สำนักหอสมุดกลาง พระจอมเกล้าลาดกระบัง

การพยากรณ์การใช้งานสตอเรจโดยใช้วิธีของอนุกรมเวลา

FORECASTING DISK STORAGE USING TIME SERIES METHOD

T131399

โดย

มงคล จันทร์เทศ

MONGKOL CHANTHADE

สจ.
ม 114 ก
2555

เลขหมู่.....
เลขทะเบียน 131399
วัน,เดือน,ปี 2 มิ.ย. 2557

b.1160998x
i.....

อาจารย์ที่ปรึกษา

ผศ.ดร.กันต์พงษ์ วรรัตน์ปัญญา

รายงานนี้เป็นส่วนหนึ่งของวิชาการศึกษาอิสระ 2

หลักสูตรวิทยาศาสตรมหาบัณฑิต สาขาวิชาเทคโนโลยีสารสนเทศ

คณะเทคโนโลยีสารสนเทศ

สถาบันเทคโนโลยีพระจอมเกล้าเจ้าคุณทหารลาดกระบัง

ภาคเรียนที่ 2 ปีการศึกษา 2555

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

FORECASTING DISK STORAGE USING TIME SERIES METHOD

MONGKOL CHANTHADE

A REPORT SUBMITTED IN PARTIAL FULFILLMENT OF THE

REQUIREMENTS OF THE COURSE

INDEPENDENT STUDY 2

MASTER OF SCIENCE PROGRAM IN INFORMATION TECHNOLOGY

FACULTY OF INFORMATION TECHNOLOGY

KING MONGKUT'S INSTITUTE OF TECHNOLOGY LADKRABANG

2/ 2012

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

COPYRIGHT 2013

FACULTY OF INFORMATION TECHNOLOGY

KING MONGKUT'S INSTITUTE OF TECHNOLOGY LADKRABANG

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

หัวข้อ	การพยากรณ์การใช้งานสตอเรจโดยใช้วิธีของอนุกรมเวลา
นักศึกษา	มงคล จันทร์เทศ
รหัสนักศึกษา	54660797
ปริญญา	วิทยาศาสตรมหาบัณฑิต
สาขาวิชา	เทคโนโลยีสารสนเทศ
แขนงวิชา	วิทยาการสารสนเทศและการจัดการ
ปีการศึกษา	2555
อาจารย์ที่ปรึกษา	ผศ.ดร.กันต์พงษ์ วรรัตน์ปัญญา

บทคัดย่อ

เนื่องด้วยข้อมูลเป็นสิ่งสำคัญของแต่ละองค์กร ซึ่งข้อมูลเหล่านี้จำเป็นต้องถูกจัดเก็บลงบนสื่อกลางที่เรียกว่าสตอเรจ แล้วแอ็ปพริเคชั่นต่างๆ ทั้งที่อยู่บนเครื่องแม่ข่ายหรือ เครื่องลูกข่ายเข้ามาเรียกใช้งาน ทำให้จำนวนข้อมูลเพิ่มขึ้นอย่างรวดเร็ว จากเหตผลดังกล่าวทุกองค์กรต่างตระหนักถึงปัญหาต่าง ๆ ที่จะตามมาเมื่อจำนวนของความต้องการมีมากขึ้น ทำให้จำนวนพื้นที่ในการจัดเก็บข้อมูลมีไม่เพียงพอต่อการใช้งาน ซึ่งอาจส่งผลกระทบต่อองค์กร โดยมิอาจประเมิณค่าความเสียหายได้ ด้วยเหตนี้ผู้เขียนจึงมีแนวความคิดที่จะนำค่าสังเกตจากการใช้งานจริงในอดิต มาทำการวิเคาะห์ ด้วยหลักการทางสถิติโดยใช้เครื่องมือการวิเคราะห์ข้อมูล ด้วยวิธีของอนุกรมเวลา เข้ามาวิเคราะห์ข้อมูลที่สังเกตได้และนำผลลัพธ์ที่ได้มาพยากรณ์ว่าในอนาคตจะใช้ขนาดของพื้นที่สตอเรจเท่าไรถึงจะเพียงพอต่อความต้องการ และมีปัจจัยใดบ้างที่มีผลต่อการเรียกใช้งาน เพื่อใช้เป็นข้อมูลให้ทางฝ่ายบริหารได้วางแผนรับมือได้ทันเวลา จากนั้นนำผลลัพธ์ที่ได้มาจัดทำเป็นโมเดลต้นแบบ และนำไปจัดทำเป็นโปรแกรมสำเร็จรูปเพื่อใช้งานในการด้านสตอเรจต่อไป

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

Title	FORECASTING DISK STORAGE USING TIME SERIES METHOD
Student	Mr.Mongkol Chanthade
Student ID	54660797
Degree	Master of Science
Program	Information Technology
Major	Information Technology Management
Academic Year	2012
Advisor	Asst.Prof.Dr. Kuntpong Wararatpanya

ABSTRACT

In the digital age, information becomes a heart of organizations and needs to be collected on various application storages, either on servers or clients, for organizational operations. However, with rapidly increasing information, it leads to insufficient storage management and is impact on organizational operations. The efficient way to solve this problem is the use of the prediction technique. With the actual use of storage spaces in the past, the demand forecasting of prospective storage capacities is estimated by using time series method. Based on the simulation, prototype program can predict the prospective storage capacities at low error rate. This proves that the proposed method can be implemented in decision support systems for monitoring, planning, budgeting, and managing processes.

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

กิตติกรรมประกาศ

โครงการศึกษาการพยากรณ์การใช้งานสตอเรจโดยใช้วิธีของอนุกรมเวลาจะสำเร็จลุล่วงไปไม่ได้เลย ถ้าไม่ได้รับการช่วยเหลือและแรงสนับสนุนจากบุคคลสำคัญหลายท่าน ดังต่อไปนี้

ข้าพเจ้าขอกราบขอบพระคุณ ผศ.ดร.กันต์พงษ์ วรรัตน์ปัญญา ซึ่งเป็นอาจารย์ที่ปรึกษาโครงการนี้ ที่ให้ความกรุณาให้คำแนะนำ และปรึกษา ข้าพเจ้ารู้สึกซาบซึ้งและขอขอบพระคุณเป็นอย่างสูง

ขอกราบพระคุณคณาจารย์ของสถาบันเทคโนโลยีพระจอมเกล้าเจ้าคุณทหารลาดกระบังทุก ๆ ท่านที่ได้ประสิทธิ์ประสาทวิชาให้กับข้าพเจ้า

ขอขอบคุณเพื่อน ๆ พี่ ๆ น้อง ๆ ทุกคนที่ให้คำแนะนำ และคอยให้กำลังใจเสมอมา

สุดท้ายนี้ข้าพเจ้าขอกราบขอบพระคุณ บิดา มารดา และครอบครัวของข้าพเจ้าที่เป็นกำลังใจ และให้การสนับสนุนในทุกเรื่องๆ ทำให้ข้าพเจ้าสามารถทำโครงการพัฒนาระบบงานฉบับนี้สำเร็จลุล่วงด้วยดี

คุณค่าและประโยชน์จากโครงการพัฒนาระบบงานฉบับนี้ ข้าพเจ้าขอมอบแด่ผู้มีพระคุณทุกท่าน

มงคล จันทร์เทศ

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

สารบัญ

หน้า

บทคัดย่อภาษาไทย I

บทคัดย่อภาษาอังกฤษ II

กิตติกรรมประกาศ III

สารบัญ IV

สารบัญตาราง VII

สารบัญรูป VIII

บทที่ 1 บทนำ 1

1.1 ความเป็นมา 1

1.2 วัตถุประสงค์ 2

1.3 ขอบเขตการดำเนินงาน 2

1.4 ขั้นตอนการศึกษา 2

1.5 ประโยชน์ที่คาดว่าจะได้รับ 3

1.6 เครื่องมือที่ใช้ในการพัฒนาระบบ 3

บทที่ 2 ทฤษฎีที่เกี่ยวข้อง 4

2.1 สตอเรจ เทคโนโลยี 4

2.1.1 DAS (Direct Attach Storage) 4

2.1.2 SAN (Storage Area Network) 4

2.1.3 NAS (Network Attach Storage) 5

2.2 RAID เทคโลโลยี 5

2.2.1 RAID พื้นฐาน 6

2.2.2 RAID ลูกผสม 6

2.2.3 RAID Proprietary 7

2.3 ปัจจัยที่มีผลต่อการใช้งานของการจัดเก็บข้อมูล 7

2.3.1 ประเภทของ Hard Disk 8

2.3.2 ประสิทธิ์ภาพของการจัดเก็บข้อมูล 8

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

สารบัญ (ต่อ)

หน้า

2.4 อนุกรมเวลาและการพยากรณ์ 10
2.4.1 อนุกรมเวลา (Time series) 10
2.4.2 การเลือกวิธีการพยากรณ์ 11
2.4.3 การเปรียบเทียบแบบจำลองที่เหมาะ 13
2.4.4 แบบจำลองการทำให้เรียบแบบง่ายด้วยวิธีค่าเฉลี่ยเคลื่อนที่ 13
2.4.5 แบบจำลองการปรับเรียบแบบเอ็กซ์โปเนนเชียล 16
2.5 ทบทวนวรรณกรรม 18

บทที่ 3 วิธีการดำเนินงาน 19
3.1 ขั้นตอนการคัดเลือกข้อมูล 19
3.2 ขั้นตอนการเตรียมข้อมูล 20
3.3 ขั้นตอนการดำเนินงาน 21
3.3.1 แบบจำลองการทำให้เรียบแบบง่ายด้วยวิธีค่าเฉลี่ยเคลื่อนที่ 22
3.3.2 แบบจำลองการปรับเรียบแบบเอ็กซ์โปเนนเชียล 24

บทที่ 4 ผลการทดลอง 27
4.1 ผลการทดสอบ อนุกรมเวลา ตามทฤษฎี 27
4.1.1 การพยากรณ์ด้วยแบบจำลอง Moving Average Method (MV) 27
4.1.2 การพยากรณ์ด้วยแบบจำลอง Single Exponential Smoothing (SES) 28
4.2 ผลการทดสอบ อนุกรมเวลา 29
4.3 ตัวอย่างหน้าต่างโปรแกรม และการใช้งาน 30
4.3.1 ใช้สำหรับการป้อนข้อมูลพื้นฐานของการใช้งานสตอเรจ 31
4.3.2 ใช้สำหรับแสดงผลลัพธ์ของการประมวลผล 32
4.4 ผลการวิเคราะห์ด้วยโปรแกรมพยากรณ์ 33

บทที่ 5 สรุปผลการทดลอง 36
5.1 สรุปผลการศึกษา 36
5.2 ประโยชน์ที่ได้รับจากการศึกษาและพัฒนาระบบ 36

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

สารบัญ (ต่อ)

หน้า

ข้อเสนอแนะ 37
บรรณานุกรม 38
ประวัติผู้เขียน 39

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

สารบัญตาราง

ตารางที่ หน้า

2.1 การเปรียบเทียบความสามารถของ RAID แต่ละระดับ 7
2.2 ค่า IOps มาตราฐานของดิสก์แต่ละประเภท 9
2.3 อธิบายเปรียบเทียบ IOps ของ RAID แต่ละประเภท 9
3.1 ข้อมูลที่ได้จากสตอเรจไอบีเอ็มรุ่น N3600 20
3.2 ข้อมูลอัตราการใช้งานสตอเรจในหน่วยวัน (GB/Day) 21
3.3 ค่าของการพยากรณ์ของวันที่ 7/17/2012 24
3.4 ผลการปรับค่าค่าถ่วงน้ำหนัก (Alpha) 26
4.1 ค่าที่ได้จากการพยากร์จากแบบจำลอง Moving Average Method ที่ N = 7 27
4.2 ค่าการพยากรณ์แบบจำลอง Single Exponential Smoothing ที่ $\alpha = 0.5$ 28
4.3 ผลลัพธ์ที่ได้จากการพยากรณ์ของ MV และ SES ของวันที่ 7/17/2012 29
4.4 ผลลัพธ์ที่ได้จากการพยากรณ์ของ SES ที่ $\alpha = 0.5$ 29

