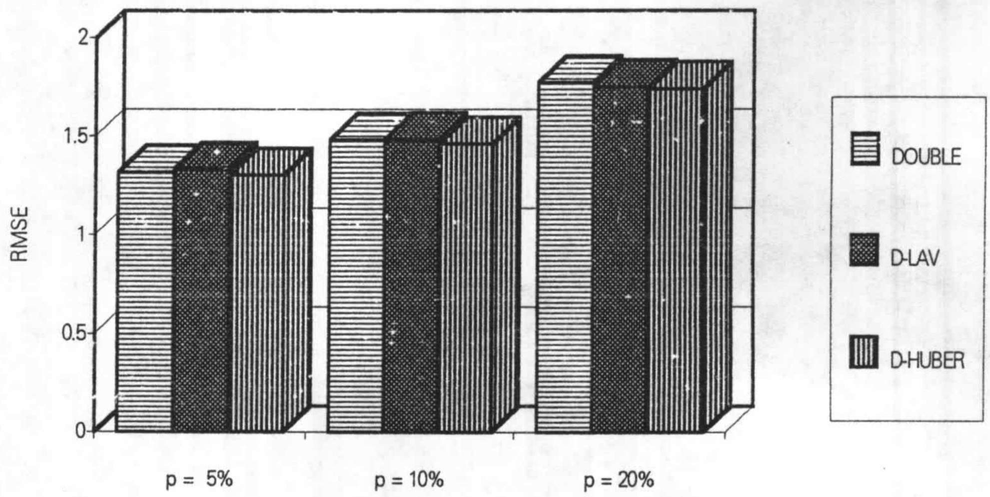


SCN(0,3*3) , n = 20

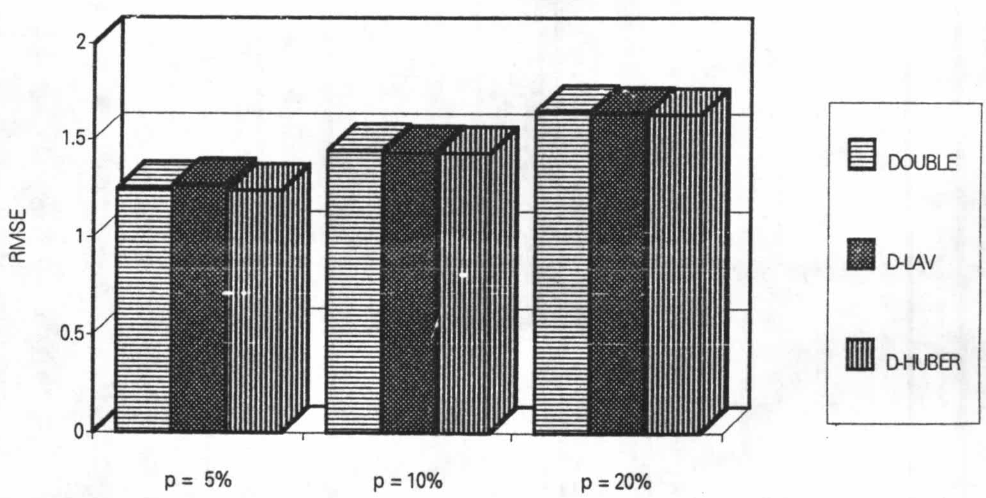


รูปที่ 4.13 (ต่อ)

SCN(0.3*3) , n = 30



SCN(0,3*3) , n = 50



จากตารางที่ 4.10 และตารางที่ 4.43 - 4.46 ได้ผลดังนี้

ในการวิเคราะห์อนุกรมเวลาที่มีลักษณะแนวโน้มเชิงเส้น เมื่อความคลาดเคลื่อนมีการแจกแจงอยู่ในรูปปลอมปนที่มีฟังก์ชันการแจกแจง $f(x) = (1-p)\ N(0,1) + p\ N(0,c^2\sigma^2)$ เมื่อสเกลแฟกเตอร์ (c) เท่ากับ 3 และ σ เท่ากับ 1 เปรียบเทียบค่า RMSE พบว่า

เมื่อขนาดตัวอย่าง 10 และ 20 วิธี D-HUBER ให้ค่า RMSE ของทุกคาบเวลาและโดยเฉลี่ยมีค่าต่ำสุด ในทุกระดับของเปอร์เซ็นต์การปลอมปน

เมื่อขนาดตัวอย่างเท่ากับ 30 วิธี D-HUBER ให้ค่า RMSE โดยเฉลี่ยต่ำสุดในทุกระดับของเปอร์เซ็นต์การปลอมปน เมื่อพิจารณาในทุกคาบเวลาพยากรณ์ พบว่าที่ระดับเปอร์เซ็นต์การปลอมปนเท่ากับ 5 วิธี D-HUBER ให้ค่าต่ำสุดในทุกคาบเวลา และยกเว้นในบางคาบเวลาที่วิธี D-LAV ให้ค่า RMSE ต่ำสุด เมื่อเปอร์เซ็นต์การปลอมปนเท่ากับ 10 และ 20

เมื่อขนาดตัวอย่างเท่ากับ 50 วิธี D-HUBER ให้ค่า RMSE โดยเฉลี่ยต่ำสุดในทุกระดับของเปอร์เซ็นต์การปลอมปน ซึ่งเมื่อพิจารณาทุกคาบเวลาพยากรณ์พบว่า โดยทั่วไปวิธี D-HUBER ให้ค่า RMSE ต่ำสุด

2.3 เมื่อความคลาดเคลื่อนมีฟังก์ชันการแจกแจง $f(x) = (1-p)N(0,1) + pN(0,c^2\sigma^2)$ เมื่อสเกลแฟกเตอร์ (c) เท่ากับ 10 และ σ เท่ากับ 1

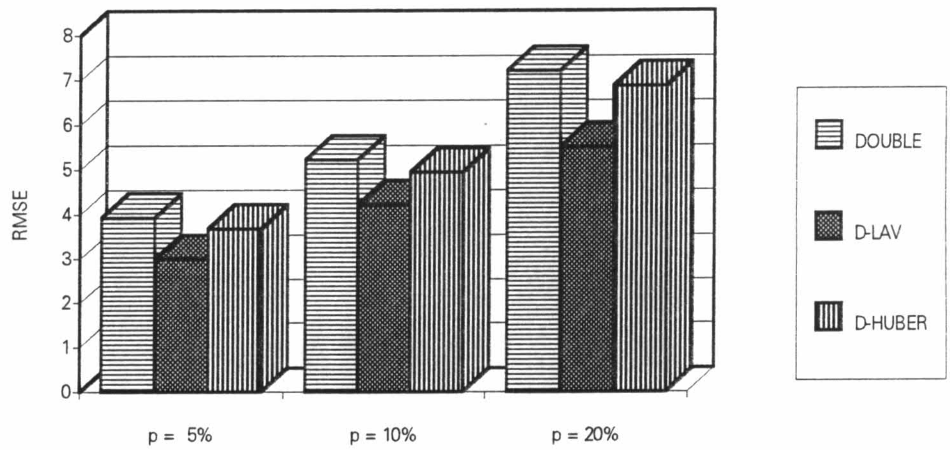
การพยากรณ์สำหรับข้อมูลที่มีลักษณะแนวโน้มเชิงเส้น ความคลาดเคลื่อนมีการแจกแจงปลอมปนที่มีฟังก์ชันการแจกแจง $f(x) = (1-p)\ N(0,1) + p\ N(0,c^2\sigma^2)$ เมื่อสเกลแฟกเตอร์ (c) เท่ากับ 10 และ σ เท่ากับ 1 เปอร์เซ็นต์การปลอม (p) เท่ากับ 5, 10 และ 20 ขนาดตัวอย่างที่ใช้ในการศึกษา คือ 10, 20, 30 และ 50 ทำการพยากรณ์ด้วยวิธีการทำให้เรียบแบบเอกซโพเนนเชียลซ้ำสองครั้ง (DOUBLE) การทำให้เรียบแบบเอกซโพเนนเชียลซ้ำสองครั้งเมื่อประมาณค่าด้วยวิธีค่าสัมบูรณ์ต่ำสุด (D--LAV) และการทำให้เรียบแบบเอกซโพเนนเชียลซ้ำสองครั้งเมื่อประมาณค่าด้วยวิธีการของฮูเบอร์ (D-HUBER) ทำการคำนวณค่ารากที่สองของค่าเฉลี่ยของความคลาดเคลื่อนกำลังสอง (RMSE) เฉลี่ย 12 คาบเวลา ได้ดังตารางที่ 4.11 และรูปที่ 4.14 และค่า RMSE ของ 12 คาบเวลาได้ในภาคผนวก ข ดังตารางที่ 4.47 - 4.50 และรูปที่ 4.53 - 4.56

ตารางที่ 4.11 แสดงค่า RMSE เฉลี่ย 12 คาบเวลาของวิธีพยากรณ์ทั้ง 3 วิธี ในการวิเคราะห์ข้อมูลอนุกรมเวลาที่มีลักษณะแนวโน้มเชิงเส้น เมื่อความคลาดเคลื่อนมีฟังก์ชันการแจกแจงอยู่ในรูป $f(x) = (1-p) N(0,1) + p N(0,c^2\sigma^2)$ เมื่อ $c = 10$ และ $\sigma = 1$ จำแนกตามขนาดตัวอย่าง (n) และ เปอร์เซ็นต์การปลอมปน (p)

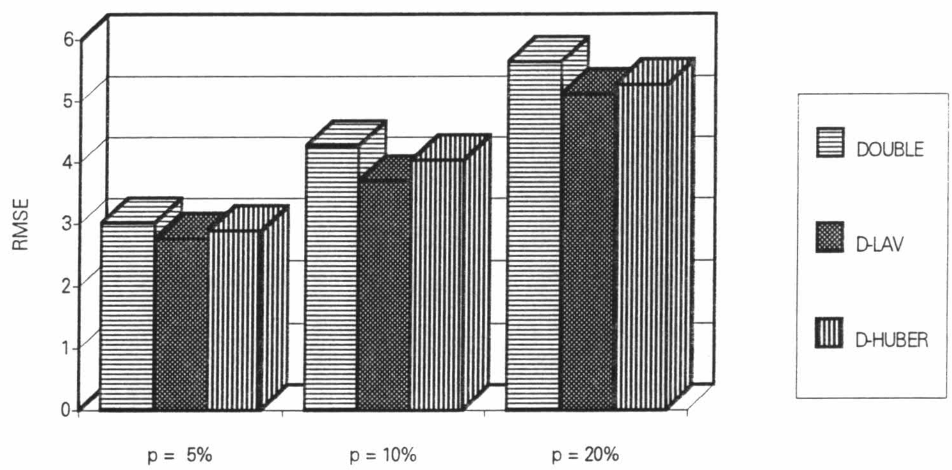
n	p	วิธีพยากรณ์		
		DOUBLE	D-LAV	D-HUBER
10	5	3.9348	3.0043	3.6901
	10	5.2256	4.2217	4.9497
	20	7.2030	5.5035	6.8696
20	5	3.0177	2.7636	2.8857
	10	4.2699	3.7011	4.0412
	20	5.6393	5.1145	5.2682
30	5	2.8369	2.6769	2.7198
	10	3.6834	3.3797	3.5535
	20	5.1036	4.6953	4.9679
50	5	2.5790	2.4814	2.5179
	10	3.4962	3.3593	3.4159
	20	4.8731	4.6256	4.7725

รูปที่ 4.14 การเปรียบเทียบค่า RMSE เฉลี่ย 12 คาบเวลาของวิธีพยากรณ์ทั้ง 3 วิธี ในการวิเคราะห์ข้อมูลอนุกรมเวลาที่มีลักษณะแนวโน้มเชิงเส้น เมื่อความคลาดเคลื่อนมีฟังก์ชันการแจกแจงอยู่ในรูป $f(x) = (1\text{-}p)\ N(0,1) + p\ N(0,c^2 .\sigma^2)$ เมื่อสเกลแฟกเตอร์ (c) เท่ากับ 10 และ σ เท่ากับ 1 จำแนกตามขนาดตัวอย่าง(n) และเปอร์เซ็นต์การปลอมปน (p)