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

สารบัญรูป

รูปที่ หน้า

2.1 ข้อมูลเปรียบเทียบ Raid แต่ละประเภท 6

2.2 การทำงานของ Raid 10 6

3.1 อัตราการใช้งานสตอเรจในหน่วยของวัน (GB/Day) 21

3.2 การเปลี่ยนแปลงของข้อมูลที่ N= 6 22

3.3 การเปลี่ยนแปลงของข้อมูลที่ N= 7 23

3.4 การเปลี่ยนแปลงของข้อมูลที่ N= 8 23

3.5 การเปลี่ยนแปลงของข้อมูลที่ N= 9 23

3.6 การเปลี่ยนแปลงของข้อมูลที่ Alpha=0.2 25

3.7 การเปลี่ยนแปลงของข้อมูลที่ Alpha=0.3 25

3.8 การเปลี่ยนแปลงของข้อมูลที่ Alpha=0.4 25

3.9 การเปลี่ยนแปลงของข้อมูลที่ Alpha=0.5 26

3.10 การเปลี่ยนแปลงของข้อมูลที่ Alpha=0.6 26

3.11 การเปลี่ยนแปลงของข้อมูลที่ Alpha=0.7 26

4.1 การเปรียบของข้อมูลที่ N=7 กับชุดข้อมูลต้นแบบ 28

4.2 การเปรียบของข้อมูลที่ Alpha=0.5 กับ ชุดข้อมูลต้นแบบ 28

4.3 การเปรียบของข้อมูลรายเดือน กับ ชุดข้อมูลต้นแบบ 30

4.4 โปรแกรมการพยากรณ์โดยใช้วิธีของอนุกรมเวลา 30

4.5 ส่วนองค์ประกอบของโปรแกรม 31

4.6 การนำข้อมูลเข้า 31

4.7 การแสดงผลของการพยากรณ์ 33

4.8 ผลลัพธ์ที่ได้จากการคำนวณ 34

4.9 ผลลัพธ์ที่ได้จากการนำเข้าของชุดข้อมูลต้นแบบ 34

4.10 ผลลัพธ์ที่ได้จากการพยากรณ์ของข้อมูลต้นแบบ 35

4.11 ผลการเปรียบเทียบโมเดลต้นแบบ กับ ชุดข้อมูลต้นแบบ 35

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

บทที่ 1
บทนำ

1.1 ความเป็นมา

ในปัจจุบันระบบสารสนเทศกลายเป็นหัวใจหลักในการขับเคลื่อนธุรกิจของหลาย ๆ องค์กร และระบบสารสนเทศ ได้มีการพัฒนาอย่างต่อเนื่องควบคู่ไปกับความต้องการของผู้ใช้งานที่มีอย่างไม่สิ้นสุด ซึ่งแต่ละองค์กรจะต้องมีการวางแผนและจัดสรรทรัพยากรของระบบสารสนเทศที่มีอยู่อย่างจำกัดให้เกิดประโยชน์สูงสุด

ด้วยระบบสารสนเทศจำเป็นต้องมีการจัดเก็บข้อมูลต่าง ๆ ลงบนพื้นที่เก็บข้อมูลขนาดใหญ่ที่เรียกว่า สตอเรจ (Storage) และต้องสามารถรองรับการใช้งานในอนาคตได้ ในปัจจุบันแต่ละองค์กรต่างประสบปัญหากับการให้บริการพื้นที่ของสตอเรจไม่เพียงพอต่อการใช้งาน เกิดความล่าช้าในการเข้าถึงข้อมูล ทั้งทางด้านการบันทึกข้อมูล และการเรียกใช้งาน จากปัญหาดังกล่าวทุกองค์กรต่างมีแนวคิดในการแก้ไข ซึ่งมีอยู่ 2 แนวทาง คือ การเพิ่มจำนวนของดิสก์ให้กับสตอเรจเพื่อเพิ่มความสามารถของสตอเรจให้มากขึ้น หรือ จัดหาสตอเรจตัวใหม่เข้ามาแทนที่ระบบเดิม ทั้งสองวิธีการต่างมีค่าใช้จ่ายทั้งสิ้น ทำให้ทุกองค์กรตระหนักถึงความสำคัญและปัญหาที่จะเกิดขึ้น นอกจากนั้น ถ้าการบริหารและการจัดการทรัพยากรของสตอเรจผิดพลาด อาจทำให้ระบบสารสนเทศที่ใช้เป็นสื่อกลางของการดำเนินธุรกิจนั้น มีผลกระทบต่อการดำเนินธุรกิจทุกภาคส่วนขององค์กร นั่นหมายความว่าธุรกิจที่ดำเนินอยู่นั้นจะต้องหยุดชะงักในการให้บริการ และจะมีผลเสียที่ตามมา ซึ่งมิอาจประเมิณค่าความเสียหายได้ ด้วยเหตุปัจจัยนี้ ทำให้หลายองค์กรให้ความสำคัญ และตระหนักถึงปัญหาที่อาจจะเกิดขึ้น จึงจำเป็นต้องหาวิธีการวางแผนและวิธีการป้องกันต่างๆ ที่อาจจะเกิดขึ้นกับแหล่งข้อมูลอันมีค่าเหล่านี้

จากการศึกษาเทคโนโลยีสตอเรจในปัจจุบันจะเป็นศูนย์รวมในการจัดเก็บข้อมูล ซึ่งข้อมูลเหล่านี้จะถูกเรียกให้ด้วยแอ็ปพริเคชั่นที่อยู่บนเครื่องแม่ข่ายหรือ เครื่องลูกข่ายเข้ามาเรียกใช้งานข้อมูล ทำให้ข้อมูลเพิ่มขึ้นอย่างรวดเร็ว ซึ่งแต่ละองค์กรต่างได้ตระหนักถึงปัญหาในการดูแลและจัดเก็บข้อมูลอันมีค่าเหล่านั้น แต่ด้วยศักยภาพของสตอเรจไม่สามารถบอกได้ว่าจะสามารถรองรับการใช้งานในอนาคตได้ถึงเมื่อไหร เพียงแต่สามารถเก็บข้อมูลของการใช้งานแบบ Real Time ว่ามีการใช้งานเป็นอย่างไร และข้อมูลดังกล่าวไม่สามารถนำเอามาประเมิณหรือคาดเดาเหตุการณ์ ในอนาคตได้

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

ด้วยเหตุผลดังกล่าว ผู้เขียนจึงมีแนวความคิดที่จะนำข้อมูลจริงจากพฤติกรรมการใช้งานของสตอเรจ ในอดีตมาทำการวิเคราะห์ด้วยหลักการทฤษฎีของอนุกรมเวลา (Time Series Method) เพื่อพยากรณ์หาความสามารถของสตอเรจ ที่จะสามารถรองรับการใช้งานในอนาคตอย่างมีแบบแผนและเป็นระบบ และเพื่อเป็นข้อมูลให้ทางฝ่ายบริหารสามารถนำไปวางแผนรับมือได้ทันเวลา

1.2 วัตถุประสงค์

1.2.1 เพื่อพัฒนาโปรแกรมต้นแบบในการพยากรณ์หาความสามารถของสตอเรจที่จะรองรับการใช้งานในอนาคตได้

1.2.2 เพื่อนำผลที่ได้จากการพยากรณ์ไปใช้เป็นข้อมูลสนับสนุนในการตัดสินใจของผู้บริหารด้านการลงทุนขององค์กรในอนาคตได้

1.3 ขอบเขตการดำเนินงาน

การศึกษาโครงงานนี้ได้กำหนดขอบเขตในการศึกษาข้อมูลจริงจากพฤติกรรมการใช้งานของสตอเรจ ในอดีตมาทำงานวิเคราะห์ด้วยหลักการทฤษฎีของอนุกรมเวลา โดยใช้ทฤษฎีการปรับเรียบแบบอย่างง่าย (Moving Average) ทฤษฎีการปรับเรียบแบบเอ็กซ์โปเนนเชียล (Exponential Smoothing) เพื่อเปรียบเทียบประสิทธิภาพผลลัพธ์ของการพยากรณ์ โดยจะใช้ข้อมูลต้นแบบชุดเดียวกัน จากนั้นจะนำทฤษฎีนั้นมาจัดทำเป็น Model ใช้งานในการคำนวณหาอัตราการใช้งานของสตอเรจต่อไป

1.4 ขั้นตอนและวิธีการดำเนินงาน

เพื่อให้การศึกษาเป็นไปตามวัตถุประสงค์ และขอบเขตที่กำหนด จึงได้กำหนดขั้นตอนในการดำเนินงานไว้ ดังนี้

1.4.1 ศึกษาทฤษฎีของสตอเรจเทคโนโลยี (Storage Technology)

1.4.2 ศึกษาทฤษฎีของอนุกรมเวลา (Time Series)

1.4.3 ศึกษาทฤษฎีที่เกี่ยวข้องกับการปรับเรียบของข้อมูลแบบอย่างง่ายแบบโดยวิธีหาค่าเฉลี่ยเคลื่อนที่ (Moving Average)

1.4.4 ศึกษาทฤษฎีที่เกี่ยวข้องกับการปรับเรียบของข้อมูลแบบเอ็กซ์โปเนนเชียล (Exponential Smoothing)

1.4.5 กำหนดวัตถุประสงค์ในการพัฒนาโปรแกรม

1.4.6 พัฒนาโปรแกรมเพื่อคำนวณหาขนาดพื้นที่เก็บข้อมูลที่จะรองรับการใช้งานในอนาคต

1.4.7 ทดสอบการใช้งานโปรแกรม

1.4.8 สรุปผลการศึกษาและข้อเสนอแนะ

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

1.5 ประโยชน์ที่คาดว่าจะได้รับ

1.5.1 สามารถใช้เทคโนโลยีสตอเรจได้อย่างมีประสิทธิภาพและเกิดประโยชน์สูงที่สุด

1.5.2 สามารถนำทฤษฎีของอนุกรมเวลามาประยุกต์ใช้งานได้

1.6 เครื่องมือที่ใช้ในการพัฒนาระบบ

ใช้ MATLAB เป็นเครื่องมือในการพัฒนาโปรแกรม เพื่อหาความสัมพันธ์ของอนุกรมเวลาในการวิเคราะห์ข้อมูลและพยากรณ์อัตราการใช้งานของสตอเรจในอนาคต

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

บทที่ 2
ทฤษฎีที่เกี่ยวข้องและทบทวนวรรณกรรม

2.1 สตอเรจเทคโนโลยี

เนื่องด้วยข้อมูลเป็นสิ่งสำคัญของแต่ละองค์กร ซึ่งข้อมูลเหล่านี้จำเป็นต้องถูกจัดเก็บลงบนสื่อกลางที่เรียกว่าสตอเรจ แล้วแอ็ปพริเคชั่นต่างๆ ทั้งที่อยู่บนเครื่องแม่ข่ายหรือ เครื่องลูกข่ายเข้ามาเรียกใช้งาน ทำให้จำนวนข้อมูลเพิ่มขึ้นอย่างรวดเร็ว ซึ่งแต่ละองค์กรต่างได้ตระหนักถึงปัญหาในการดูแลและจัดเก็บข้อมูลอันมีค่าเหล่านั้น และจำเป็นต้องมีการวางแผนการใช้งานสตอเรจอย่างไรและ จะต้องคำนึงถึงปัจจัยต่าง ๆ ดังนี้ ปัจจัยที่หนึ่งการเพิ่มขึ้นอย่างรวดเร็วของข้อมูล รวมทั้งข้อมูลที่เป็นรายการทางธุรกรรมและข้อมูลอ้างอิงต่างๆ ปัจจัยที่สองข้อกำหนดที่เพิ่มมากขึ้นในการจัดเก็บข้อมูลเพื่อให้พร้อมสำหรับการเรียกใช้งานและสามารถกู้ข้อมูล ได้อย่างรวดเร็ว เพื่อเป็นหลักประกันให้ธุรกิจสามารถดำเนินต่อไปได้และสามารถกอบกู้ระบบจากภัยพิบัติต่าง ๆ ได้ ปัจจัยที่สามการเข้าถึงข้อมูล จากส่วนต่าง ๆ ขององค์กร เช่น ผู้ใช้งานทั่วไประบบต่าง ๆ หรือโปรแกรมแอ็ปพริเคชั่นต่าง ๆ จากทั่วทั้งองค์กร สามารถเข้าถึงข้อมูลและใช้บริการได้ ทำให้เกิดแรงผลักดันในการพัฒนาด้านต่างๆ ให้ดีขึ้นด้วย สำหรับการเลือกใช้งานของสตเรจนั้นจำเป็นต้องคำนึงถึงองค์ประกอบหลักของทางด้าน ฮาร์ดแวร์ว่าควรมีหลักการพิจารณาในการเลือกใช้งานสตอเรจให้เหมาะสมกับงานแต่ละประเภท โดยจะแบ่งออกเป็น 3 ประเภทใหญ่ ๆ ดังนี้

2.1.1 DAS (Direct Attach Storage)

ด้วยการให้บริการของเครื่องแม่ข่ายต่างให้บริการพื้นที่ให้กับการใช้งานของแอ็ปพลิเคชั่นต่างๆ ของแต่ละองค์กรและเป็นปัจจัยพื้นฐานของการใช้งานสตอเรจ ซึ่งการใช้งานแบบนี้เรียกว่า DAS ซึ่งลักษณะการให้บริการแบบ Block-Level แต่การเติบโตทางธุรกิจบวกกับการใช้งานของแอ็ปพลิเคชันต่างๆ ส่งผลให้ DAS ไม่สามารถตอบสนองความต้องการเท่าที่ควร เนื่องจาก DAS มีข้อจำกัดของการเพิ่มขนาดพื้นที่และความสามารถการเรียกใช้งานจากหลาย ๆ แอ็ปพลิเคชันในเวลาเดียวกัน รวมไปถึงการดูแลเครื่องแม่ข่ายให้สามารถให้บริการ หากมีเหตุการณ์ที่ทำเครื่องแม่ข่ายไม่สามารถให้บริการอาจส่งผลกระทบต่อการดำเนินธุระกิจขององค์กรได้