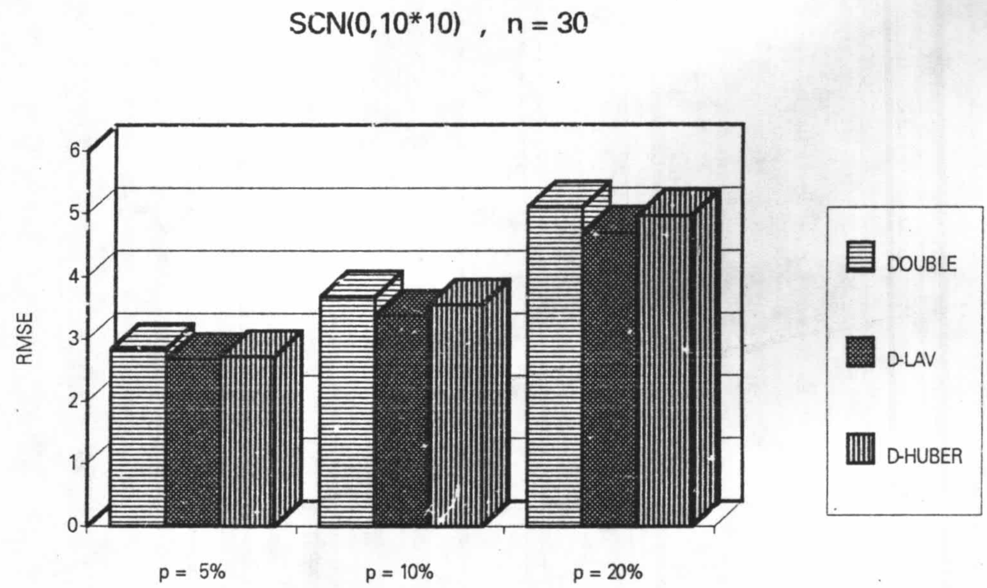
SCN(0,10*10) , n = 10

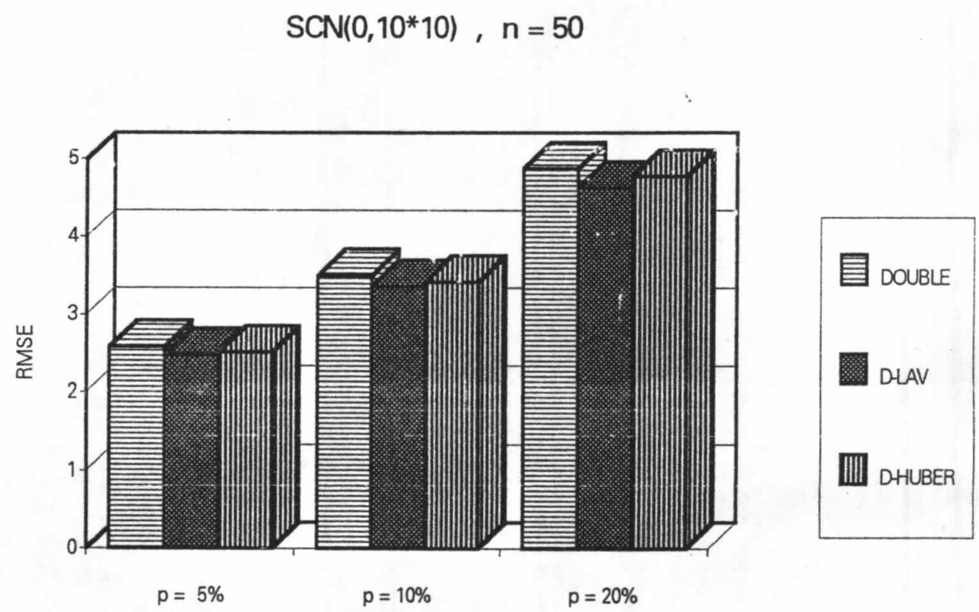


SCN(0,10*10) , n = 20



รูปที่ 4.14 (ต่อ)


SCN(0,10*10) , n = 30
RMSE
0
1
2
3
4
5
6
p = 5%
p = 10%
p = 20%
DOUBLE
D-LAV
D-HUBER


SCN(0,10*10) , n = 50
RMSE
0
1
2
3
4
5
p = 5%
p = 10%
p = 20%
DOUBLE
D-LAV
D-HUBER

จากตารางที่ 4.11 และตารางที่ 4.47 - 4.50 ได้ผลดังนี้

ในการวิเคราะห์อนุกรมเวลาที่มีลักษณะแนวโน้มเชิงเส้น เมื่อความคลาดเคลื่อนมีการแจกแจงอยู่ในรูปปลอมปนที่มีฟังก์ชันการแจกแจง $f(x) = (1\text{-}p)\ N(0,1) + p\ N(0,c^2\sigma^2)$ เมื่อสเกลแฟกเตอร์ (c) เท่ากับ 10 และ σ เท่ากับ 1 เปรียบเทียบค่า RMSE พบว่า

วิธี D-LAV ให้ค่า RMSE ของทุกคาบเวลาและโดยเฉลี่ยต่ำสุดในทุกระดับของเปอร์เซ็นต์การปลอมปน และทุกขนาดตัวอย่าง โดยความแตกต่างของความคลาดเคลื่อนของการพยากรณ์จะต่ำในคาบเวลาแรก ๆ และเพิ่มมากขึ้น เมื่อคาบเวลาเพิ่มขึ้น

2.4 เมื่อความคลาดเคลื่อนมีฟังก์ชันการแจกแจง $f(x) = (1\text{-}p)\ N(0,1) + p\ L(0,\beta)$ เมื่อ β เท่ากับ 1

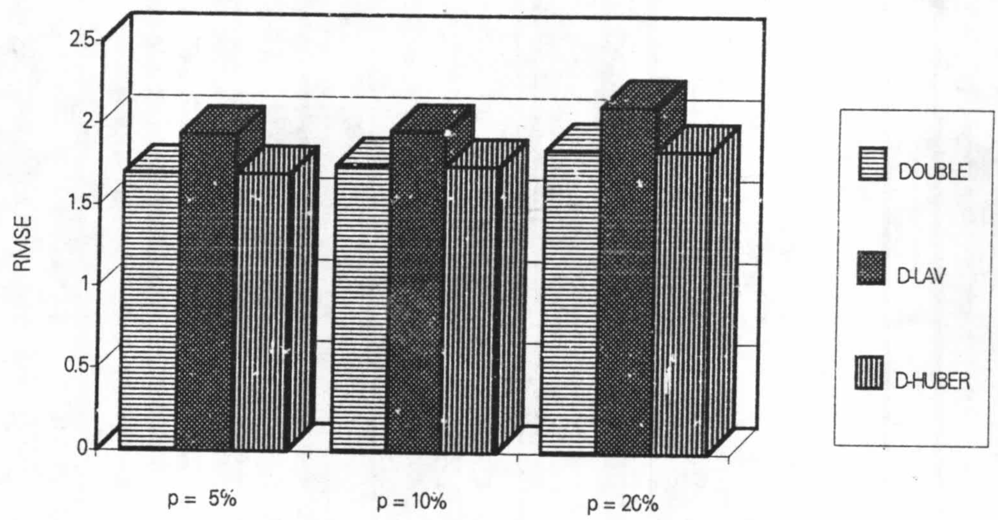
การพยากรณ์สำหรับข้อมูลที่มีลักษณะแนวโน้มเชิงเส้น ความคลาดเคลื่อนมีการแจกแจงปลอมปนที่มีฟังก์ชันการแจกแจง $f(x) = (1\text{-}p)\ N(0,1) + p\ L(0,\beta)$ เมื่อ β เท่ากับ 1 ระดับเปอร์เซ็นต์การปลอม (p) เท่ากับ 5, 10 และ 20 ขนาดตัวอย่างที่ใช้ในการศึกษา คือ 10, 20, 30 และ 50 ทำการพยากรณ์ด้วยวิธีการทำให้เรียบแบบเอกซโพเนนเชียลซ้ำสองครั้ง (DOUBLE) การทำให้เรียบแบบเอกซโพเนนเชียลซ้ำสองครั้งเมื่อประมาณค่าด้วยวิธีค่าสัมบูรณ์ต่ำสุด (D--LAV) และการทำให้เรียบแบบเอกซโพเนนเชียลซ้ำสองครั้งเมื่อประมาณค่าด้วยวิธีการของฮูเบอร์ (D-HUBER) ทำการคำนวณค่ารากที่สองของค่าเฉลี่ยของความคลาด-เคลื่อนกำลังสอง (RMSE) เฉลี่ย 12 คาบเวลา ได้ดังตารางที่ 4.12 และรูปที่ 4.15 และค่า RMSE ของ 12 คาบเวลาได้ในภาคผนวก ข ดังตารางที่ 4.51 - 4.54 และรูปที่ 4.57 - 4.60

ตารางที่ 4.12 แสดงค่า RMSE เฉลี่ย 12 คาบเวลาของวิธีพยากรณ์ทั้ง 3 วิธี ในการวิเคราะห์ข้อมูลอนุกรมเวลาที่มีลักษณะแนวโน้มเชิงเส้น เมื่อความคลาดเคลื่อนมีฟังก์ชันการแจกแจงอยู่ในรูป $f(x) = (1-p)\,N(0,1) + p\,L(0,\beta)$ เมื่อ $\beta = 1$ จำแนกตามขนาดตัวอย่าง (n) และ เปอร์เซ็นต์การปลอมปน (p)

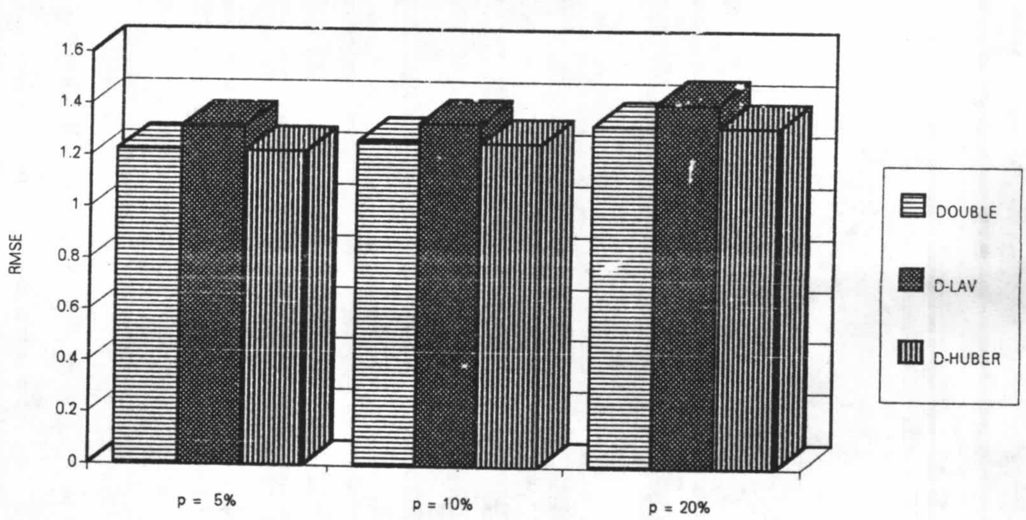
n	p	วิธีพยากรณ์		
		DOUBLE	D-LAV	D-HUBER
10	5	1.6959	1.9363	1.6935
	10	1.7478	1.9583	1.7455
	20	1.8504	2.1131	1.8453
20	5	1.2236	1.3149	1.2201
	10	1.2606	1.3306	1.2551
	20	1.3279	1.4114	1.3221
30	5	1.1499	1.1946	1.1493
	10	1.1625	1.2104	1.1582
	20	1.2340	1.2602	1.2285
50	5	1.0781	1.1035	1.0781
	10	1.1253	1.1497	1.1241
	20	1.1554	1.1660	1.1516

รูปที่ 4.15 การเปรียบเทียบค่า RMSE เฉลี่ย 12 คาบเวลาของวิธีพยากรณ์ทั้ง 3 วิธี ในการวิเคราะห์ข้อมูลอนุกรมเวลาที่มีลักษณะแนวโน้มเชิงเส้น เมื่อความคลาดเคลื่อนมีฟังก์ชันการแจกแจงอยู่ในรูป $f(x) = (1-p)\,N(0,1) + p\,L(0,\beta)$ เมื่อ $\beta = 1$ จำแนกตามขนาดตัวอย่าง (n) และเปอร์เซ็นต์การปลอมปน