2.1.2 SAN (Storage Area Network)

ด้วยรูปแบบการให้บริการที่ตอบสนองความต้องการและลดความเสียงจากเครื่องแม่ข่ายเพียงเครื่องเดียว ทุกองค์กรได้ให้ความสำคัญต่อการวางแผนรับมือและเพิ่มช่องทางในการอุปกรณ์และเทคโนโลยีที่สามารถทำให้สามารถดำเนินธุรกิจต่อไป เช่น การทำ Cluster Service ให้กับเครื่องแม่ข่าย ด้วยเทคโนโลยี ของ SAN เข้ามาช่วยในการจัดการข้อมูลและการใช้งานของแอ็ปพลิเคชัน เนื่องจาก SAN เป็นการให้บริการเบบ Block-Level เช่นเดียวกันกับ DAS แต่เทคโนโลยี SAN

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

สามารถเพิ่มขยายจำนวนของพื้นที่ให้บริการได้ตามต่อความต้องการขององค์กร และลักษณะการใช้งานจะใช้การเชื่อมต่อแบบไฟเบอร์ ซึ่งจะช่วยลดปัญหาเรื่องของระยะทางในการเชื่อมต่อละการสูญหายของข้อมูลจากสัญญาณรบกวนต่างๆได้ แต่การใช้งานของ SAN เทคโนโลยีจะมีค่าใช้จ่ายค่อนข้างสูง ในการเชื่อมต่อระหว่างเครื่องแม่ข่ายกับสตอเรจซึ่งการใช้งานเป็นแบบไฟเบอร์ทั้งหมด

2.1.3 NAS (Network Attach Storage)

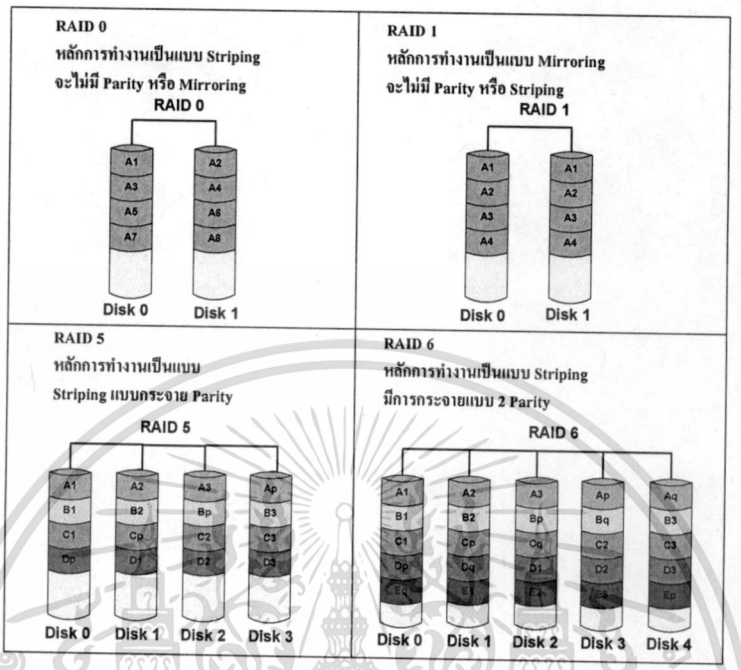
เป็นรูปแบบการให้บริการของสตอเรจแบบ File-Level ซึ่งการให้บริการของ NAS นั้นหมายความว่าจะต้องมีเชื่อมต่ออุปกรณ์ของ NAS สตอเรจเข้ากับเน็ตเวอค์ขององค์กรได้โดยตรง และการใช้งานจะต้องทำผ่านเน็ตเวอค์ซึ่งลักษณะการทำงานคล้ายกับไฟล์เซิฟเวอร์ สามารถเข้าเรียกใช้งานได้ในเวลาเดียวกันได้ โดยไม่ทำให้เกิดความเสียหายข้อมูล แต่ยังมีข้อจำกัดของการใช้งานแบบ NAS สตอเรจ ซึ่งยังคงเป็นปัญหาเรื่องของเน็ตเวอค์ ด้วยความเร็วที่ใช้ในปัจจุบันบนสายสัญญาณ UTP (Unshielded Twisted Pair) ยังคงมีความเร็วอยู่ที่ 100 Mbps, 1 Gbps หรือ 10 Gbps ในกรณีที่มีการใช้งานแบบเน็ตเวอร์คโปรโตคอล และอีกหนึ่งปัจจัยที่ส่งผลกระทบต่อการใช้งาน คือเรื่องของระยะทางในการเชื่อมต่อ ซึ่งจะใช้งานที่ระยะทางไม่ไกลมากนัก เนื่องจากมีปัญหาเรื่องการสูญหายของข้อมูลระหว่างทาง ที่อาจเกิดจากสายสัญญาณรบกวนหรือปัจจัยจากสภาวะแวดล้อมอื่น ๆ ทำให้มีผลกระทบต่อความเสถียรภาพของข้อมูล

2.2 RAID เทคโลโลยี

สำหรับการเลือกใช้งานสตอเรจแต่ละประเภทนั้น จำเป็นต้องคำนึงถึงหัวใจหลักในด้านการลงทุน ซึ่งแต่ละองค์กรต่างมีความต้องการที่จะจัดเก็บข้อมูลและความคล่องตัวในการเลือกใช้งานที่แตกต่างกัน ถ้าข้อมูลของลูกค้า มีความสำคัญมาก จะต้องมีการลงทุนที่สูงในการเลือกใช้เทคโนโลยีเพื่อการรองรับการเรียกคืนข้อมูลได้อย่างรวดเร็ว หากมีเหตุการณ์ที่ไม่ได้คาดคิดเกิดขึ้น เพื่อเป็นหลักประกันว่า ธุรกิจจะต้องดำเนินต่อไปได้ ทางด้านเทคโนโลยีก็มีการพัฒนาและมีการป้องกันความผิดพลาดที่อาจเกิดจากตัวอุปกรณ์เพื่อสร้างความมั่นใจให้กับผู้ใช้งาน โดยมีแนวคิดที่จะนำเอาจำนวนฮาร์ดดิสก์มากกว่า 1 ลูก มาเชื่อมต่อกัน เพื่อเพิ่มขนาดของพื้นที่ใช้งานให้มากขึ้น และเพิ่มประสิทธิ์ภาพในการให้บริการได้เร็วมากขึ้น รวมไปถึงการเพิ่มความน่าเชื่อถือ และความปลอดภัยของข้อมูลมากขึ้น ซึ่งเรียกว่าระบบนี้ว่า RAID (Redundant Arrays Of Independent Discs) โดยมีองค์ประกอบหลัก คือ RAID Controller ที่ทำหน้าที่เป็นเสมือนโครงข่ายหลัก ที่ทำงานอยู่เบื้องหลังทำหน้าที่ควบคุมและการจัดการของฮาร์ดดิสก์ ทำให้มีความคล่องตัวมากยิ่งขึ้น สำหรับการเลือกให้งานระบบ RAID นั้นจะต้องคำนึงถึง แอ็ปพลิเคชัน และความต้องการขององค์กรว่าต้องการประสิทธิภาพในการใช้งานและความปลอดภัยในข้อมูลนั้น ๆ อย่างไร โดยแบ่งออกเป็น 3 กลุ่มดังนี้

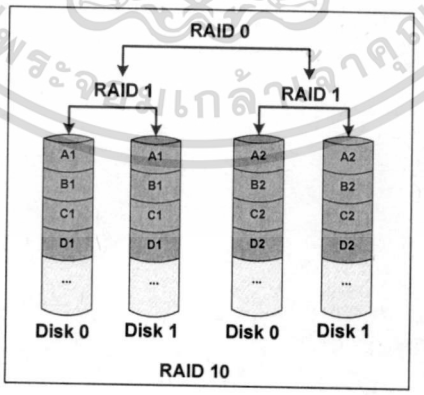
เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

2.2.1 RAID พื้นฐาน ปัจจุบันมีตั้งแต่ RAID 0 ถึง RAID 6 แต่ที่ได้รับความนิยมคือ RAID 0, 1, 5, 6 ซึ่ง RAID แต่ละระดับนั้นต่างให้ประสิทธิภาพการใช้งานที่ต่างกัน และความเนื่อเชื่อในการป้องความปลอดภัยที่แตกต่างกัน ในกรณีที่ จำนวนฮาร์ดดิสก์ชำรุดหรือเสียหาย



รูปที่ 2.1 ข้อมูลเปรียบเทียบ Raid แต่ละประเภท

2.2.2 RAID ลูกผสม การนำ RAID แบบพื้นฐานมาเชื่อมต่อกันด้วยเป็น RAID อีกแบบหนึ่ง เช่น การนำ RAID 1 สองชุดมาเชื่อมต่อกัน แบบ RAID 0 เลยเรียกว่า RAID 10 หรือการนำ RAID สองชุดมาเชื่อมต่อกันแบบ RAID 0 เลยเรียกว่า RAID 50 เป็นต้น



รูปที่ 2.2 การทำงานของ Raid 10

2.2.3 RAID Proprietary คือ RAID เฉพาะแบบของแต่ละผู้ผลิต เช่น RAID 1E โดย IBM, RAID S โดย EMC, RAID-DP โดย NetAPP สตอเรจ เป็นต้น ดังนั้น การคำนวณความจุรวมของ

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

ระบบ จะต้องเข้าใจวิธีคำนวณพื้นที่ ที่จะสามารถใช้งานได้จริงของ RAID แต่ระดับ ด้วยวิธีการเชื่อมต่อฮาร์ดดิสก์เข้าระบบด้วยว่าต่อแบบธรรมดา หรือใช้ RAID แบบไหน

จากการพิจารณาความสามารถของ RAID แต่ละระดับต่างให้ประสิทธิ์ภาพของการใช้งานที่แตกต่างกัน ดังนั้นจำเป็นต้องทำความเข้าใจโครงสร้างการทำงานของ RAID และวิธีการคำนวณหาจำนวนพื้นที่สุทธิหลังจากการทำ RAID และความสามารถในการป้องกันข้อมูลเมื่อมีจำนวนของฮาร์ดดิสก์ที่สามารถยอมให้เสียได้จำนวนเท่าไร แต่ไม่ส่งผลกระทบต่อข้อมูล ดังตารางที่ 2.1

ตารางที่ 2.1 การเปรียบเทียบความสามารถของ RAID แต่ละระดับ

ข้อเปรียบเทียบ	Raid 0	Raid 1	Raid 5	Raid 6
1.จำนวนฮาร์ดดิสก์ขั้นต่ำ	2	2	3	4
2.พื้นที่เก็บข้อมูลหลังหัก Overhead	N*K	N*K/2	(N-1)*K	(N-2)*K
3.จำนวนฮาร์ดดิสก์ที่พังได้ โดยข้อมูลไม่เสียหาย	0	N/2	1	2
4.อัตราการอ่านข้อมูล	N*X	N*X	(N-1)*X	(N-1)*X
5.อัตราการเขียนเข้อมูล	N*X	X	(N-1)*X	(N-1)*X
N=จำนวนฮาร์ดดิสก์ K=Capacity ต่อฮาร์ดดิสก์ 1 ลูก X=อัตราการอ่าน หรือเขียนข้อมูลของฮาร์ดดิสก์ 1 ลูก โดยปรกติ อัตราการเขียนข้อมูลของ RAID 5 กับ RAID 6 มักจะช้ากว่าอัตราการอ่านข้อมูล เนื่องจากต้องคำนวณ Parity ด้วย แต่สำหรับ Hard Disk Controller ในระดับไฮเอนด์ อัตราการเขียนข้อมูล อาจมีค่าใกล้เคียงกับอัตราการอ่าน				

2.3 ปัจจัยที่มีผลต่อการใช้งานของการจัดเก็บข้อมูล

สำหรับองค์กรที่มีระบบการทำงานที่มีอัตราการเรียกใช้ข้อมูลจำนวนมาก จะต้องให้ความสนใจเกี่ยวกับความสามารถของเทคโนโลยี ว่าสามารถรองรับอัตราการอ่านเขียนของข้อมูล IOps (Input-Output Operations/Second) เป็นพิเศษ เนื่องจากเป็นข้อจำกัดของฮาร์ดดิสก์ที่ไม่สามารถรองรับจำนวน IOps จำนวนมหาศาลได้ ถึงแม้ว่าเครื่องคอมพิวเตอร์จะมีหน่วยประมวลผลเร็วเพียงใดก็ตามก็ไม่สามารถช่วยให้ระบบทำงานได้เร็วขึ้นได้ ดังนั้นจำเป็นต้องมีการเลือกใช้อุปกรณ์ให้เหมาะสมกับงานแต่ละประเภทด้วย ซึ่งความสามารถของอุปกรณ์ที่จะพิจารณาได้จากเทคโนโลยีของฮาร์ดดิสก์แต่ละประเภทดังนี้

2.3.1 ประเภทของ Hard Disk ปัจจุบันมีหน่วยความจุ 3 ประเภทที่สามารถนำมาใช้งานเป็นสตอเรจ ประเภทที่หนึ่ง SATA (Serial Advanced Technology Attachment) หรือ Serial ATA ฮาร์ดดิสก์ประเภท SATA เป็นอุปกรณ์มาตรฐานในเครื่องคอมพิวเตอร์ส่วนบุคคล และได้มีพัฒนาการจนมีน่าเชื่อถือเพียงพอที่จะนำมาใช้ในเซิร์ฟเวอร์และสตอเรจ ประเภทที่สอง. SAS

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

(Serial Attached SCSI) ฮาร์ดดิสก์ประเภท SAS ได้พัฒนาขึ้นทดแทนฮาร์ดดิสก์ประเภท SCSI ซึ่งในปัจจุบันได้นำมาใช้งานกับเซิร์ฟเวอร์และสตอเรจ ประเภทที่สาม. SSD (Solid-State Drive) หน่วยความจุประเภท SSD เริ่มได้รับความนิยมมากขึ้น ในปัจจุบันก็มีการนำมาใช้งานบนเครื่องคอมพิวเตอร์เดสก์ทอป, เซิร์ฟเวอร์ และสตอเรจ หรือสำหรับงานเฉพาะด้านที่มีต้องการความเร็วในการเข้าถึงข้อมูล และการตอบสนองต่อการใช้งานแบบต่อเนื่อง ซึ่งฮาร์ดดิสก์ประเภท SSD ก็เป็นอีกหนึ่งทางเลือกสำหรับการเลือกใช้งาน

2.3.2 ประสิทธิ์ภาพของการจัดเก็บข้อมูล การประสิทธิ์ของการจัดเก็บข้อมูลนั้นไม่ว่าจะเป็น Hard Disk, SSD หรือ SAN (Storage Area Network) นิยมวัดสามารถตอบสนองการอ่านเขียนข้อมูลได้กี่ครั้งต่อวินาที หรือกี่ IOPS (Input/Output Operations Per Second) โดยหลักการคำนวณค่า IOps เบื้องต้นนั้นคือ

เวลาที่ ใช้ในการอ่านเขียนฮาร์ดดิสก์ครั้งหนึ่งๆ คือ

ให้ A = (Latency Time)

B = (Seek Time)

จะได้สมการเวลาที่ใช้ในการอ่านเขียนต่อครั้ง คือ

$$A + B \tag{2.1}$$

หรือ เวลาใช้ในการอ่านเขียนเฉลี่ยต่อครั้ง

ให้ Av = (Average Latency Time)

Bv = (Average Seek Time)

จะได้สมการเวลาที่ใช้ในการอ่านเขียนเฉลี่ยต่อครั้ง คือ

$$(Av) + (Bv) \tag{2.2}$$

จากอัตราการอ่านเขียนจากค่าของ IOps นั้นไม่ได้ขึ้นอยู่กับความเร็วของอินเทอร์เฟซของฮาร์ดดิสก์ เช่น SATA II, SATA III, SAS แต่จะขึ้นอยู่กับ Spin Speed ซึ่งสรุปได้ดังตารางที่ 2.2

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

ตารางที่ 2.2 ค่า IOps มาตราฐานของดิสก์แต่ละประเภท

Spin Speed (RPM)	AVR Latency (ms)	Seek Time (ms)	IOPS
5,400	5.5	7.0 - 14.4	50 - 80
7,200	4.2	5.9 - 9.1	75 - 100
10,000	3	3.6 - 5.0	125 - 150
15,000	2	2.7 - 3.7	175 - 210

ค่ากลาง (Safe Value) ที่ใช้อ้างอิงในการคำนวณ Storage IOPS	
ประเภท	IOPS
5,400	50
7,200	75
10,000	125
15,000	175
SSD	6000

สำหรับวิธีการพิจารณา หาค่า IOps จาก RAID ไม่ว่าจะใช้งาน RAID แบบ Local Storage หรือ SAN ก็ตาม โดยหลักการประเมิน IOps นั้นขึ้นอยู่กับประเภทของ RAID

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

ตารางที่ 2.3 อธิบายเปรียบเทียบ IOps ของ RAID แต่ละประเภท

RAID	Read IOPS	Write IOPS (Write Factor)	คำอธิบาย
0	1	1	การอ่านหรือเขียนข้อมูลจะเกิด 1 IOPS
1	1	2	ในการเขียนข้อมูลเกิดที่ดิสก์ทั้งสองชุดคู่กันเสมอ
2	1	4	ในการเขียนข้อมูลระบบต้องอัพเดต Parity Data ด้วยดังนี้ -อ่าน 1 ครั้ง สำรับข้อมูลใน Stripe Set -อ่าน 1 ครั้ง สำหรับ Parity Data -เขียน 1 ครั้ง สำหรับข้อมูลใน Strip Set -เขียน 1 ครั้ง สำหรับข้อมูลใน Parity Data
6	1	6	ใช้หลักการเดียวกับ RAID-5 แต่เนื่องจาก มี Parity Disk 2 ชุด จึงต้องอ่าน 3 ครั้งและเขียน 3 ครั้ง
10	1	2	ใช้หลักการเดียวกับ RAID-1
ดังนั้น Total IOPS=n*IOPS/[R+(W*F)]			
โดย : n คือ จำนวนฮาร์ดดิสก์ IOPS คืออัตรา IOPS ของฮาร์ดดิสก์ 1 ลูก R = อัตราการอ่านข้อมูล(%) W = อัตราการเขียนข้อมูล(%) F = Write Factor			

2.4 อนุกรมเวลาและการพยากรณ์

จากข้อมูลของการดำเนินชีวิตในปัจจุบันนั้น ต่างมีความผูกพันธ์เกี่ยวกับการพยากรณ์ตลอดเวลา เช่นการคาดการณ์ลักษณะอากาศ การคาดการณ์การซื้อขายของตลาดหุ้น ซึ่งการพยากรณ์ดังกล่าวต่างใช้ข้อมูลจากอดีตมาช่วยสนับสนุนในการทำนายอนาคตว่าจะเป็นอย่างไร แต่ค่าของการพยากรณ์ที่แม่นยำนั้น ไม่ได้ขึ้นอยู่กับข้อมูลในอดีตเพียงอย่างเดียว แต่ยังขึ้นอยู่ประสบการณ์ของผู้พยากรณ์ และความชำนาญในเรื่องนั้น ๆ ด้วย ทั้งนี้การพยากรณ์ที่มีประสิทธิภาพนั้นจะต้องอาศัยข้อมูลในอดีตที่เกี่ยวข้องรวมกับประสบการณ์ของผู้พยากรณ์และปัจจัยรอบด้านที่มีผลต่อค่าความแม่นยำของการพยากรณ์ ดังนั้นการพยากรณ์ที่มีประสิทธิ์จะต้องอาศัยการวิเคราะห์ข้อมูลทางสถิติเข้ามาช่วย คือ อนุกรมเวลา เพื่อเพิ่มความแม่นยำในการพยากรณ์ให้มากขึ้นและมีค่าความคลาดเคลื่อนน้อยที่สุด

2.4.1 อนุกรมเวลา (Time series) เป็นเซตข้อมูลเชิงปริมาณที่จัดเก็บในช่วงเวลาหนึ่ง ข้อมูลของอนุกรมเวลาอาจอยู่ในลักษณะที่เป็นข้อมูลรายปี รายไตรมาส หรือรายเดือนก็ได้ ทั้งนี้ขึ้นอยู่กับความเหมาะสมในการนำไปใช้ประโยชน์ซึ่งการพยากรณ์อนุกรมเวลา เป็นวิธีการทางสถิติที่ใช้

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

ข้อมูลในอดีตมาวิเคราะห์หาต้นแบบอธิบายลักษณะความสัมพันธ์ระหว่างค่าสังเกตที่เก็บตามลำดับเวลา และใช้ต้นแบบนั้นในการพยากรณ์ค่าสังเกตในอนาคต วิธีการวิเคราะห์อนุกรมเวลามีหลายวิธี ซึ่งแบ่งออกเป็น 2 กลุ่มใหญ่ ๆ

1) รูปแบบการพยากรณ์แบบเส้นตรง (Linear Approach) สำหรับค่าพยากรณ์โดยการวิเคราะห์แนวโน้มด้วยวิธีกำลังสองน้อยที่สุดในรูปแบบเส้นตรง สามารถแสดงสมการได้ดังนี้

$$\hat{Z}_t(\ell) = \alpha + \beta t \tag{2.3}$$

โดยที่

$\hat{Z}_t(\ell)$ คือ ค่าพยากรณ์ของ $\hat{Z}_t$

α, β คือ ค่าพารามิเตอร์

t คือ เวลา โดยที่ $t = 1,2,3,\dots,n$

2) รูปแบบการพยากรณ์แบบที่ไม่ใช้เส้นตรง (Non-Linear Approach) สมการค่าพยากรณ์โดยการวิเคราะห์แนวโน้มในข้างต้นนั้น เป็นสมการแนวโน้มแบบเส้นตรง แต่ในความเป็นจริงแล้วสมการแนวโน้มอาจเป็นเส้นโค้งก็ได้ ดังนั้นการใช้สมการแนวโน้มเส้นตรงในการพยากรณ์อาจทำให้เกิดความคลาดเคลื่อน สำหรับรูปแบบสมการแนวโน้มที่ไม่เป็นเส้นตรงที่มักพบ เสมอ ได้แก่ แนวโน้มโพลิโนเมียล (Polynomial Trend) มีรูปแบบสมการดังนี้

$$\hat{Z}_t(\ell) = \alpha_0 + \beta_1 t + \beta_2 t^2 + \beta_3 t^3 + \cdots + \beta_n t^n \tag{2.4}$$

โดยที่

$\hat{Z}_t(\ell)$ คือ ค่าพยากรณ์ของ $\hat{Z}_t$

α, β คือ ค่าพารามิเตอร์

t คือ เวลา โดยที่ $t = 1,2,3,\dots,n$

2.4.2 การเลือกวิธีการพยากรณ์ ในการพยากรณ์ที่ให้ได้คุณภาพ จะต้องมีการศึกษาปัจจัยต่าง ๆ และประสบการณ์ของผู้พยากรณ์ ในการพยากรณ์ ดังนั้นการวิเคราะห์ปัญหาจากข้อมูลในอดีตเป็นสิ่งสำคัญ และจำเป็นต้องเข้าใจปัญหาและจะต้องเลือกใช้วิธีการพยากรณ์ให้ถูกต้องด้วย เพื่อให้ได้ผลลัพธ์ที่ถูกต้องและแม่นยำ โดยการพิจารณาข้อมูลในอดีตนั้นว่าเป็นข้อมูลประเภทไหน เพื่อเป็นแนวทางในการคัดเลือกวิธีในการวิเคราะห์ข้อมูลให้ถูกต้องเพื่อให้ได้ผลลัพธ์ที่ตรงกับความต้องการของผู้ใช้งานมากที่สุด เช่น วิธีการพยากรณ์เชิงคุณภาพ หรือ วิธีการพยากรณ์เชิงปริมาณ ซึ่งทั้งสองต่างใช้หลักการณ์และแนวคิดที่แตกต่างกัน ซึ่งแบบจำลองแต่ละแบบมีรูปแบบในการศึกษาที่ต่างๆ กัน เช่น วิธีการปรับเรียบแบบเอ็กซ์โปเนนเชียล เป็นแบบจำลองที่ให้ความสำคัญกับข้อมูลใน

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

ปัจจุบัน และอดีต แต่ให้ความสำคัญกับข้อมูลในปัจจุบันมากกว่า ยังมีการพิจารณาเพิ่มเติมในส่วนของ แนวโน้ม และฤดูกาล และค่าความคลาดเคลื่อนของข้อมูลในอดีตมาใช้อธิบายข้อมูลในปัจจุบันและอนาคต ซึ่งในการพยากรณ์จะเลือกใช้แบบจำลองใดมาใช้นั้น ผู้พยากรณ์จำเป็นต้องพิจารณาสถานการณ์และลักษณะปัญหาที่ต้องการพยากรณ์ให้ชัดเจน ซึ่งอาจพิจารณาจากส่วนประกอบหลักต่อไปนี้

1) รูปแบบของค่าพยากรณ์ โดยทั่วไปแล้วเมื่อมีปัญหาที่ต้องการทราบค่าพยากรณ์ สิ่งหนึ่งที่ผู้พยากรณ์จำเป็นต้องทราบคือ รูปแบบของค่าพยากรณ์ที่ต้องการนำไปใช้ประโยชน์ต่อไปนั้นเป็นค่าพยากรณ์แบบจุดหรือแบบช่วง ทั้งนี้เพราะวิธีการพยากรณ์บางวิธีจะให้ค่าพยากรณ์แบบจุด