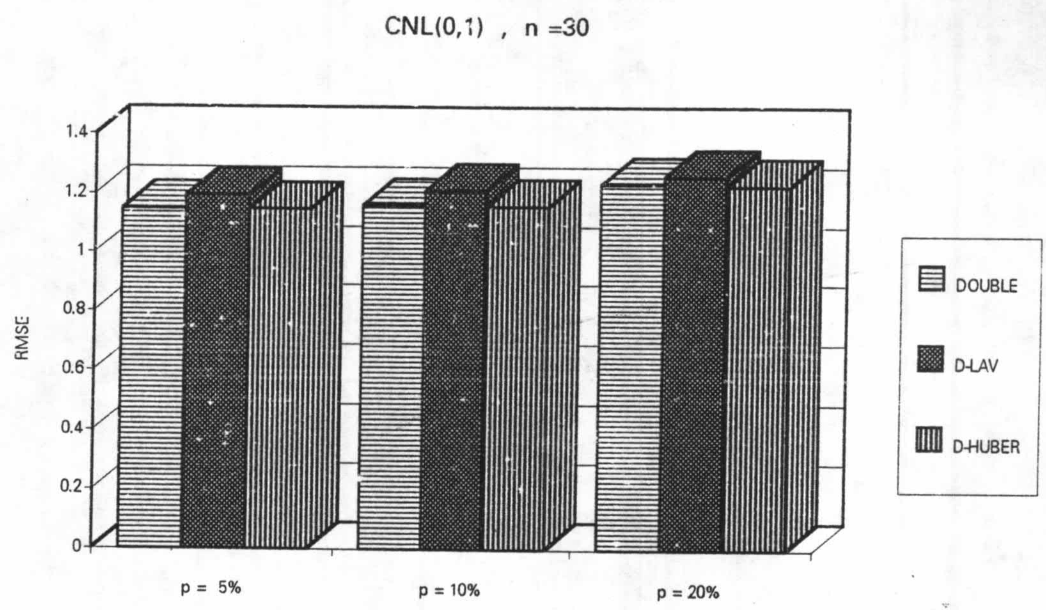
CNL(0,1) , n = 10

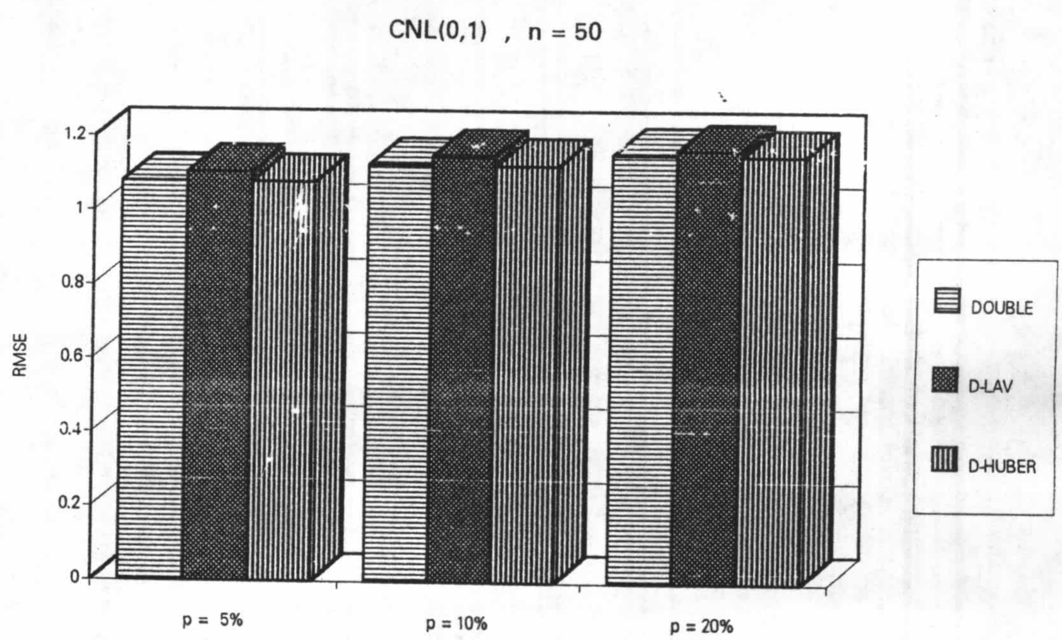


CNL(0,1) , n = 20



รูปที่ 4.15 (ต่อ)


CNL(0,1) , n =30
1.4
1.2
1
0.8
0.6
0.4
0.2
0
RMSE
p = 5%
p = 10%
p = 20%
DOUBLE
D-LAV
D-HUBER


CNL(0,1) , n = 50
1.2
1
0.8
0.6
0.4
0.2
0
RMSE
p = 5%
p = 10%
p = 20%
DOUBLE
D-LAV
D-HUBER

จากตารางที่ 4.12 และตารางที่ 4.51 - 4.54 ได้ผลดังนี้

ในการวิเคราะห์อนุกรมเวลาที่มีลักษณะแนวโน้มเชิงเส้น เมื่อความคลาดเคลื่อนมีการแจกแจงอยู่ในรูปปลอมปนที่มีฟังก์ชันการแจกแจง $f(x) = (1-p)\ N(0,1) + p\ L(0,\beta)$ เมื่อ β เท่ากับ 1 เปรียบเทียบค่า RMSE พบว่า

เมื่อขนาดตัวอย่าง 10 และ 30 วิธี S-HUBER ให้ค่า RMSE โดยเฉลี่ยต่ำสุดในทุกระดับของเปอร์เซ็นต์การปลอมปน ซึ่งจากการพิจารณาในทุกคาบเวลาพยากรณ์พบว่า คาบเวลาโดยทั่วไปวิธี D-HUBER ให้ค่า RMSE ต่ำสุดในทุกระดับของเปอร์เซ็นต์การปลอมปน ยกเว้นบางคาบเวลาที่วิธี DOUBLE ให้ค่า RMSE ต่ำสุด

เมื่อขนาดตัวอย่างเท่ากับ 20 วิธี S-HUBER ให้ค่า RMSE โดยเฉลี่ยต่ำสุดในทุกระดับของเปอร์เซ็นต์การปลอมปน และเมื่อพิจารณาทุกคาบเวลาของทุกระดับเปอร์เซ็นต์การปลอมปนพบว่า วิธี S-HUBER ให้ค่า RMSE ต่ำสุด

เมื่อขนาดตัวอย่างเท่ากับ 50 ณ ระดับเปอร์เซ็นต์การปลอมปนเท่ากับ 5 วิธี D-HUBER และวิธี DOUBLE ให้ค่า RMSE โดยเฉลี่ยต่ำสุด เมื่อพิจารณาในทุกคาบเวลาพยากรณ์พบว่าวิธีทั้งสองให้ค่า RMSE เกือบเท่ากันและเท่ากันในบางคาบเวลา เมื่อเปอร์เซ็นต์การปลอมปนเท่ากับ 10 และ 20 วิธี D-HUBER ให้ค่า RMSE โดยเฉลี่ยต่ำสุด เมื่อพิจารณาทุกคาบเวลาพบว่าโดยส่วนใหญ่วิธี D-HUBER ให้ค่า RMSE ต่ำสุด

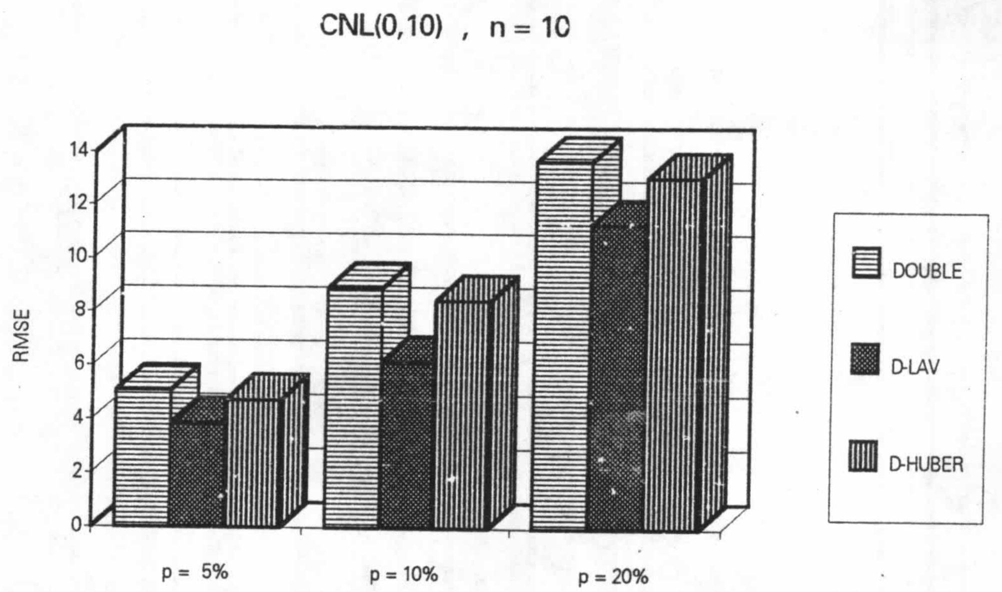
2.5 เมื่อความคลาดเคลื่อนมีฟังก์ชันการแจกแจง $f(x) = (1-p)N(0,1) + pL(0,\beta)$ เมื่อ β เท่ากับ 10

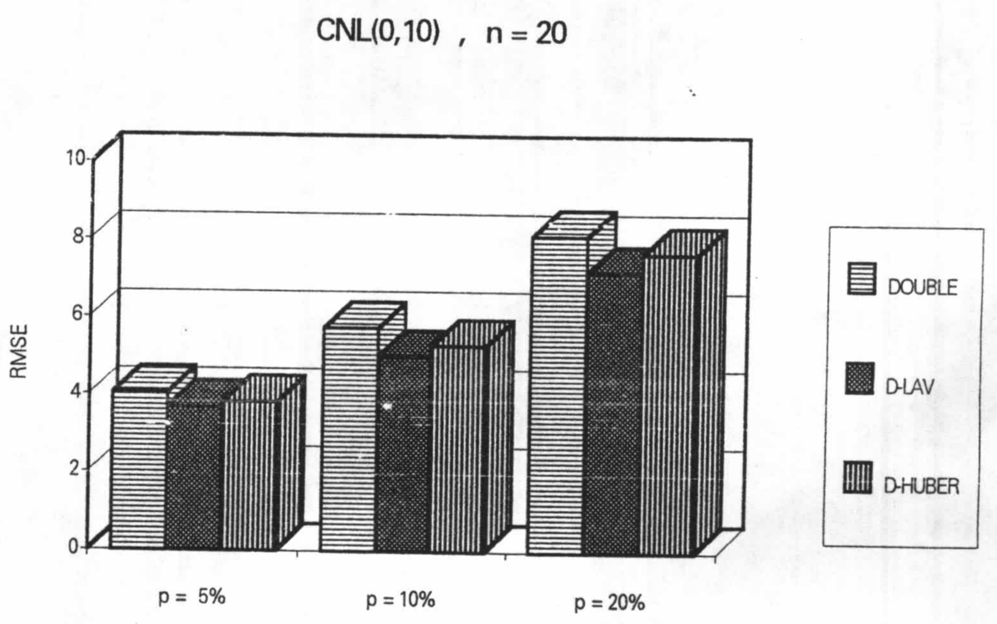
การพยากรณ์สำหรับข้อมูลที่มีลักษณะแนวโน้มเชิงเส้น ความคลาดเคลื่อนมีการแจกแจงปลอมปนที่มีฟังก์ชันการแจกแจง $f(x) = (1-p)N(0,1) + p\ L(0,\beta)$ เมื่อ β เท่ากับ 10 เปอร์เซ็นต์การปลอมเท่ากับ 5, 10 และ 20 ขนาดตัวอย่างที่ใช้ในการศึกษาคือ 10, 20, 30 และ 50 ทำการพยากรณ์ด้วยวิธีการทำให้เรียบแบบเอกซโพเนนเชียลซ้ำสองครั้ง (DOUBLE) การทำให้เรียบแบบเอกซโพเนนเชียลซ้ำสองครั้งเมื่อประมาณค่าด้วยวิธีค่าสัมบูรณ์ต่ำสุด (D--LAV) และการทำให้เรียบแบบเอกซโพเนนเชียลซ้ำสองครั้งเมื่อประมาณค่าด้วยวิธีการของฮูเบอร์ (D-HUBER) ทำการคำนวณค่ารากที่สองของค่าเฉลี่ยของความคลาดเคลื่อนกำลังสอง (RMSE) เฉลี่ย 12 คาบเวลาได้ดังตารางที่ 4.13 และรูปที่ 4.16 และค่า RMSE ของ 12 คาบเวลาได้ในภาคผนวก ข ดังตารางที่ 4.55 - 4.58 และรูปที่ 4.61 - 4.64