2) ความแม่นยำ ความแม่นยำของค่าพยากรณ์ เป็นส่วนประกอบหนึ่งที่ช่วยให้เลือกวิธีการพยากรณ์ที่เหมาะสมโดยในบางกรณีต้องการค่าพยากรณ์คร่าวๆ ในขณะที่บางปัญหาต้องการค่าพยากรณ์ที่แม่นยำ นอกจากนี้คุณสมบัติความแม่นยำของการพยากรณ์แต่ละวิธีแตกต่างกัน บางวิธีให้ค่าแม่นยำเฉพาะในช่วงระยะเวลาสั้นในอนาคต แต่บางวิธีอาจสามารถพยากรณ์ในระยะยาวได้ด้วย

3) กรอบเวลา เนื่องจากการหาค่าพยากรณ์ คือ การหาค่าของตัวแปรที่พิจารณาในอนาคตที่ช่วงเวลาห่างไปจากปัจจุบันเป็นวัน สัปดาห์ เดือน หรือปี เป็นต้น ช่วงเวลาเหล่านี้เรียกว่า กรอบเวลา ซึ่งโดยทั่วไปอาจจำแนกกรอบเวลาตามความยาวของกรอบเวลาเช่า

- ค่าพยากรณ์ระยะใกล้ หมายถึง ค่าพยากรณ์ในอนาคตในช่วงเวลาห่างจากปัจจุบันน้อยกว่า 1 เดือน
- ค่าพยากรณ์ระยะสั้น หมายถึง ค่าพยากรณ์ในอนาคตในช่วงเวลาห่างจากปัจจุบัน 1 เดือนแต่ไม่มากกว่า 3 เดือน
- ค่าพยากรณ์ระยะกลาง หมายถึง ค่าพยากรณ์ในอนาคตในช่วงเวลาห่างจากปัจจุบันมากกว่า 3 เดือน แต่น้อยกว่า 2 ปี ค่าพยากรณ์ระยะยาว หมายถึง ค่าพยากรณ์ในอนาคตในช่วงเวลาห่างจากปัจจุบัน 2 ปี หรือมากกว่า 2 ปี

2.4.3 การเปรียบเทียบแบบจำลองที่เหมาะ จากเกณฑ์การคัดเลือกแบบจำลอง ผู้พยากรณ์ มีความคลาดหวังว่าแบบจำลอง จะมาช่วยเรื่องผลของการพยากรณ์ที่จะทำให้ได้มีความแม่นยำ และมีความน่าเชื่อ และผลการพยากรณ์จะต้องมีความคลาดเคลื่อนน้อยที่สุดเมื่อ เปรียบเทียบระหว่างค่าพยากรณ์ล่วงหน้ากับข้อมูลจริง ซึ่งเกณฑ์ที่ใช้ในการเปรียบเทียบวิธีการพยากรณ์ จะพิจารณาจาก ค่าเปอร์เซ็นต์ความคลาดเคลื่อนสัมบูรณ์เฉลี่ย (Mean Absolute Percentage Error : MAPE) วิธีการพยากรณ์ที่ให้ค่า MAPE ต่ำที่สุดแสดงว่าวิธีการพยากรณ์นั้นมีค่าแม่นยำที่สุด

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

1) Mean Square Error (MSE)

$$\text{MSE} = \frac{1}{n}\sum_{i=1}^{n}(Z_t - \hat{Z}_t)^2 \qquad (2.5)$$

เมื่อ $\text{MSE} \geq 0$

เป็นวิธีที่ใช้กันทั่วไป ข้อเสียคือมีฐานเปรียบเทียบ และการที่ MSE มีค่าสูงแปลว่า คลาดเคลื่อนสูงหรือไม่ก็ไม่อาจแน่ใจได้ ขึ้นอยู่กับขนาดของข้อมูลด้วย

2) Mean Absolute Percentage Error (MAPE)

$$MAPE = \frac{1}{n}\sum_{i=1}^{n}\left[\frac{Z_t - \hat{Z}_t}{Z_t}\right]^2 \times 100 \qquad (2.6)$$

โดยที่ $PE_n = \frac{Z_t - \hat{Z}_t}{Z_t} \times 100$ ซึ่งได้มาจากการเทียบบัญญัติไตรยางค์นี้

ค่าจริงเท่ากับ z_n วิธีพยากรณ์ให้ค่าพยากรณ์คลาดเคลื่อนไป $z_t - \hat{z}_t$

ค่าจริงเท่ากับ 100 วิธีพยากรณ์ให้ค่าพยากรณ์คลาดเคลื่อนไป $\frac{Z_t - \hat{Z}_t}{Z_t} \times 100$

ข้อเสีย วิธีพยากรณ์ที่ให้ MAPE สูงกว่าไม่แน่ใจว่าจะเลวร้ายกว่าวิธีอื่น ควรมีวิธีการพยากรณ์วิธีหนึ่งรอให้วิธีอื่นไปเปรียบเทียบ MAPE จึงจะทราบได้ว่าวิธีใดดีกว่า

จะสังเกตได้ว่าการใช้เทคนิคการพยากรณ์ที่ แตกต่างกันจะทำให้ได้ ค่าการพยากรณ์อุปสงค์ที่แตกต่างกันไปด้วยทั้งนี้นักพยากรณ์จำเป็นต้องมีวิธีในการตรวจสอบว่าเทคนิคที่ใช้มีความแม่นยำในการพยากรณ์ ข้อมูลหนึ่ง ๆ มากน้อยเพียงไรเนื่องจากไม่มีเทคนิคการพยากรณ์ใดที่สามารถให้ผลแม่นยำตลอดเวลาเทคนิคการพยากรณ์นั้นๆ

2.4.4 แบบจำลองการทำให้เรียบแบบง่ายด้วยวิธีค่าเฉลี่ยเคลื่อนที่ (Moving Average Method) การวิเคราะห์ข้อมูลอนุกรมเวลา จะพิจารณาจากส่วนประกอบของข้อมูลอนุกรมเวลาว่าเป็นมีลักษณะเป็นอย่างไร โดยให้หลักการพิจารณาจากแนวโน้มของข้อมูลว่ามีการเปลี่ยนแปลง ไปทิศทางใด บางครั้งอาจมีทิศทางเพิ่มขึ้น หรือลดลง หรือมีทั้งสองแบบในชุดข้อมูลเดียวกัน ด้วยปัจจัยต่าง ๆ ที่มีอิทธิพลทำให้ลักษณะของข้อมูลเปลี่ยนแปลงไป ดังนั้นการพยากรณ์จำเป็นต้องนำข้อมูลมาทำการปรับเรียบเสียก่อน ก่อนที่จะทำการพยากรณ์ ซึ่งลักษณะของข้อมูลที่ใช้ในการวิเคราะห์จะต้องไม่มีอิทธิพลจากแนวโน้ม และ การผันแปรตามฤดูกาล และข้อมูลที่นำมาใช้ต้องมีลักษณะคงที่ โดยวิธีการปรับเรียบจะที่มีการพิจารณาค่าเฉลี่ยคงที่ ซึ่งตัวแบบถดถอยจะมีตัวแปรอิสระเป็นค่าของ อนุกรมเวลาที่สอดคล้องกับข้อมูล และการจะมีค่าเปลี่ยนแปลงไปมากน้อยขึ้นอยู่

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

กับ รอบระดับค่าเฉลี่ย ซึ่งอาจคงที่ หรือเปลี่ยนแปลงอย่างช้า ๆ เมื่อเวลาเปลี่ยนแปลงไป การพยากรณ์ค่าของอนุกรมเวลาจะมีสองแนวทาง คือการพยากรณ์ภายใต้ข้อตกลงที่ β เป็นค่าคงที่เสมอ และการพยากรณ์ภายใต้การยอมรับว่า β เปลี่ยนแปลงไปอย่างช้า ๆ ตามเวลาที่เปลี่ยนไป (หรือ β คงที่เฉพาะแห่ง)

$$Z_{t+j} = \beta + a_{n+j} \tag{2.7}$$

ถ้าพิจารณาค่า $Z_n(\ell)$ แทนค่าพยากรณ์เข้าไปที่สมการ (2.3) ค่าสังเกตในอนาคตที่เวลา $n + \ell$ ($Z_{n+\ell} = \beta + a_{n+j}$) แสดงได้ว่าค่าพยากรณ์ที่ให้ความคลาดเคลื่อน

เมื่อทราบค่า β

$$Z_n(\ell) = \beta \tag{2.8}$$

เมื่อไม่ทราบค่า β

$$Z_n(\ell) = \beta = \bar{Z} = \frac{1}{n}\sum_{t=1}^{n} Z_t \tag{2.9}$$

ถ้าต้องการปรับค่าพยากรณ์เมื่อได้ค่าสังเกตที่ $n + 1$ เพิ่มขึ้น อาจพยากรณ์ค่าจริงที่ ℓ หน่วยเวลาล่วงหน้า โดยเปลี่ยนจุดเริ่มต้นทำการพยากรณ์ไปเป็นที่ $n + 1$ ได้

$$Z_n(\ell) = \frac{1}{n+1}(Z_1 + \cdots + Z_n + Z_{n+1}) \tag{2.10}$$

$$= \frac{n}{n+1}\bar{Z} + \frac{1}{n+1}Z_{n+1} \tag{2.11}$$

$$= \bar{Z} - \frac{1}{n+1}\bar{Z} + \frac{1}{n+1}Z_{n+1}$$

$$= \bar{Z} + \frac{1}{n+1}(Z_{n+1} - \bar{Z})$$

$$= \hat{Z}_n(\ell+1) + \frac{1}{n+1}[Z_{n+1} - \hat{Z}(1)] \tag{2.12}$$

เพราะว่า

$$\hat{Z}_n(\ell+1) = \hat{Z}(1) = \bar{Z} \tag{2.13}$$

จากสมการที่ (2.10) สรุปได้ว่าค่าพยากรณ์ที่ ℓ หน่วยเวลาล่วงหน้า เมื่อเริ่มพยากรณ์ $n + 1$ หรือ $\hat{Z}_{n+1}(\ell)$ คือ ผลรวมเชิงเส้นของค่าพยากรณ์เมื่อเริ่มต้นพยากรณ์ที่ n และค่าสังเกตที่เวลา $n + 1$ ถ้า

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

พิจารณาจากสมาการที่ (2.13) สรุปได้ว่าค่าพยากรณ์ $\hat{Z}_{n(\ell+1)}$ คือค่าพยากรณ์ $\hat{Z}_{n+1(\ell)}$ ที่ปรับด้วยความคลาดเคลื่อนของค่าพยากรณ์ที่เวลา $n+1$ ปัจจุบันและความคลาดเคลื่อนของค่าพยากรณ์ ณ เวลานี้เท่านั้น ในกรณีที่ผู้วิเคราะห์ยอมรับว่าค่า β ในสมการที่ (2.13) มีการเปลี่ยนแปลงไปอย่างช้า ๆ ตามเวลาหรือ β คงที่เฉพาะแห่ง ผู้วิเคราะห์อาจให้ความสนในกับข้อมูลที่เป็นปัจจุบันมากกว่าในการพยากรณ์ แนวทางที่ทำได้ คือ จากข้อมูล $Z_1, \ldots Z_n$ ใช้ค่าเฉลี่ยจากข้อมูลจำนวณ N ค่าสุดท้าย ($N < n$) เป็นค่าพยากรณ์ของ $\hat{Z}_{n+1}$ และเมื่อได้ข้อมูลใหม่มาจะทิ้งข้อมูลเก่าที่สุดไปหนึ่งค่า แล้วหาค่าพยากรณ์โดยใช้ข้อมุลใหม่แทน ด้วยวิธีนี้จำนวนข้อมูลในการเฉลี่ยทุกครั้งจะไม่เปลี่ยนแปลง (เท่ากับ) ขณะที่เวลาเปลี่ยนไป วิธีการพยากรณ์เช่นนี้เรียกว่า วิธีเฉลี่ยเคลื่อนที่ ดังนั้นจึงหาค่าพยากรณ์ที่ละหนึ่งหน่วยเวลาด้วยวิธีค่าเฉลี่ยเคลื่อนที่ ได้ดังนี้

ให้ $\hat{Z}_{n+1} = \hat{Z}_n(1)$ แทนค่าพยากรณ์ที่เวลา $n+1$ เมื่อพยากรณ์ ณ เวลา n