ตารางที่ 4.13 แสดงค่า RMSE เฉลี่ย 12 คาบเวลาของวิธีพยากรณ์ทั้ง 3 วิธี ในการวิเคราะห์ข้อมูลอนุกรมเวลาที่มีลักษณะแนวโน้มเชิงเส้น เมื่อความคลาดเคลื่อนมีฟังก์ชันการแจกแจงอยู่ในรูป $f(x) = (1-p)\ N(0,1) + p\ L(0,\beta)$ เมื่อ $\beta = 10$ จำแนกตามขนาดตัวอย่าง (n) และเปอร์เซ็นต์การปลอมปน (p)

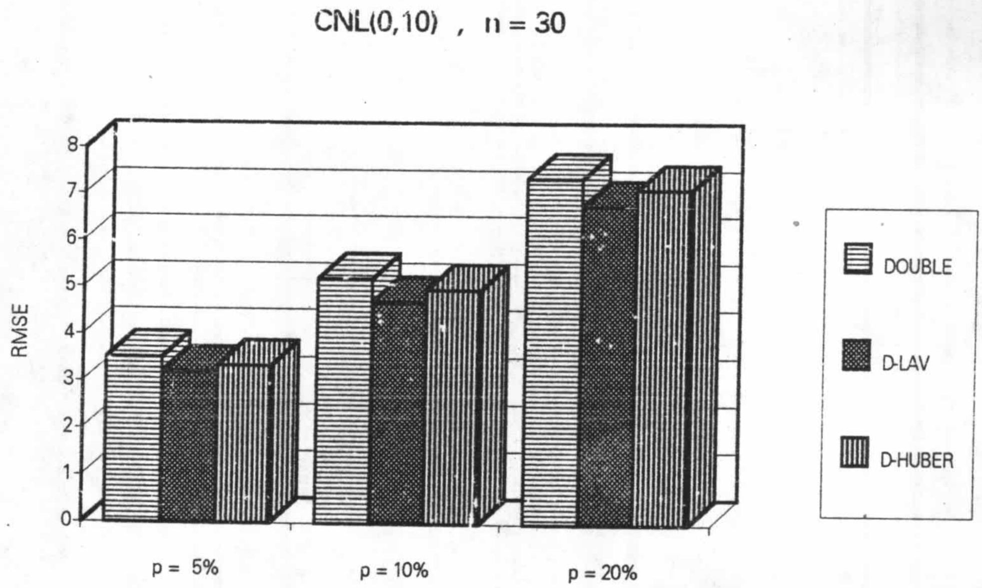
n	p	วิธีพยากรณ์		
		DOUBLE	D-LAV	D-HUBER
10	5	5.0894	3.8593	4.7331
	10	8.9082	6.1627	8.4539
	20	13.6921	11.3029	13.0803
20	5	4.0484	3.6742	3.8201
	10	5.7313	5.0031	5.2924
	20	8.1149	7.1932	7.6709
30	5	3.5194	3.2083	3.3460
	10	5.2001	4.7053	4.9644
	20	7.3394	6.7685	7.1244
50	5	3.3879	3.2149	3.2617
	10	4.3886	4.1315	4.2363
	20	6.5341	6.1680	6.3624

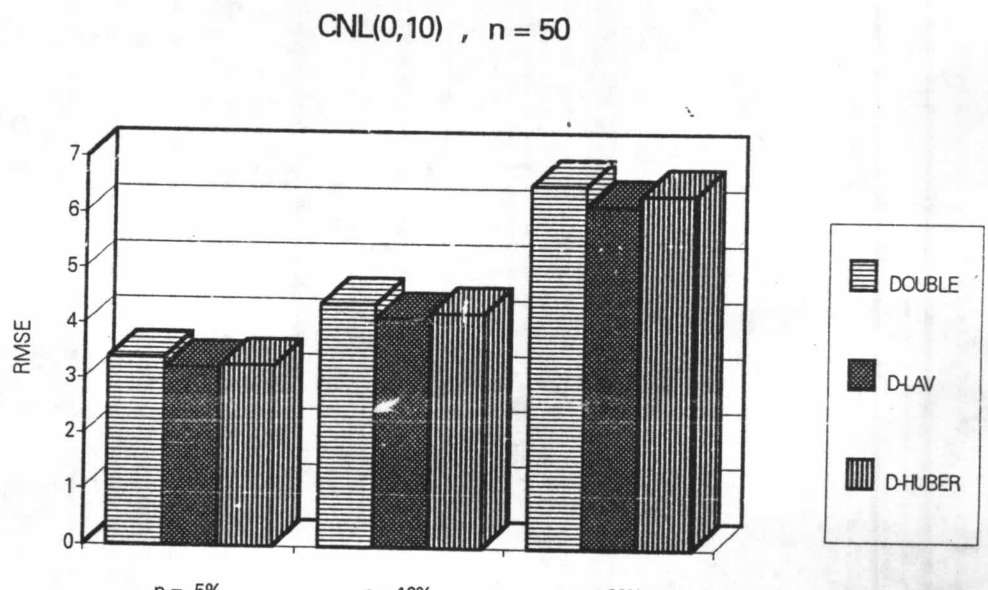
รูปที่ 4.16 การเปรียบเทียบค่า RMSE เฉลี่ย 12 คาบเวลาของวิธีพยากรณ์ทั้ง 3 วิธี ในการวิเคราะห์ข้อมูลอนุกรมเวลาที่มีลักษณะแนวโน้มเชิงเส้น เมื่อความคลาดเคลื่อนมีฟังก์ชันการแจกแจงอยู่ในรูป $f(x) = (1-p)\,N(0,1) + p\,L(0,\beta)$ เมื่อ $\beta = 10$ จำแนกตามขนาดตัวอย่าง (n) และเปอร์เซ็นต์การปลอมปน (p)





รูปที่ 4.16 (ต่อ)


CNL(0,10) , n = 30
RMSE
0
1
2
3
4
5
6
7
8
p = 5%
p = 10%
p = 20%
DOUBLE
D-LAV
D-HUBER


CNL(0,10) , n = 50
RMSE
0
1
2
3
4
5
6
7
p = 5%
p = 10%
p = 20%
DOUBLE
D-LAV
D-HUBER

จากตารางที่ 4.13 และตารางที่ 4.55 - 4.58 ได้ผลดังนี้

ในการวิเคราะห์อนุกรมเวลาที่มีลักษณะแนวโน้มเชิงเส้น เมื่อความคลาดเคลื่อนมีการแจกแจงอยู่ในรูปปลอมปนที่มีฟังก์ชันการแจกแจง f(x) = (1-p) N(0,1) + p L(0,β) เมื่อ β เท่ากับ 10 เปรียบเทียบค่า RMSE พบว่า

วิธี D-LAV ให้ค่า RMSE ของทุกคาบเวลาและโดยเฉลี่ยต่ำสุด ในทุกระดับของเปอร์เซ็นต์การปลอมปน ของทุกขนาดตัวอย่าง

2.6 เมื่อความคลาดเคลื่อนมีฟังก์ชันการแจกแจง f(x) = (1-p)N(0,1) + pU(-5,5)

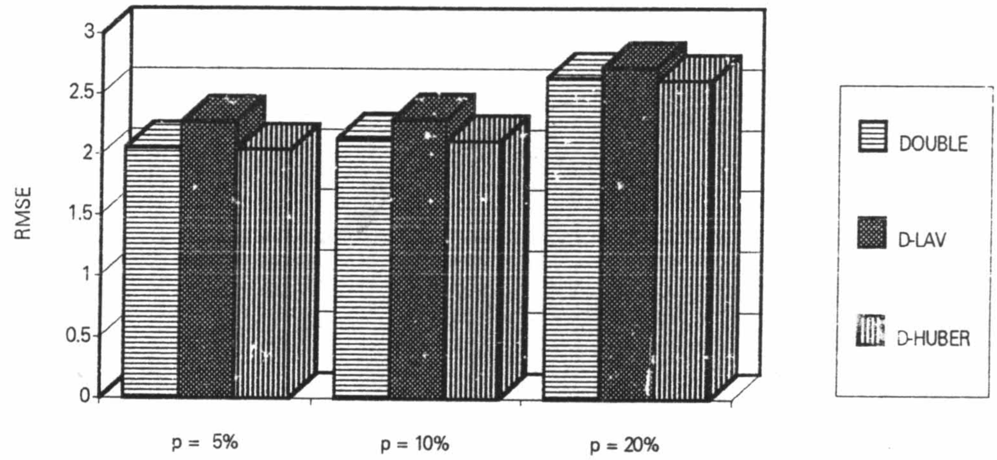
การพยากรณ์สำหรับข้อมูลที่มีลักษณะแนวโน้มเชิงเส้น ความคลาดเคลื่อนมีการแจกแจงปลอมปนที่มีฟังก์ชันการแจกแจง f(x) = (1-p)N(0,1) + pU(-5,5) เปอร์เซ็นต์การปลอม (p) เท่ากับ 5, 10 และ 20 ขนาดตัวอย่างที่ใช้ในการศึกษา คือ 10, 20, 30 และ 50 ทำการพยากรณ์ด้วยวิธีการทำให้เรียบแบบเอกซโพเนนเชียลซ้ำสองครั้ง (DOUBLE) การทำให้เรียบแบบเอกซโพเนนเชียลซ้ำสองครั้งเมื่อประมาณค่าด้วยวิธีค่าสัมบูรณ์ต่ำสุด (D--LAV) และการทำให้เรียบแบบเอกซโพเนนเชียลซ้ำสองครั้งเมื่อประมาณค่าด้วยวิธีการของฮูเบอร์ (D-HUBER) ทำการคำนวณค่ารากที่สองของค่าเฉลี่ยของความคลาดเคลื่อนกำลังสอง (RMSE) เฉลี่ย 12 คาบเวลา ได้ดังตารางที่ 4.14 และรูปที่ 4.17 และค่า RMSE ของ 12 คาบเวลาได้ในภาคผนวก ข ดังตารางที่ 4.59 - 4.62 และรูปที่ 4.65 - 4.68

ตารางที่ 4.14 แสดงค่า RMSE เฉลี่ย 12 คาบเวลาของวิธีพยากรณ์ทั้ง 3 วิธี ในการวิเคราะห์ข้อมูลอนุกรมเวลาที่มีลักษณะแนวโน้มเชิงเส้น เมื่อความคลาดเคลื่อนมีฟังก์ชันการแจกแจงอยู่ในรูป $f(x) = (1-p)\, N(0,1) + p\, U(-5,5)$ จำแนกตามขนาดตัวอย่าง (n) และ เปอร์เซ็นต์การปลอมปน (p)

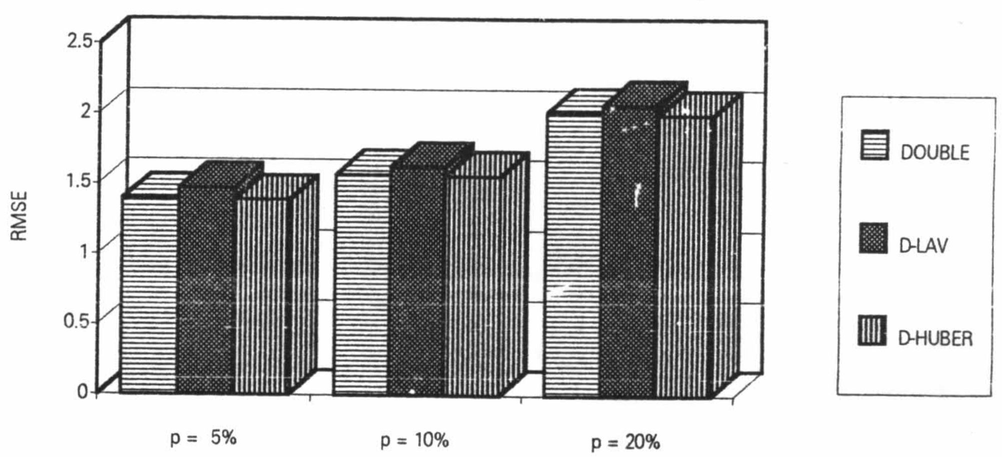
n	p	วิธีพยากรณ์		
		DOUBLE	D-LAV	D-HUBER
10	5	2.0547	2.2695	2.0425
	10	2.1318	2.2890	2.1136
	20	2.6336	2.7243	2.6141
20	5	1.3961	1.4719	1.3877
	10	1.5692	1.6250	1.5573
	20	2.0141	2.0655	1.9960
30	5	1.2919	1.3258	1.2842
	10	1.4603	1.4720	1.4421
	20	1.7227	1.7287	1.7075
50	5	1.2410	1.2498	1.2317
	10	1.4037	1.4061	1.3969
	20	1.6631	1.6590	1.6548

รูปที่ 4.17 การเปรียบเทียบค่า RMSE เฉลี่ย 12 คาบเวลาของวิธีพยากรณ์ทั้ง 3 วิธี ในการวิเคราะห์ข้อมูลอนุกรมเวลาที่มีลักษณะแนวโน้มเชิงเส้น เมื่อความคลาดเคลื่อนมีฟังก์ชันการแจกแจงอยู่ในรูป f(x) = (1-p) N(0,1) + p U(-5,5) จำแนกตามขนาดตัวอย่าง (n) และเปอร์เซ็นต์การปลอมปน (p)

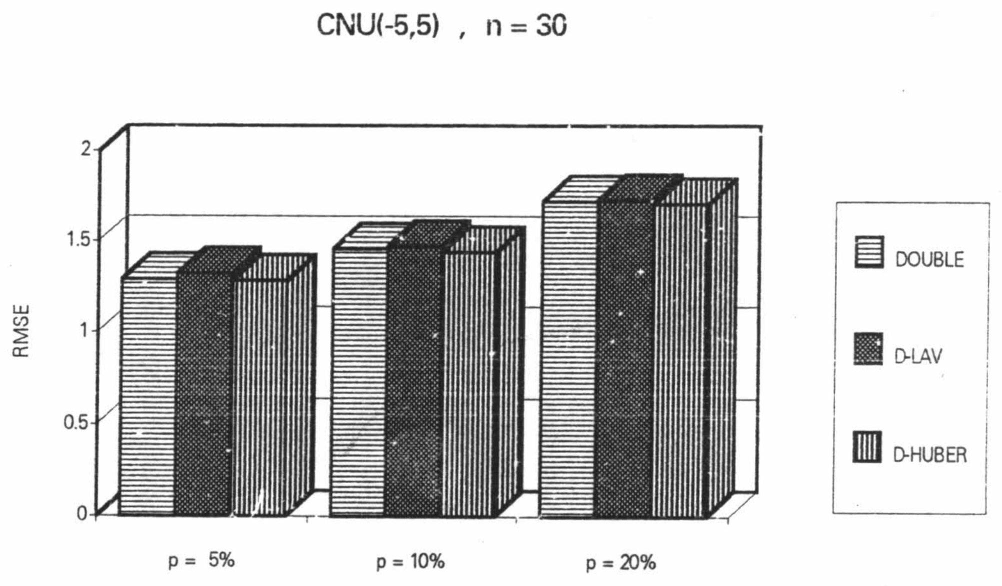
CNU(-5,5) , n = 10

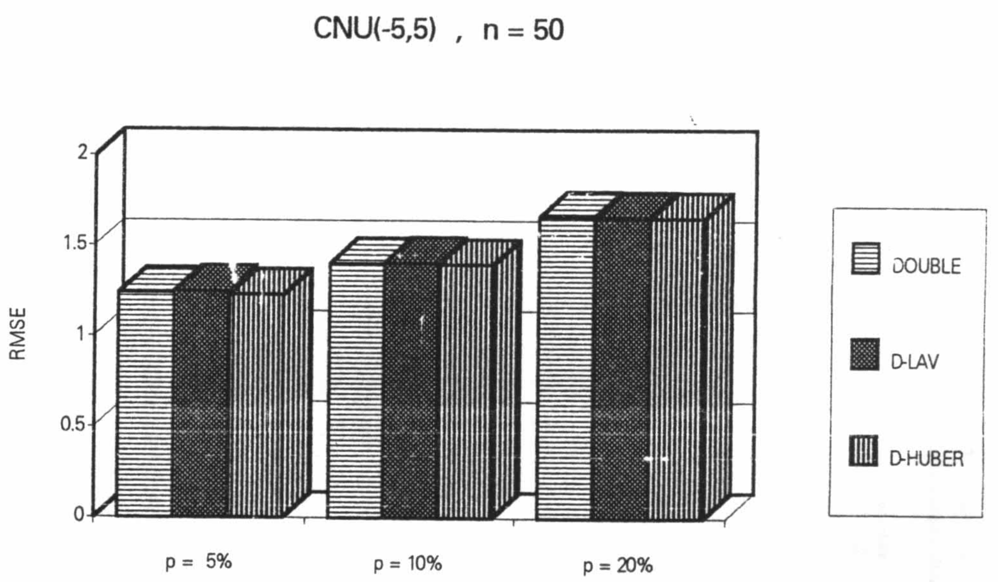


CNU(-5,5) , n = 20



รูปที่ 4.17 (ต่อ)


CNU(-5,5) , n = 30
2
1.5
1
0.5
0
RMSE
p = 5%
p = 10%
p = 20%
DOUBLE
D-LAV
D-HUBER


CNU(-5,5) , n = 50
2
1.5
1
0.5
0
RMSE
p = 5%
p = 10%
p = 20%
DOUBLE
D-LAV
D-HUBER

จากตารางที่ 4.14 และตารางที่ 4.59 - 4.62 ได้ผลดังนี้

ในการวิเคราะห์อนุกรมเวลาที่มีลักษณะแนวโน้มเชิงเส้น เมื่อความคลาดเคลื่อนมีการแจกแจงอยู่ในรูปปลอมปนที่มีฟังก์ชันการแจกแจง f(x) = (1-p) N(0,1) + p U(-5,5) เปรียบ-เทียบค่า RMSE พบว่า

เมื่อขนาดตัวอย่าง 10, 20 และ 30 วิธี D-HUBER ให้ค่า RMSE โดยเฉลี่ยต่ำสุด และให้ค่า RMSE ต่ำสุดในทุกคาบเวลา ในทุกระดับของเปอร์เซ็นต์การปลอมปน

เมื่อขนาดตัวอย่างเท่ากับ 50 วิธี D-HUBER ให้ค่า RMSE โดยเฉลี่ยต่ำสุด ในทุกระดับของเปอร์เซ็นต์การปลอมปน ซึ่งจากการพิจารณาทุกคาบเวลาพยากรณ์พบว่า โดยทั่วไป วิธี D-HUBER ให้ค่า RMSE ต่ำสุด

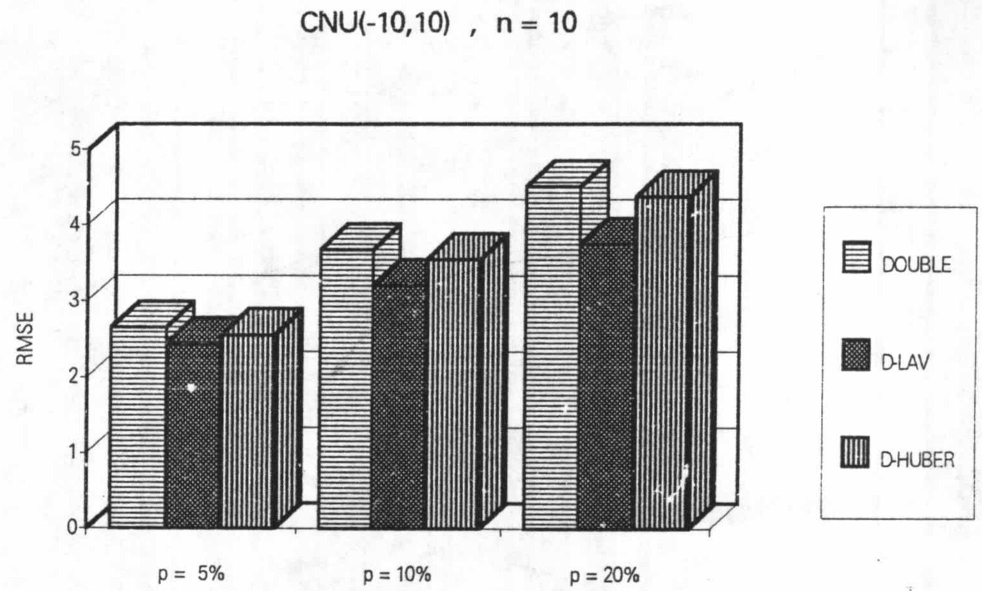
2.7 เมื่อความคลาดเคลื่อนมีฟังก์ชันการแจกแจง f(x) = (1-p)N(0,1) + p U(-10,10)

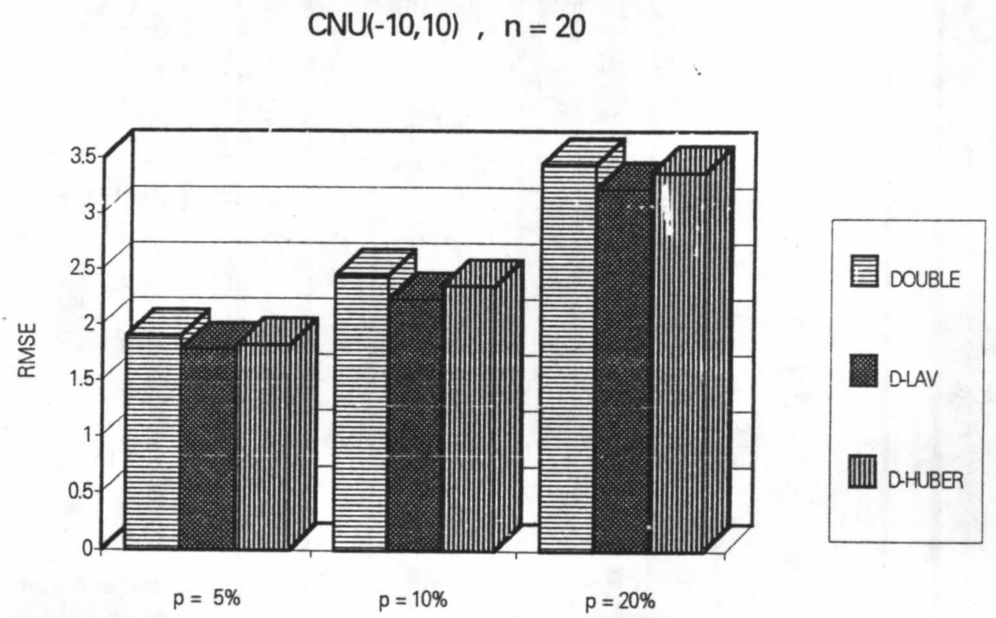
การพยากรณ์สำหรับข้อมูลที่มีลักษณะแนวโน้มเชิงเส้น ความคลาดเคลื่อนมีการแจกแจงปลอมปนที่มีฟังก์ชันการแจกแจง f(x) = (1-p) N(0,1) + p U(-10,10) เปอร์เซ็นต์การปลอม (p) เท่ากับ 5, 10 และ 20 ขนาดตัวอย่างในการศึกษา คือ 10, 20, 30 และ 50 ทำการพยากรณ์ด้วยวิธีการทำให้เรียบแบบเอกซโพเนนเชียลซ้ำสองครั้ง (DOUBLE) การทำให้เรียบแบบเอกซโพเนนเชียลซ้ำสองครั้งเมื่อประมาณค่าด้วยวิธีค่าสัมบูรณ์ต่ำสุด (D--LAV) และการทำให้เรียบแบบเอกซโพเนนเชียลซ้ำสองครั้งเมื่อประมาณค่าด้วยวิธีการของฮูเบอร์ (D-HUBER) ทำการคำนวณค่ารากที่สองของค่าเฉลี่ยของความคลาดเคลื่อนกำลังสอง (RMSE) เฉลี่ย 12 คาบเวลาได้ดังตารางที่ 4.15 และรูปที่ 4.18 และค่า RMSE ของ 12 คาบเวลาได้ในภาคผนวก ข ดังตารางที่ 4.63 - 4.66 และรูปที่ 4.69 - 4.72

ตารางที่ 4.15 แสดงค่า RMSE เฉลี่ย 12 คาบเวลาของวิธีพยากรณ์ทั้ง 3 วิธี ในการวิเคราะห์ข้อมูลอนุกรมเวลาที่มีลักษณะแนวโน้มเชิงเส้น เมื่อความคลาดเคลื่อนมีฟังก์ชันการแจกแจงอยู่ในรูป f(x) = (1-p) N(0,1) + p U(-10,10) จำแนกตามขนาดตัวอย่าง (n) และเปอร์เซ็นต์การปลอมปน (p)

n	p	วิธีพยากรณ์		
		DOUBLE	D-LAV	D-HUBER
10	5	2.6542	2.4384	2.5564
	10	3.6820	3.2116	3.5505
	20	4.5170	3.7514	4.3763
20	5	1.8972	1.7780	1.8205
	10	2.4449	2.2324	2.3538
	20	3.4431	3.2271	3.3721
30	5	1.7613	1.6859	1.7093
	10	2.3166	2.1985	2.2531
	20	3.0775	2.8834	3.0112
50	5	1.6834	1.6493	1.6519
	10	2.1240	2.0500	2.0794
	20	2.8399	2.7665	2.8070