N เป็นจำนวนค่าสังเกตในการเฉลี่ยเคลื่อนที่

Z_n แทนค่าที่สังเกตที่เวลา n

$$\hat{Z}(1) = \frac{1}{N}(Z_n + Z_{n-1} + \cdots + Z_{n-N+1}) \tag{2.14}$$

จากสูตรข้างต้นเมื่อพิจารณาในรูปของการประมาณค่า β อาจกล่าวได้ว่าค่าประมาณของ β ที่หน่วยเวลา t ใด ๆ จะได้มาจากข้อมูลจำนวน N ค่าย้อนหลังจากเวลา t ซึ่งจะได้ค่าประมาณของ β ที่หน่วยเวลาต่างกันมีค่าต่างกันตามลักษณะของการเฉลี่ยเคลื่อนที่ตามลำดับจนถึงค่าสุดท้ายจะได้

$$\hat{\beta} = (Z_n + Z_{n-1} + \cdots + Z_{n-N+1}) \tag{2.15}$$

ซึ่งจะเห็นว่าสอดคล้องกับการที่ยอมรับว่าค่า β เปลี่ยนไปอย่างช้า ๆ ตามเวลา ฉะนั้นถ้าเลือกใช้จำนวนค่าสังเกตในการเคลื่อนที่เท่ากับ N ค่าพยากรณ์ค่าแรกสุดที่จะได้คือ $\hat{Z}_{N+1}$ ค่าที่สองที่หาได้คือ $\hat{Z}_{N+2}$ ตามลำดับไปจนถึงค่าสุดท้ายคือ $\hat{Z}_{N+1}$ ดังต่อไปนี้

$$\hat{Z}_{N+1} = \frac{1}{N}(Z_N + \cdots + Z_1)$$

$$\hat{Z}_{N+2} = \frac{1}{N}(Z_{N+1} + \cdots + Z_2)$$

$$\vdots$$

$$\hat{Z}_{n+1} = \hat{Z}(1) = \frac{1}{N}(Z_n + Z_{n-1} + \cdots + Z_{n-N+1})$$

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

จะได้สมการ

$$\hat{Z}_{n+1} = \frac{1}{N}(Z_n + Z_{n-1} + \cdots + Z_n + Z_{n-N+1}) \quad (2.16)$$

2.4.5 แบบจำลองการปรับเรียบแบบเอ็กซ์โปเนนเชียล (Exponential Smoothing) วิธีการปรับเรียบแบบเอ็กซ์โปเนนเชียล เป็นวิธีการพยากรณ์ที่เหมาะกับการพยากรณ์ในระยะสั้นและปานกลาง โดยวิธีนี้เป็นการพยากรณ์โดยคำนึงอิทธิพลของข้อมูลในอดีตที่แตกต่างกันไปตามเวลา กล่าวคือ การพยากรณ์โดยวิธีนี้จะให้ความสำคัญหรือค่าถ่วงน้ำหนักของข้อมูลในปัจจุบันมากกว่าข้อมูลในอดีตที่ห่างไกลโดยอาจสรุป ได้ว่าเมื่อเวลาเปลี่ยนแปลงไปข้อมูลในปัจจุบันจะมีอิทธิพลต่อค่าพยากรณ์มากกว่าข้อมูลในอดีต โดยใช้วิธีการปรับเรียบแบบเอ็กซ์โปเนนเชียล เป็นวิธีการทำให้เรียบอย่างง่าย โดยอาศัยการหาค่าเฉลี่ยแบบถ่วงน้ำหนัก และสมมติให้ค่าน้ำหนักหรือค่าความสำคัญของข้อมูล คือ α (Alpha) อย่างไรก็ตามวิธีนี้มีเงื่อนไขที่สำคัญคือ ข้อมูลที่ใช้ในการวิเคราะห์จะต้องไม่มีอิทธิพลจากแนวโน้ม (Trend) และ การผันแปรตามฤดูกาล (Seasonal Variations) ซึ่งหมายความว่าข้อมูลที่นำมาใช้ต้องมีลักษณะคงที่ โดยพิจารณาจากการเปลี่ยนแปลงไปอย่างช้าตามเวลาที่เปลี่ยนแปลงไป หรือกล่าวได้ว่าอนุกรมเวลามีค่าเฉลี่ยคงที่เฉพาะแห่ง เมื่อพิจารณาค่าพยากรณ์จากค่าจริงของ Z_n ในสมการที่ (2.16) จะได้

$$\text{ให้} \quad \hat{Z}_{n(\ell+1)} = \hat{Z}(1) = \bar{Z} \quad (2.17)$$

ฉะนั้นสำหรับการพยากรณ์โดยวิธีการทำให้เรียบแบบเอ็กซ์โปเนนเชียลหนึ่งครั้ง สมการที่ใช้ในการพยากรณ์ที่ ℓ หน่วยเวลาล่วงหน้าเมื่อพยากรณ์ ณ เวลา n (ทุกค่า ℓ) คือ

$$\hat{Z}_{n(\ell+1)} = \alpha Z_n + (1-\alpha)\hat{Z}_{n-1} \quad (2.18)$$

การหาค่าพยากรณ์ Z_n และค่า $Z_{n-1}, Z_{n-2}, Z_{n-3}, \ldots Z_2, Z_1$ ทำได้ตามลำดับ

$$\hat{Z}_{n(1)} = \alpha Z_n + (1-\alpha)\hat{Z}_{n-1}$$

$$Z_{n-1} = \alpha Z_{n-1} + (1-\alpha)\hat{Z}_{n-2}$$
$$\vdots$$
$$\hat{Z}_1 = \alpha Z_1 + (1-\alpha)\hat{Z}_0$$

เมื่อ Z_n คือ สถิติจากการทำให้เรียบ

และ α คือ ค่าคงที่ในการทำให้เรียบ

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

โดยมี Z_0 เป็นค่าเริ่มต้น

โดยที่ $Z_0 = \bar{Z}$

โดยทั่วไปข้อมูลในการพยากรณ์จะมีเพียง n จำนวน การดำเนินการหาค่าพยากรณ์โดยวิธีนี้ จึงต้องกำหนดค่าเริ่มต้น Z_0 และค่าคงที่ในการทำให้เรียบก่อนที่จะปรับค่าพยากรณ์ครั้งแรก การเลือกกำหนดค่า Z_0 และ α อาจพิจารณาจากจข้อสังเกตต่อไปนี้ประกอบการตัดสินใจของผู้พยากรณ์ โดยพิจารณาจากสมการ (2.18)

$$\hat{Z}_n(\ell) = \alpha Z_{n-1} + (1-\alpha)\hat{Z}_{n-1} \tag{2.19}$$

เมื่อทำการพิจารณา ในหนึ่งหน่วย หรือ ℓ หน่วยเวลาล่วงหน้าโดยพยากรณ์จาก $t = n$ ได้แล้ว การพยากรณ์ควรปรับค่าพยากรณ์เพิ่มจากค่าที่สังเกตได้ ที่ n + 1 โดยการปรับค่าพยากรณ์ จากสมาการ (2.19)

$$\hat{Z}_{n+1}(\ell) = Z_{n+1} = \alpha Z_n + (1-\alpha)\hat{Z}_n \tag{2.20}$$

จะได้ค่าพยากรณ์ที่ ℓ หน่วยเวลาล่วงหน้าใด ๆ เมื่อพยากรณ์ ณ เวลา $t = n+1$ ซึ่งคือ

จะเป็น $\hat{Z}_{n+1}(1) = Z_{n+1}$

$$\hat{Z}_{n+1} = \alpha Z_n + (1-\alpha)\hat{Z}_n \tag{2.21}$$

เมื่อ $\hat{Z}_n$ คือ สถิติจากการทำให้เรียบ

และ α คือ ค่าคงที่ในการทำให้เรียบ

โดยมี Z_0 เป็นค่าเริ่มต้น

จากสมการที่ (2.21) จะเห็นได้ว่า ค่าพยากรณ์ใหม่ $\hat{Z}_{n+1}$ คือผลรวมเชิงเส้นของค่าพยากรณ์เดิมและค่าสังเกตที่ได้เพิ่มมานั้นเอง โดยที่ถ้า α มาก น้ำหนักของค่าสังเกตที่เพิ่มมาจะมากและแสดงอิทธิพลของค่าพยากรณ์เดิมที่มีต่อค่าพยากรณ์ใหม่น้อยกว่า ถ้า α ใกล้ 0 แสดงว่าค่าสังเกตใหม่ที่เพิ่มมามีอิธีพลน้อยมากต่อค่าพยากรณ์ใหม่

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

131399

2.5 ทบทวนงานวรรณกรรม

จากการศึกษาจากการพยากรณ์จากข้อมูลต้นแบบจากงานวิจัยอื่น ได้มีการนำเสนอตัวต้นแบบและทฤษฎีที่แตกต่างกันกันออกไปขึ้นอยู่กับตัวข้อมูลต้นแบบที่จะนำมาใช้งาน เช่น

1. การนำเสนอผลงานการพยากรณ์ดัชนีผลผลิตอุตสาหกรรม ก็ได้ให้เหตุผลการวิเคราะห์ข้อมูลจากข้อมูลต้นแบบว่าข้อมูลแต่ละประเภทมีความแตกต่างกัน ทั้งนี้ขึ้นอยู่กับปัจจัยด้ายโครงสร้างและอิทธิพลที่มีผลต่อผลผลิตทางด้านอุตสาหกรรม โดยผลงานนี้จะใช้ แบบจำลองการวิเคราะห์แนวโน้ม (Time Trend) แบบจำลองการปรับเรียบแบบเอ็กซ์โปเนนเชียล (Exponential Smoothing) แบบจำลอง Autoregressive Integrated Moving Average Model (ARIMA) และแบบจำลอง Vector Auto Regressions (VARs)

2. การพยากรณ์ความต้องการพลังไฟฟ้าสูงสุดในเขตภาคกลาง ของประเทศไทย สามารถนำไ ป ใ ช้ประโยชน์ สำหรับการตัดสินใจ วางแผนผลิตพลังงานไฟฟ้า และระบบสายส่ง เพื่อตอบสนองต่อความต้องการพลังงานไฟฟ้าทีมี แนวโน้มเพิ่มมากขึ้นโดยใช้ทฤษฎี วิธีการปรับให้เรียบเอ็กซ์โปเนนเชียลแบบ Holt-Winter การวิเคราะห์ การถดถอยที่ ใช้ตัวแปรดัมมี่วิธีของบอกซ์-เจนกินส์ และการวิเคราะห์การถดถอยแบบฟัซซีที่ใช้ตัวแปรดัมมี่การเลือกวิธีการพยากรณ์ ที่เหมาะสมพิจารณาจากค่า MAPE ที่ต่ำที่สุด ผลการศึกษา พบว่า วิธีการพยากรณ์ ที่ให้รูปแบบที่เหมาะสมที่สุด คือ วิธีการปรับให้เรียบเอ็กซ์โปเนนเชียลแบบ Holt-Winter

3. แบบจำลองการพยากรณ์การใช้ไฟฟ้า โดยใช้วิธีผสมของวิธีการปรับให้เรียบเอกซ์โพเนนเชียลแบโฮลต์-วินเทอร์และขั้นตอนวิธีอาณานิคมผึ้งเทียม ศึกษา การพยากรณ์การใช้ไฟฟ้าล่วงหน้า นำข้อมูลไปใช้ในการ วางแผนบริหารจัดการเกี่ยวกับการผลิตไฟฟ้าเพื่อให้เพียงพอต่อความต้องการในอนาคตได้

4. การพยากรณ์อนุกรมเวลาด้วยตัวแบบผสมระหว่าง ARIMA และ เครือข่ายประสาทเทียม ทำการศึกษาเปรียบเทียบการพยากรณ์อนุกรมเวลาด้วยตัวแบบ ARIMA และ ANN ว่าตันแบบไหนที่ใช้ในการพยากรณ์ราคาปิดของหุ้นในตลาดหลักทรัพย์แห่งประเทศได้ผลมากที่สุด

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

บทที่ 3
วิธีการดำเนินงาน

ในการศึกษาของโครงการนี้ จะใช้ทฤษฎีของอนุกรมเวลา มาประยุกต์ใช้งาน ด้านการทำนายพยากรณ์ เพื่อหาอัตราการใช้งานของสตอเรจในอนาคต โดยจะนำค่าสังเกตจากข้อมูลจริงที่ได้จากการใช้งานสตอเรจในอดีต มาทำการวิเคราะห์ด้วยวิธีของอนุกรมเวลา แล้วนำผลลัพธ์ที่ได้ หาค่าความคลาดเคลื่อนของการพยากรณ์ ถ้าผลลัพธ์ใดมีค่า Mean Square Error (MSE) น้อยที่สุดก็จะนำแบบจำลองนั้น จัดทำเป็นโปรแกรมพยากรณ์สำเร็จรูป เพื่อที่จะนำไปใช้งาน ทำนายอัตราการใช้งานของสตอเรจในอนาคตต่อไป

3.1 ขั้นตอนการคัดเลือกข้อมูล

ในโครงการนี้จะรับข้อมูลการใช้งานสตอเรจจาก ซอฟแวร์ของสตอเรจ ที่เรียกว่า Operation Manager ซึ่งเป็นซอฟแวร์ที่ใช้ควบคุมและดูแลการใช้งานของสตอเรจประเภท NAS (Network Area Storage) ซึ่งเป็นสตอเรจที่ให้บริการ แบบ File-Level ทำการเรียกข้อมูลการใช้งานจากปัจจุบันย้อนหลัง 1 ปี ตั้งแต่วันที่ 14-Sep-2011 ถึงวันที่ 17-Jul-2012 มาทำการวิเคราะห์ ซึ่งจำแนกข้อมูลออกเป็น 3 ประเภทดังนี้ ประเภทที่หนึ่ง อัตราการใช้งานแบบ Written (kb/sec) ประเภทที่สอง อัตราการใช้งานแบบ Read (kb/sec) ประเภทที่สาม อัตราการใช้งานแบบ Throughput (kb/sec)

จากการพิจราณาข้อมูลทั้งสามประเภทพบว่า ชุดข้อมูลการใช้งานแบบ Written มีความเหมาะสมที่จะนำมาทำการวิเคราะห์หาความสัมพันธ์กับหน่วยเวลาที่เปลี่ยนแปลงไป ซึ่งเป็นชุดข้อมูลที่มีพฤติกรรมการใช้งานคงที่ และอัตราการใช้งานมีแนวโน้มเพิ่มขึ้นเรื่อย ๆ เมื่อเวลาผ่านไป และเมื่อเปรียบเทียบความสัมพันธ์กับอัตราการการเรียกใช้ข้อมูลแบบ Read ชุดข้อมูลนี้จะมีค่าเปลี่ยนแปลงตามชุดข้อมูลแบบ Written กล่าวคือเมื่อมีข้อมูลมากขึ้นก็จะมีการเรียกใช้ข้อมูลมากขึ้นเช่นกัน ส่วนอัตราการใช้งานแบบ Throughput ก็จะมีค่าเปลี่ยนแปลงไปตามอัตราการใช้งานของ Written กับ Read

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

ตารางที่ 3.1 ข้อมูลที่ได้จากสตอเรจไอบีเอ็มรุ่น N3600A

Time	Data_read Unit :Kb/sec	Data_written Unit:Kb/sec	Data_throughput Unit:Kb/sec
Wed 14 Sep 2011 15:32	7203.2610	3780.6520	10983.9130
Wed 14 Sep 2011 15:33	8323.5560	13415.7040	21739.2590
Wed 14 Sep 2011 15:34	5232.0000	6616.8670	11848.8670
Wed 14 Sep 2011 15:35	3496.4670	6324.8000	9821.2670
Wed 14 Sep 2011 15:36	2251.0670	3072.3330	5323.4000
Wed 14 Sep 2011 15:37	1983.9330	2745.6000	4729.5330
Wed 14 Sep 2011 15:38	1811.0000	2967.8670	4778.8670
Wed 14 Sep 2011 15:39	2575.8000	2181.0670	4756.8670
Wed 14 Sep 2011 15:40	3884.7330	3279.6670	7164.4000
Wed 14 Sep 2011 15:41	2600.6000	4335.3330	6935.9330
Wed 14 Sep 2011 15:42	2763.5330	3502.8000	6266.3330
Wed 14 Sep 2011 20:30	80450.3330	24980.5330	105430.8670

3.2 ขั้นตอนการเตรียมข้อมูล

เมื่อพิจารณาจากข้อมูลต้นแบบจากอัตราการใช้งานของสตอเรจแบบ Written นั้นจะมีการบันทึกข้อมูลเป็นหน่วยวินาที (kb/s) ตั้งแต่ช่วงเวลา 14-Sep-2011 ถึง 16-Jul-2012 ซึ่งเป็นจำนวนข้อมูลที่ละเอียดมาก เมื่อเทียบกับหน่วยเวลาเป็นวินาที ดังนั้นเพื่อความเหมาะสมในการวิเคราะห์ข้อมูลและการประมวลของโปรแกรมที่ใช้คำนวณ ทำการแปลงชุดข้อมูลที่เป็นหน่วยวินาทีไปเป็นหน่วยวัน แล้วทำการศึกษาว่าในแต่ละวันมีการใช้งานของการบันทึกข้อมูลเป็นอย่างไร โดยการหาค่าเฉลี่ยของชุดข้อมูลดังกล่าว ตั้งแต่ช่วงเวลา 14-Sep-2011 ถึง 16-Jul-2012 เป็นเวลาทั้งหมด 307 วันและอัตราการใช้งานคือ (GB/Day)

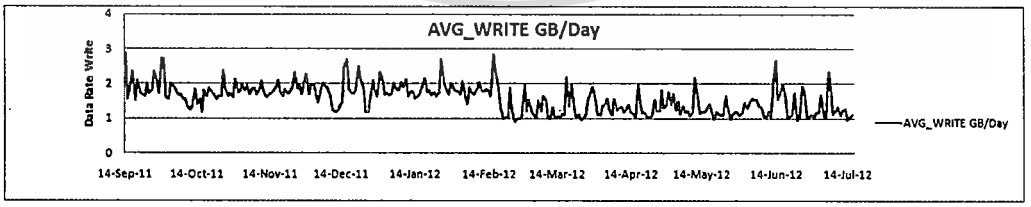
เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

ตารางที่ 3.2 ข้อมูลอัตราการใช้งานสตอเรจในหน่วยวัน (GB/Day) มีจำนวนทั้งหมด 307 วัน

TIME	AVG_WRITE Kb/Day	AVG_WRITE KB/Day	AVG_WRITE GB/Day
14-Sep-11	23605.6031	2950.7004	2.8815
15-Sep-11	12780.5673	1597.5709	1.5601
16-Sep-11	16420.6077	2052.5760	2.0045
17-Sep-11	19174.5779	2396.8222	2.3406
18-Sep-11	12548.3425	1568.5428	1.5318
19-Sep-11	17165.2383	2145.6548	2.0954
11-Jul-12	9061.8834	1132.7354	1.1062
12-Jul-12	9945.2923	1243.1615	1.2140
13-Jul-12	10396.0753	1299.5094	1.2691
14-Jul-12	7785.7896	973.2237	0.9504
15-Jul-12	8466.2289	1058.2786	1.0335
16-Jul-12	9064.5307	1133.0663	1.1065

3.3 ขั้นตอนการดำเนินงาน

ในการศึกษาของโครงงานนี้จะดำเนินการศึกษาโดยจะนำชุดข้อมูลการใช้งานของสตอเรจในหน่วยวัน (GB/Day) มาทำเป็นชุดข้อมูลต้นแบบ เพื่อจะนำไปหาอัตราการใช้งานสตอเตจในอนาคต โดยจะพิจารณาจากข้อมูลว่าจะนำแบบจำลองใดมาทำการวิเคราะห์ โดยทำการนำข้อมูลต้นแบบแสดงเป็นกราฟเพื่อทำการวิเคราะห์หาแบบจำลองที่มีคุณลักษณะที่สอดคล้องกับชุดข้อมูลของการใช้งานสตอเรจ



รูปที่ 3.1 อัตราการใช้งานสตอเรจในหน่วยของวัน (GB/Day)

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

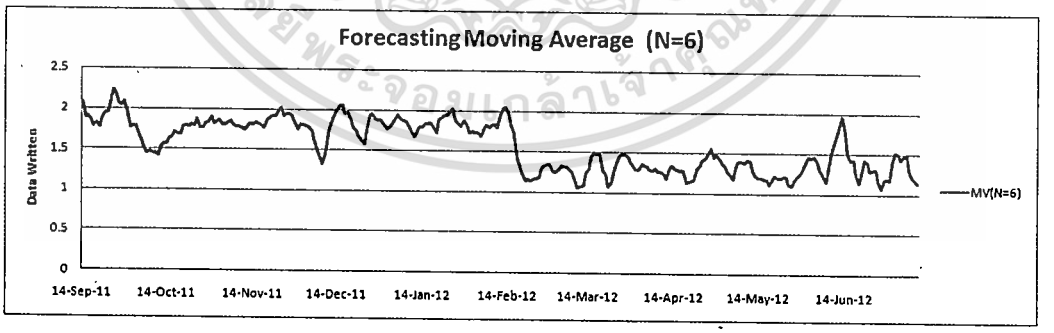
จากการวิเคราะห์ข้อมูลจากกราฟของอัตราการใช้งานสตอเรจ ทำให้ทราบว่าช่วงเวลาในแต่ละคาบของหน่วยเวลาไม่มีผลต่อการใช้งานของสตอเรจและแนวโน้มของอัตราการใช้งานสตอเรจมีค่าคงที่ และชุดข้อมูลไม่มีอิทธิพลจากแนวโน้ม (Trend) และ การผันแปรตามฤดูกาล (Seasonal Variations) เข้ามาเกี่ยวข้อง ซึ่งในทางทฤษฎีของอนุกรมเวลาถ้าชุดข้อมูล เข้ามาเกี่ยวข้องนั้นโดยที่นิยมนำมาใช้วิเคราะห์หาค่าพยากรณ์ในอนาคตนั้นมีด้วยกัน 2 ประเภทคือ ประเภทที่หนึ่ง แบบจำลองการทำให้เรียบแบบง่ายด้วยวิธีค่าเฉลี่ยเคลื่อนที่ (Moving Average Method) สำหรับประเภทที่สอง แบบจำลองการปรับเรียบแบบเอ็กซ์โปเนนเชียล (Exponential Smoothing)

3.3.1 แบบจำลองการทำให้เรียบแบบง่ายด้วยวิธีค่าเฉลี่ยเคลื่อนที่ (Moving Average Method) อนุกรมเวลาที่สอดคล้องตัวแบบนี้โดยทั่วไปคืออนุกรมเวลาที่ไม่มีส่วนประกอบแนวโน้มเรื่องของปัจจัยภายนอกเข้ามาเกี่ยวข้อง และจะมีค่าเปลี่ยนแปลงไปมากน้อยขึ้นอยู่กับรอบระดับค่าเฉลี่ย ซึ่งอาจไม่คงที่ เมื่อเวลาเปลี่ยนแปลงไป

จากสมการที่ (2.16) คือ

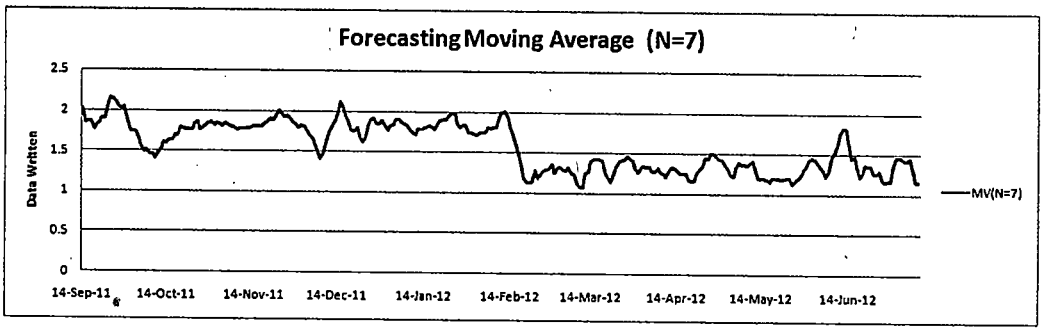
$$\hat{Z}_{n+1} = \frac{1}{N}(Z_n + Z_{n-1} + \cdots + Z_n + Z_{n-N+1})$$

จากสมการ เมื่อเวลาที่เปลี่ยนแปลงไปในช่องเวลา ค่าของ $\hat{Z}_{n+1}$ เมื่อทำการวัดผลค่าในแต่ละช่วงเวลาคงที่โดยทำการทดสอบจากสมการที่ โดยทำการทดสอบแทนค่าของ N เป็นจำนวนค่าสังเกตจากข้อมูลจริงในอดีตในการเฉลี่ยเคลื่อนที่ ซึ่งแบ่งออกเป็นช่วงเวลา N= 6 ถึง N=11 ได้ผลดังนี้

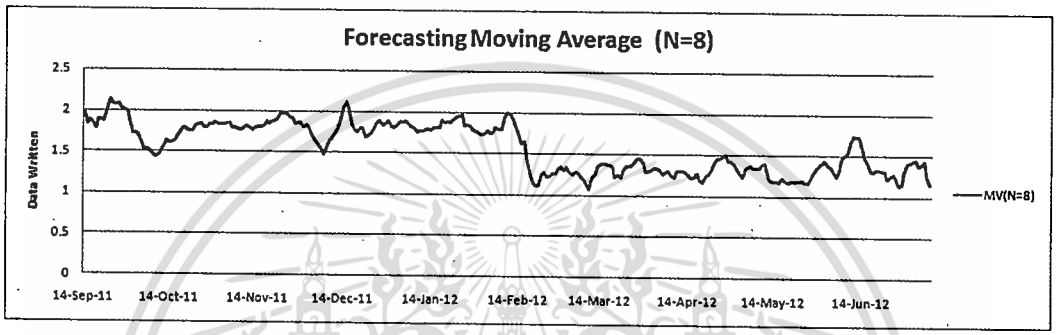


รูปที่ 3.2 การเปลี่ยนแปลงของข้อมูลที่ N= 6

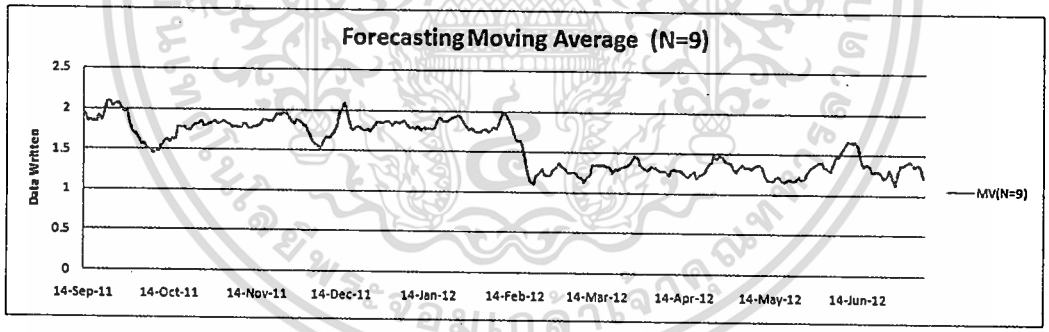
เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้