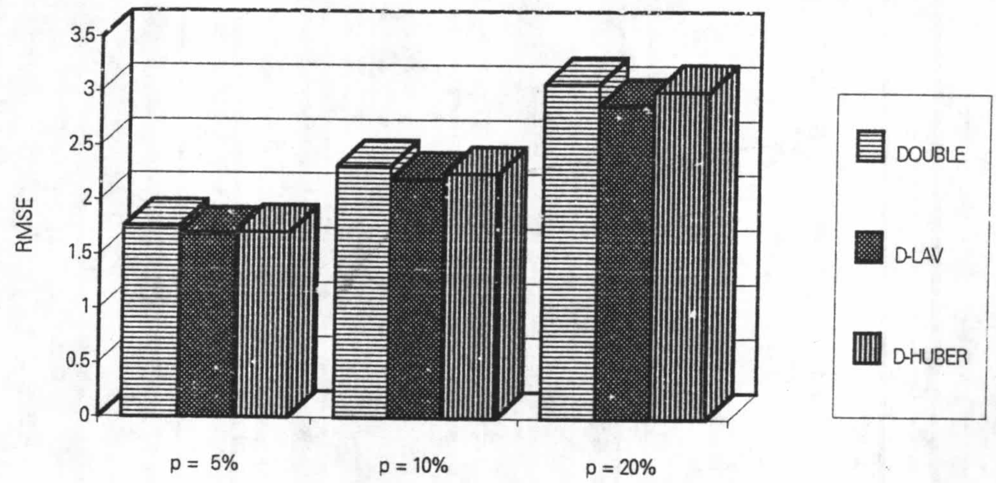
รูปที่ 4.18 การเปรียบเทียบค่า RMSE เฉลี่ย 12 คาบเวลาของวิธีพยากรณ์ทั้ง 3 วิธี ในการวิเคราะห์ข้อมูลอนุกรมเวลาที่มีลักษณะแนวโน้มเชิงเส้น เมื่อความคลาดเคลื่อนมีฟังก์ชันการแจกแจงอยู่ในรูป f(x) = (1-p) N(0,1) + p U(-10,10) จำแนกตามขนาดตัวอย่าง (n) และเปอร์เซ็นต์การปลอมปน (p)



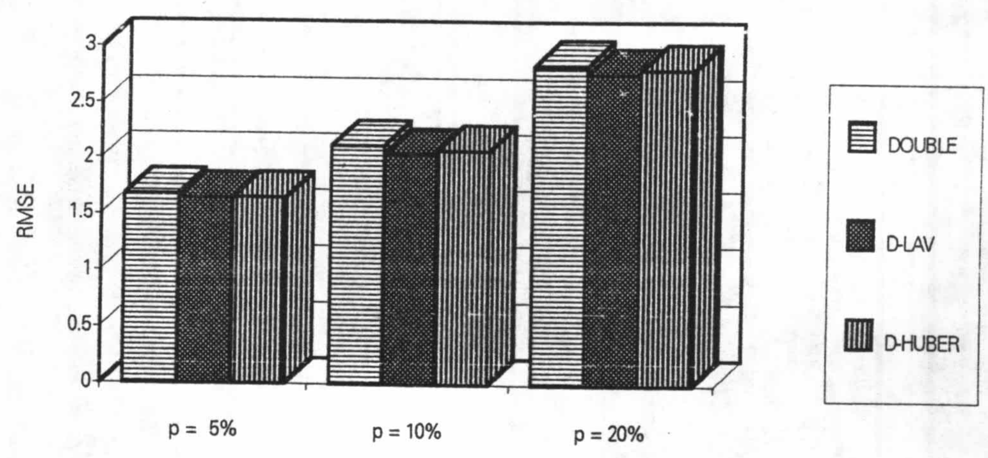


รูปที่ 4.18 (ต่อ)

CNU(-10,10) , n = 30



CNU(-10,10) , n = 50



จากตารางที่ 4.15 และตารางที่ 4.63 - 4.66 ได้ผลดังนี้

ในการวิเคราะห์อนุกรมเวลาที่มีลักษณะแนวโน้มเชิงเส้น เมื่อความคลาดเคลื่อนมีการแจกแจงอยู่ในรูปปลอมปนที่มีฟังก์ชันการแจกแจง f(x) = (1-p) N(0,1) + p U(-10,10) เปรียบ-เทียบค่า RMSE พบว่า

วิธี D-LAV ให้ค่า RMSE ของทุกคาบเวลาและโดยเฉลี่ยต่ำสุด ในทุกระดับของเปอร์เซ็นต์การปลอมปน ของทุกขนาดตัวอย่าง

2.8 สรุปผลการวิจัยสำหรับข้อมูลที่มีลักษณะเคลื่อนไหวอยู่ในระดับค่าเฉลี่ย เมื่อความคลาดเคลื่อนมีฟังก์ชันการแจกแจงแบบต่าง ๆ

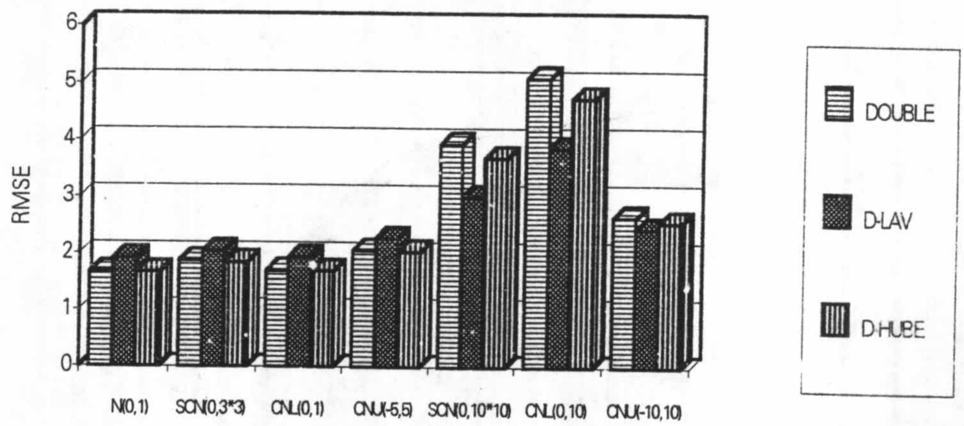
จากการวิเคราะห์ผลข้อมูลที่มีลักษณะแนวโน้มเชิงเส้น เมื่อความคลาด-เคลื่อนมีการแจกแจงแบบต่าง ๆ ซึ่งสามารถสรุปผลการเปรียบเทียบค่า RMSE โดยเฉลี่ย 12 คาบเวลาของวิธีพยากรณ์ทั้ง 3 วิธี เปอร์เซ็นต์การปลอมปนเท่ากับ 5, 10 และ 20 และขนาดตัวอย่างเท่ากับ 10, 20, 30 และ 50 ได้ดังตารางที่ 4.16 และรูปที่ 4.19 - 4.22

ตารางที่ **4.16** แสดงค่า RMSE เฉลี่ย 12 คาบเวลาพยากรณ์ของวิธีพยากรณ์ทั้ง 3 วิธี ในการวิเคราะห์อนุกรมเวลาที่มีลักษณะแนวโน้มเชิงเส้น จำแนกตามขนาดตัวอย่าง (n) เปอร์เซ็นต์การปลอมปน (p) และการแจกแจงของความคลาดเคลื่อน

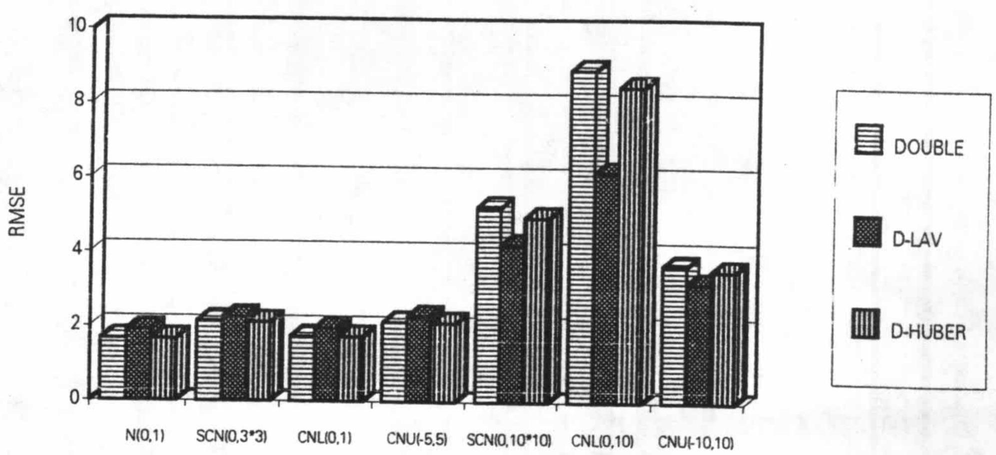
n	p	การแจกแจงของความคลาดเคลื่อน																	
		SCN(0,3*3)			CNL(0,1)			CNU(-5,5)			SCN(0,10*10)			CNL(0,10)			CNU(-10,10)		
		DOUBLE	D-LAV	D-HUBER	DOUBLE	D-LAV	D-HUBER	DOUBLE	D-LAV	D-HUBER	DOUBLE	D-LAV	D-HUBER	DOUBLE	D-LAV	D-HUBER	DOUBLE	D-LAV	D-HUBER
10	5	1.8765	2.0471	1.8472	1.6959	1.9363	1.6935	2.0547	2.2695	2.0425	3.9348	3.0043	1.8472	5.0894	3.8593	4.7331	2.6542	2.4384	2.5564
	10	2.1596	2.2546	2.1137	1.7478	1.9583	1.7455	2.1318	2.2890	2.1136	5.2256	4.2217	4.9497	8.9082	6.1627	8.4539	3.6820	3.2116	3.5505
	20	2.7447	2.7474	2.6884	1.8504	2.1131	1.8453	2.6336	2.7243	2.6141	7.2030	5.5035	6.8696	13.6921	11.3029	13.0803	4.5170	3.7514	4.3763
20	5	1.4247	1.4799	1.4094	1.2236	1.3149	1.2201	1.3961	1.4719	1.3877	3.0177	2.7636	2.8857	4.0484	3.6742	3.8201	1.8972	1.7780	1.8205
	10	1.6565	1.6943	1.6314	1.2606	1.3306	1.2551	1.5692	1.6250	1.5573	4.2699	3.7011	4.0412	5.7313	5.0031	5.2924	2.4449	2.2324	2.3538
	20	1.9318	1.9150	1.8962	1.3279	1.4114	1.3221	2.0141	2.0655	1.9960	5.6393	5.1145	5.2682	8.1149	7.1932	7.6709	3.4431	3.2271	3.3721
30	5	1.3184	1.3326	1.3042	1.1499	1.1946	1.1493	1.2919	1.3258	1.2842	2.8369	2.6769	2.7198	3.5194	3.2083	3.3460	1.7613	1.6859	1.7093
	10	1.4819	1.4827	1.4616	1.1625	1.2104	1.1582	1.4603	1.4720	1.4421	3.6834	3.3797	3.5535	5.2001	4.7053	4.9644	2.3166	2.1985	2.2531
	20	1.7709	1.7490	1.7398	1.2340	1.2602	1.2285	1.7227	1.7287	1.7075	5.1036	4.6953	4.9679	7.3394	6.7685	7.1244	3.0775	2.8834	3.0112
50	5	1.2497	1.2668	1.2405	1.0781	1.1035	1.0781	1.2410	1.2498	1.2317	2.5790	2.4814	2.5179	3.3879	3.2149	3.2617	1.6834	1.6493	1.6519
	10	1.4469	1.4347	1.4324	1.1253	1.1497	1.1241	1.4037	1.4061	1.3969	3.4962	3.3593	3.4159	4.3886	4.1315	4.2363	2.1240	2.0500	2.0794
	20	1.6475	1.6386	1.6327	1.1554	1.1660	1.1516	1.6631	1.6590	1.6548	4.8731	4.6256	4.7725	6.5341	6.1680	6.3624	2.8399	2.7665	2.8070

รูปที่ 4.19 แสดงค่า RMSE เฉลี่ย 12 คาบเวลาพยากรณ์ของวิธีทั้ง 3 วิธี ในการวิเคราะห์ข้อมูลอนุกรมเวลาที่มีลักษณะแนวโน้มเชิงเส้น ขนาดตัวอย่าง (n) เท่ากับ 10 จำแนกตามเปอร์เซ็นต์การปลอมปน (p) เมื่อความคลาดเคลื่อนมีการแจกแจงปลอมปนแบบต่าง ๆ เปรียบเทียบกับเมื่อมีการแจกแจงปกติมาตรฐาน

n = 10, p = 5%

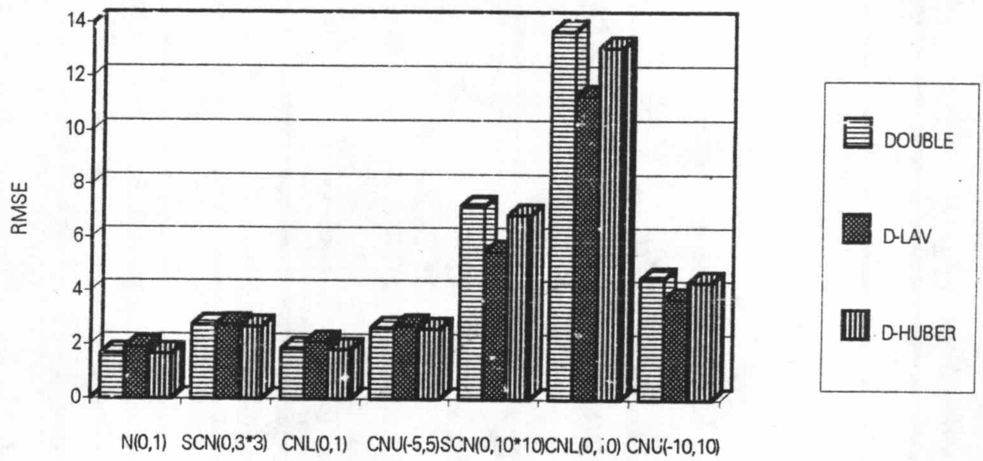


n = 10, p = 10%



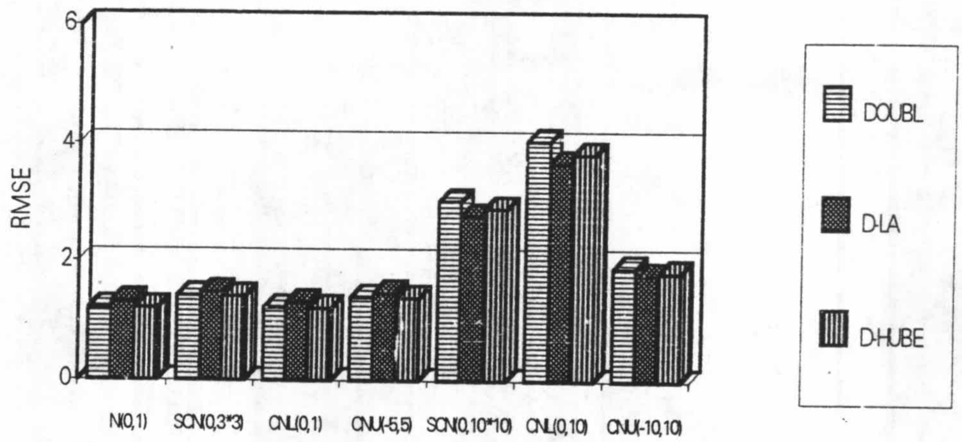
รูปที่ 4.19 (ต่อ)

n = 10, p = 20%

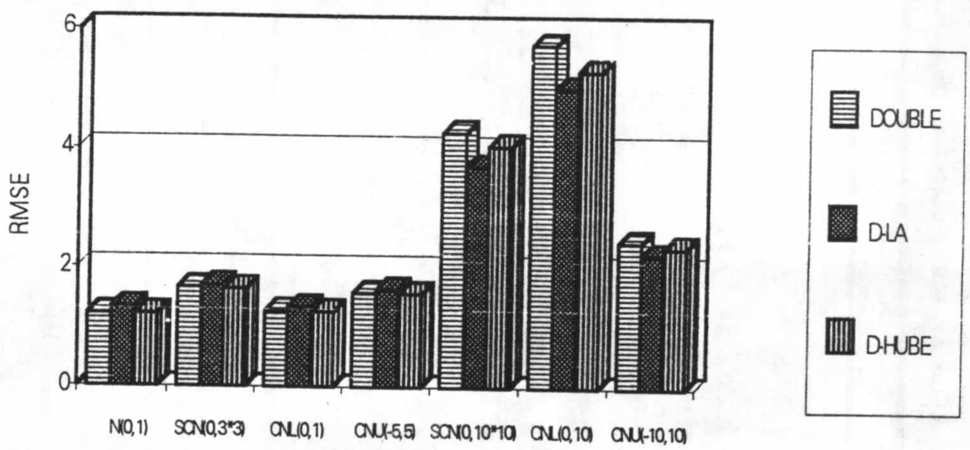


รูปที่ 4.20 แสดงค่า RMSE เฉลี่ย 12 คาบเวลาพยากรณ์ของวิธีทั้ง 3 วิธี ในการวิเคราะห์ข้อมูลอนุกรมเวลาที่มีลักษณะแนวโน้มเชิงเส้น ขนาดตัวอย่าง (n) เท่ากับ 20 จำแนกตามเปอร์เซ็นต์การปลอมปน (p) เมื่อความคลาดเคลื่อนมีการแจกแจงปลอมปนแบบต่าง ๆ เปรียบเทียบกับเมื่อมีการแจกแจงปกติมาตรฐาน

n = 20, p = 5%

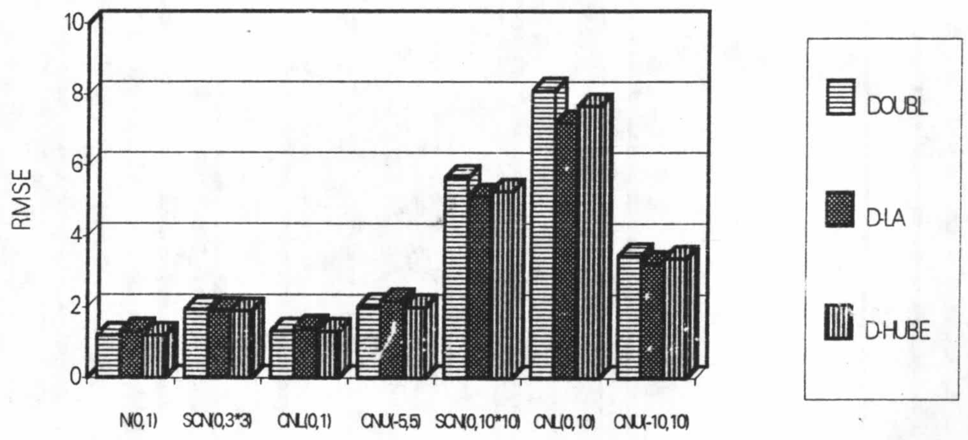


n = 20, p = 10%

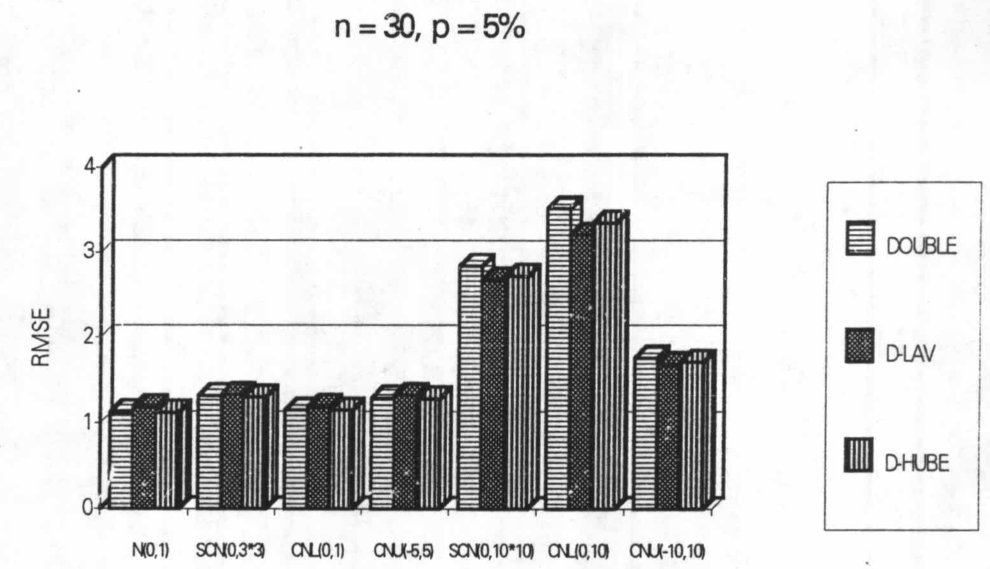


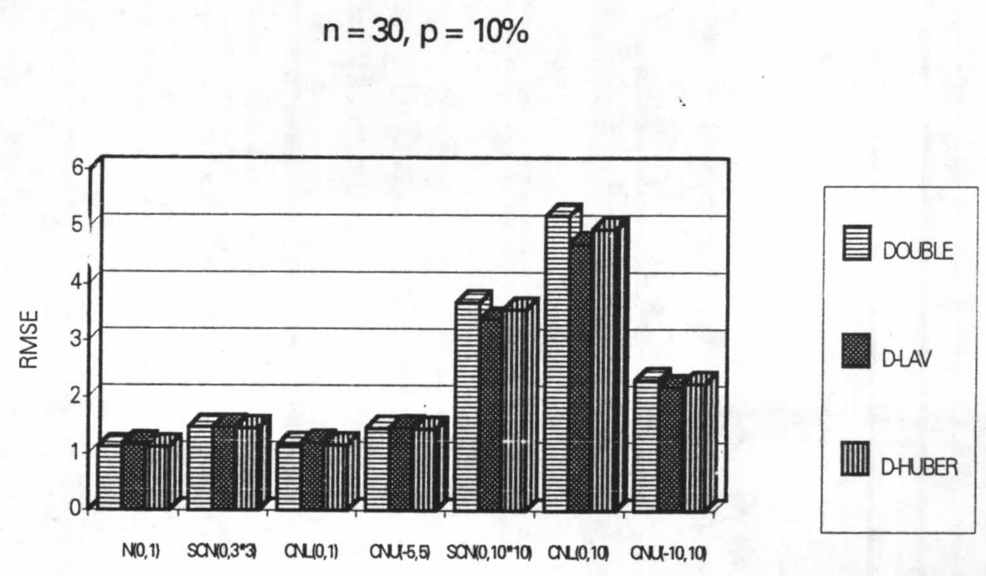
รูปที่ 4.20 (ต่อ)

n = 20, p = 20%



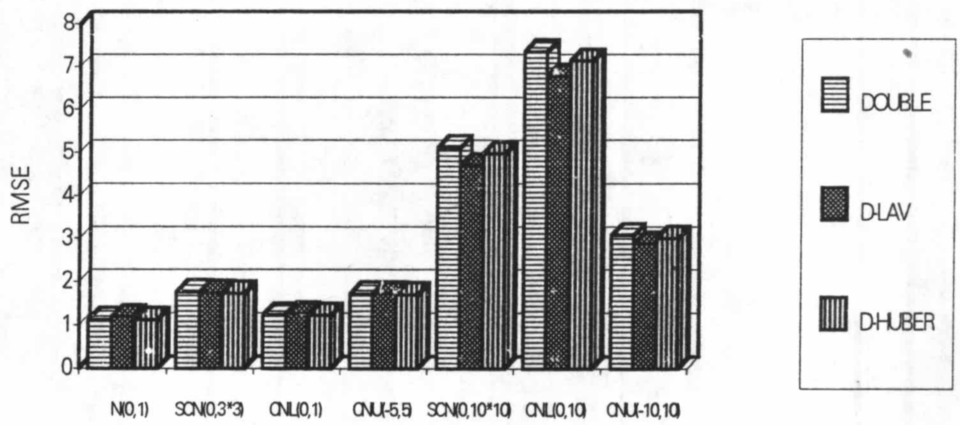
รูปที่ 4.21 แสดงค่า RMSE เฉลี่ย 12 คาบเวลาพยากรณ์ของวิธีทั้ง 3 วิธี ในการวิเคราะห์ข้อมูลอนุกรมเวลาที่มีลักษณะแนวโน้มเชิงเส้น เมื่อขนาดตัวอย่าง (n) เท่ากับ 30 จำแนกตามเปอร์เซ็นต์การปลอมปน (p) เมื่อความคลาดเคลื่อนมีการแจกแจงปลอมปนแบบต่าง ๆ เปรียบเทียบกับเมื่อมีการแจกแจงปกติมาตรฐาน



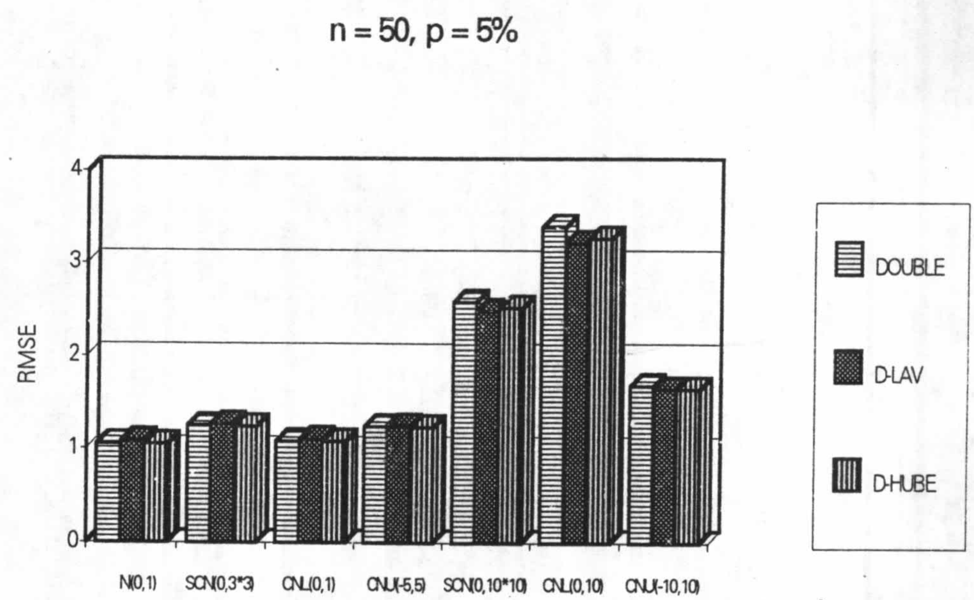


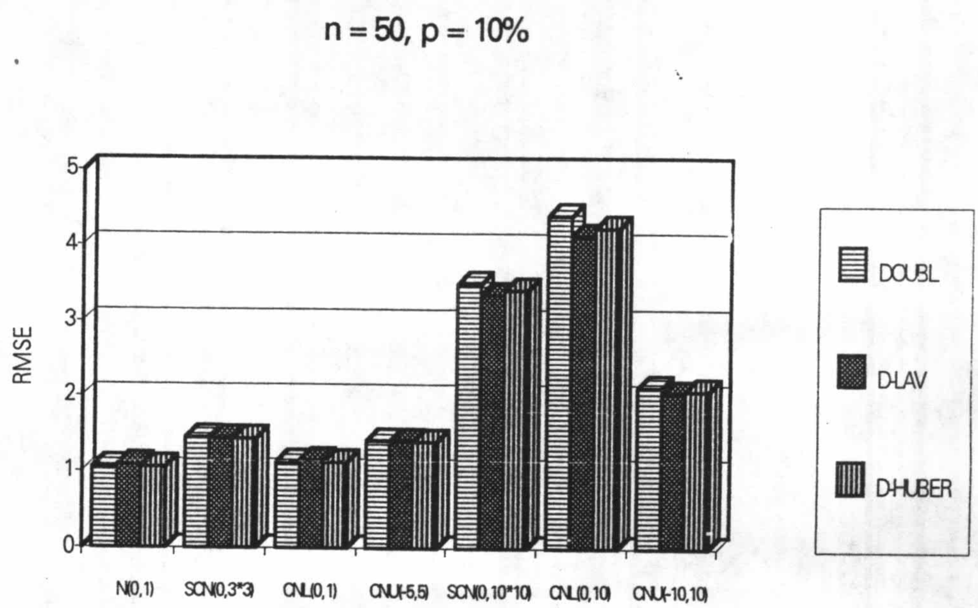
รูปที่ 4.21 (ต่อ)

n = 30, p = 20%

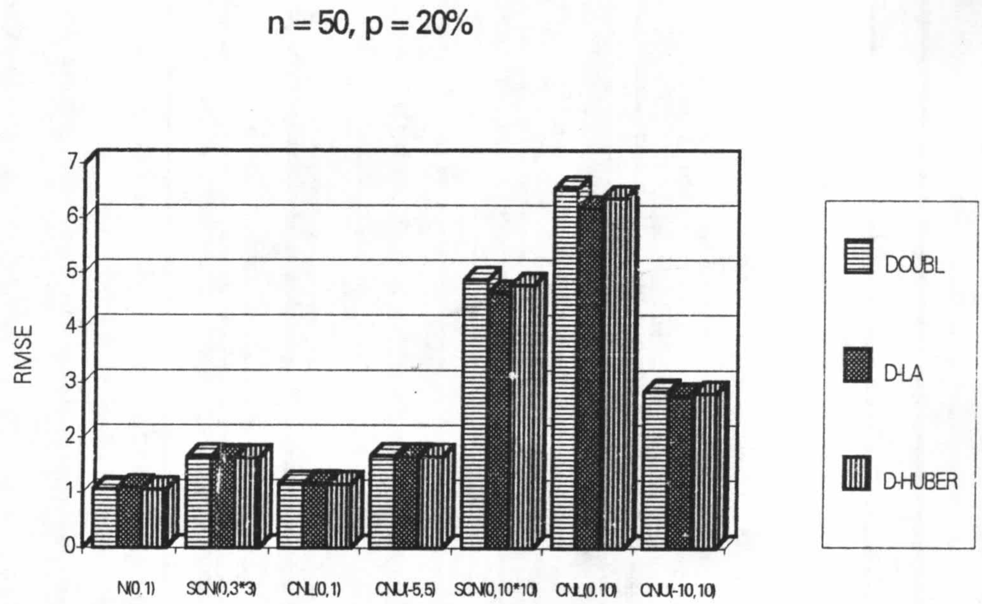


รูปที่ 4.22 แสดงค่า RMSE เฉลี่ย 12 คาบเวลาพยากรณ์ของวิธีทั้ง 3 วิธี ในการวิเคราะห์ข้อมูลอนุกรมเวลาที่มีลักษณะแนวโน้มเชิงเส้น ขนาดตัวอย่าง (n) เท่ากับ 50 จำแนกตามเปอร์เซ็นต์การปลอมปน (p) เมื่อความคลาดเคลื่อนมีการแจกแจงปลอมปนแบบต่าง ๆ เปรียบเทียบกับเมื่อมีการแจกแจงปกติมาตรฐาน





รูปที่ 4.22 (ต่อ)


n = 50, p = 20%
RMSE
0
1
2
3
4
5
6
7
N(0.1)
SCN(0,3*3)
CNL(0,1)
CNU(-5,5)
SCN(0,10*10)
CNL(0,10)
CNU(-10,10)
DOUBL
D-LA
D-HUBER

จากตารางที่ 4.16 ผลการศึกษาเปรียบเทียบวิธีพยากรณ์ 3 วิธี ได้แก่ การทำให้เรียบแบบเอกซโพเนนเชียลซ้ำสองครั้ง การทำให้เรียบแบบเอกซโพเนนเชียลซ้ำสองครั้งเมื่อประมาณค่าด้วยวิค่าสัมบูรณ์ต่ำสุด และการทำให้เรียบแบบเอกซโพเนนเชียลซ้ำสองครั้งเดียวเมื่อประมาณค่าด้วยวิธีการของฮูเบอร์ เมื่อความคลาดเคลื่อนมีการแจกแจงแบบต่าง ๆ ซึ่งสามารถสรุปผลได้ดังนี้

(1) เมื่อมีความคลาดเคลื่อนมีการแจกแจงมาตรฐาน ซึ่งเป็นกรณีที่ไม่ปรากฏข้อมูลผิดปกติในข้อมูลอนุกรมเวลา วิธีการทำให้เรียบแบบเอกซโพเนนเชียลซ้ำสองครั้งเมื่อประมาณค่าด้วยวิธีการของฮูเบอร์ ให้ค่าความคลาดเคลื่อนโดยเฉลี่ยต่ำสุด ซึ่งใกล้เคียงกับเมื่อใช้วิธีการทำให้เรียบแบบเอกซโพเนนเชียลซ้ำสองครั้ง ในทุกขนาดตัวอย่าง (n = 10, 20, 30 และ 50) และเมื่อพิจารณาค่าความคลาดเคลื่อนเมื่อขนาดตัวอย่างเพิ่มขึ้น พบว่าค่าความคลาดเคลื่อนของวิธีพยากรณ์ทั้ง 3 วิธีลดลง

(2) เมื่อความคลาดเคลื่อนมีการแจกแจงปกติปลอมปน $f(x) = (1\text{-}p)\ N(0,1) + p\ N(0,c^2\sigma^2)$ เมื่อ c เท่ากับ 3 และ σ เท่ากับ 1 $f(x) = (1\text{-}p)\ N(0,1) + p\ L(0,\beta)$ เมื่อ β เท่ากับ 1 และ $f(x) = (1\text{-}p)\ N(0,1) + p\ U(-5,5)$ ซึ่งกำหนดค่าพารามิเตอร์ของการแจกแจงปลอมปนให้เกิดค่าของข้อมูลผิดปกติไม่เด่นชัดในข้อมูลอนุกรมเวลา พบว่า วิธีการทำให้เรียบแบบเอกซโพเนนเชียลซ้ำสองครั้งเมื่อประมาณค่าด้วยวิธีการของฮูเบอร์ ให้ค่าความคลาดเคลื่อนในการพยากรณ์โดยทั่วไปต่ำสุดในทุกกรณี และเมื่อพิจารณาค่าความคลาดเคลื่อน ณ. ขนาดตัวอย่างหนึ่ง ๆ พบว่าความคลาดเคลื่อนของวิธีพยากรณ์ทั้ง 3 วิธีเพิ่มขึ้นเมื่อเปอร์เซ็นต์การปลอมปนเพิ่มขึ้น และ ณ ระดับการปลอมปนหนึ่ง ๆ ค่าความคลาดเคลื่อนของวิธีพยากรณ์ทั้ง 3 วิธีลดลงเมื่อขนาดตัวอย่างเพิ่มขึ้น

(3) เมื่อความคลาดเคลื่อนมีการแจกแจงปกติปลอมปน $f(x) = (1\text{-}p)\ N(0,1) + p\ N(0,c^2\sigma^2)$ เมื่อ c เท่ากับ 10 และ σ เท่ากับ 1 $f(x) = (1\text{-}p)\ N(0,1) + p\ L(0,\beta)$ เมื่อ β เท่ากับ 10 และ $f(x) = (1\text{-}p)\ N(0,1) + p\ U(-10,10)$ ซึ่งกำหนดค่าพารามิเตอร์ของการแจกแจงปลอมปนให้เกิดค่าของข้อมูลผิดปกติเด่นชัด พบว่าวิธีการทำให้เรียบแบบเอกซโพเนนเชียลซ้ำสองครั้งเมื่อประมาณค่าด้วยวิธีค่าสัมบูรณ์ต่ำสุด ให้ค่าความคลาดเคลื่อนโดยเฉลี่ยในการพยากรณ์ต่ำสุดในทุกกรณี เมื่อพิจารณาค่าความคลาดเคลื่อน ณ. ขนาดตัวอย่างหนึ่ง ๆ แล้ว พบว่าความคลาดเคลื่อนของวิธีพยากรณ์ทั้ง 3 วิธีเพิ่มขึ้นเมื่อเปอร์เซ็นต์การปลอมปนเพิ่มขึ้น และเมื่อพิจารณา ณ ระดับการปลอมปนหนึ่ง ๆ เมื่อขนาดตัวอย่างเพิ่มขึ้นค่าความคลาดเคลื่อนของวิธีพยากรณ์ทั้ง 3 วิธีลดลง