รูปที่ 3.3 การเปลี่ยนแปลงของข้อมูลที่ N=7



รูปที่ 3.4 การเปลี่ยนแปลงของข้อมูลที่ N=8



รูปที่ 3.5 การเปลี่ยนแปลงของข้อมูลที่ N=9

จากการศึกษาและแทนค่าตามทฤษฎีของแบบจำลองทำให้เรียบแบบง่ายด้วยวิธีค่าเฉลี่ยเคลื่อนที่นั้นทำให้ทราบว่าแต่ละค่าของ N ที่ใช้ในการพยากรณ์นั้นมีผลต่อการเปลี่ยนแปลงของข้อมูลและทำให้ได้ผลการพยากรณ์ของวันที่ 17-Jul-2012 ได้ผลแตกต่างกัน

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

ตารางที่ 3.3 ค่าของการพยากรณ์ของวันที่ 7/17/2012

Forecasting Moving Average (MV)						
ค่าคงที่ N	**6 Day**	**7 Day**	**8 Day**	**9 Day**	**10 Day**	**11 Day**
ค่าพยากรณ์วันที่ 7/17/2012	**1.1133**	**1.1448**	**1.1539**	**1.1490**	**1.2179**	**1.3197**

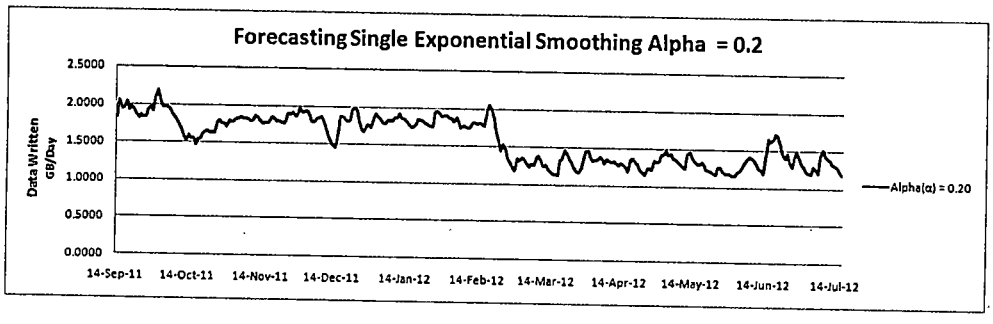
3.3.2 แบบจำลองการปรับเรียบแบบเอ็กซ์โปเนนเชียล (Exponential Smoothing) เป็นวิธีการพยากรณ์ที่เหมาะกับการพยากรณ์ในระยะสั้นและปานกลาง โดยวิธีนี้เป็นการพยากรณ์โดยคำนึงอิทธิพลของข้อมูลในอดีตที่แตกต่างกันไปตามเวลา กล่าวคือ การพยากรณ์โดยวิธีนี้จะให้ความสำคัญหรือค่าถ่วงน้ำหนักของข้อมูล ในปัจจุบันมากกว่าข้อมูลในอดีต สรุปได้ว่าเมื่อเวลาเปลี่ยนแปลงไปข้อมูลในปัจจุบันจะมีอิทธิพลต่อค่าพยากรณ์มากกว่าข้อมูลในอดีต โดยแบบจำลองที่ใช้ในการพยากรณ์แบบวิธีการทำให้เรียบอย่างง่ายคือ แบบจำลองการปรับเรียบแบบเอ็กซ์โปเนนเชียล โดยอาศัยการหาค่าเฉลี่ยแบบถ่วงน้ำหนัก ถ้าสมมติให้ค่าถ่วงน้ำหนักหรือค่าความสำคัญของข้อมูล คือ (Alpha) อย่างไรก็ตามวิธีนี้มีเงื่อนไขที่สำคัญคือ ข้อมูลที่ใช้ในการวิเคราะห์จะต้องไม่มีอิทธิพลจากแนวโน้ม (Trend) และ การผันแปรตามฤดูกาล (Seasonal variations) ซึ่งหมายความว่าข้อมูลที่นำมาใช้ต้องมีลักษณะคงที่ สำหรับสมการที่ใช้ในการพยากรณ์แสดงได้ดังนี้
จากสมการที่ (2.21) คือ

$$\hat{Z}_{n+1} = \alpha Z_n + (1-\alpha)\hat{Z}_n$$

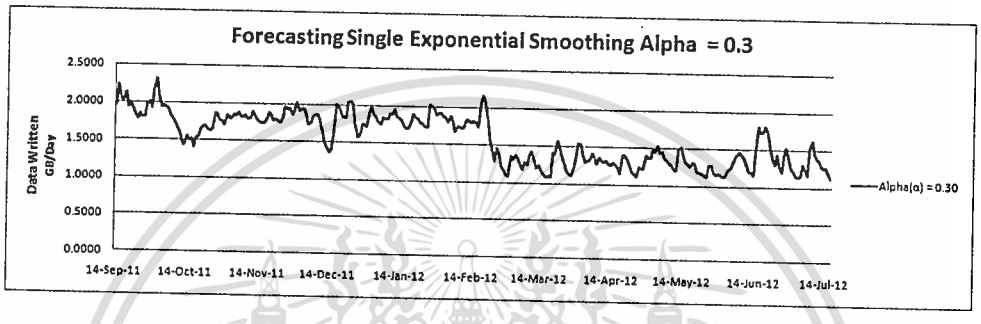
เมื่อ $\hat{Z}_{n+1}$ คือ ค่าสังเกตุของข้อมูล
เมื่อ $\hat{Z}_n$ คือ สถิติจากการทำให้เรียบ
เมื่อ Z_n คือ สถิติจากค่าในอดีต
และ α คือ ค่าคงที่ในการทำให้เรียบ

จากสมการ เมื่อทำการแทนค่าและทดสอบความสารถของค่าถ่วงน้ำหนักที่ $0 \leq \alpha \leq 1$ โดยเริ่มจาก $\alpha = 0.1$ จนถึง 1 โดยเพิ่มจำนวนที่ละ 0.1 แล้วบันทึกค่าที่ได้ ทำให้ทราบว่าแต่ละค่าของ Alpha ที่ใช้ในการพยากรณ์นั้นมีผลต่อการเปลี่ยนแปลงของข้อมูลดังรูป

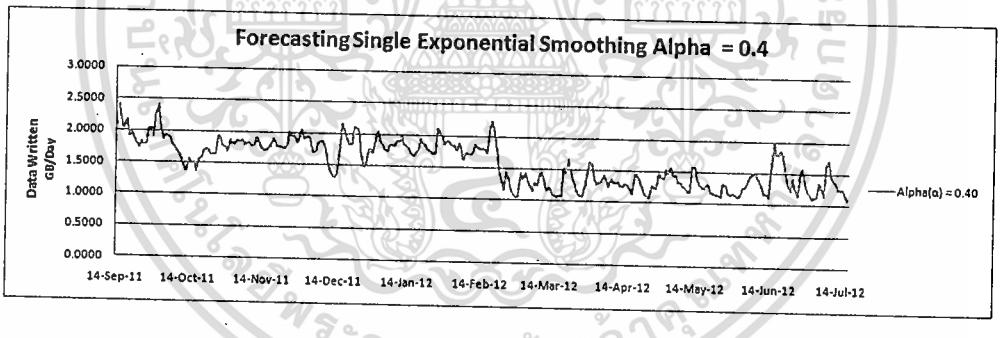
เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้



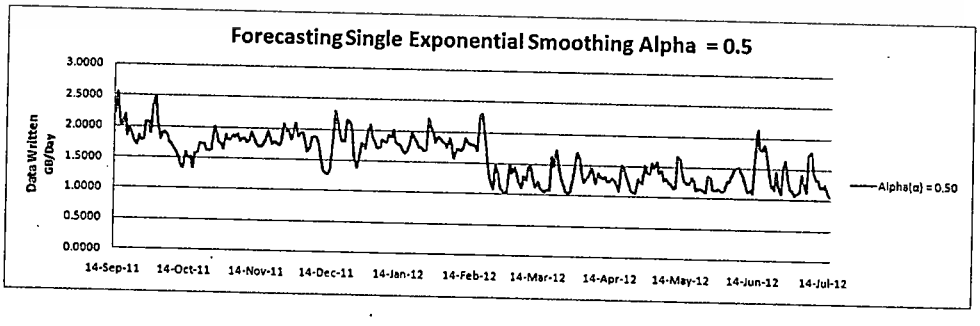
รูปที่ 3.6 การเปลี่ยนแปลงของข้อมูลที่ Alpha=0.2



รูปที่ 3.7 การเปลี่ยนแปลงของข้อมูลที่ Alpha=0.3

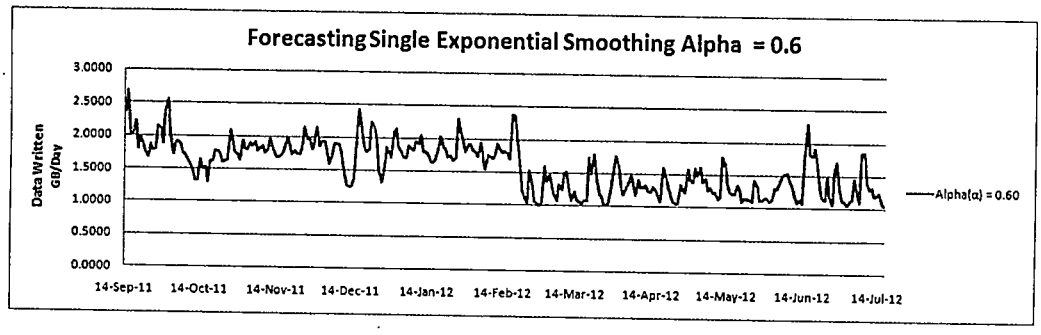


รูปที่ 3.8 การเปลี่ยนแปลงของข้อมูลที่ Alpha=0.4

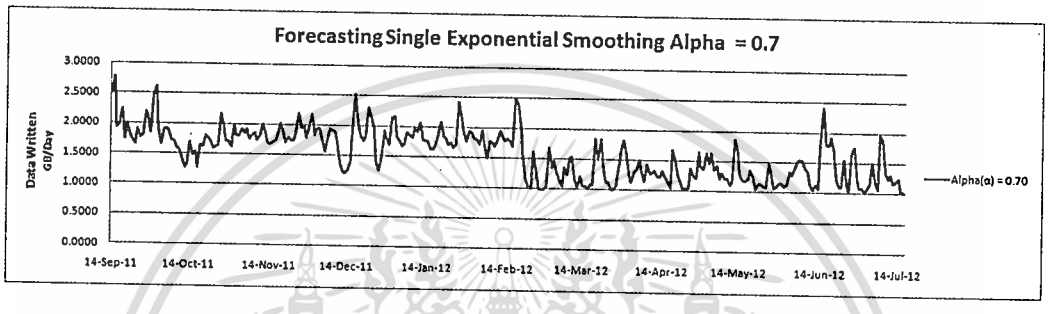


รูปที่ 3.9 การเปลี่ยนแปลงของข้อมูลที่ Alpha=0.5

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้



รูปที่ 3.10 การเปลี่ยนแปลงของข้อมูลที่ Alpha=0.6



รูปที่ 3.11 การเปลี่ยนแปลงของข้อมูลที่ Alpha=0.7

จากการศึกษาและแทนค่าตาม แบบจำลองการปรับเรียบแบบเอ็กซ์โปเนนเชียล นั้นทำให้ทราบว่าแต่ละค่าถ่วงน้ำหนัก ที่ใช้ในวิเคราะห์ข้อมูลนั้นมีผลต่อการเปลี่ยนแปลงของข้อมูลและทำให้ได้ผลการพยากรณ์ของวันที่ 17-Jul-2012 ได้ผลที่แตกต่างกัน

ตารางที่ 3.4 ผลการปรับค่าค่าถ่วงน้ำหนัก (Alpha)

Forecasting Single Exponential Smoothing (SES)							
ค่าคงที่	α =0.1	α =0.2	α =0.3	α =0.4	α =0.5	α =0.6	α =0.70
ค่าพยากรณ์วันที่ 7/17/2012	1.2460	1.2525	1.2580	1.1241	1.0977	1.0858	1.0827

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

บทที่ 4

ผลการทดลอง

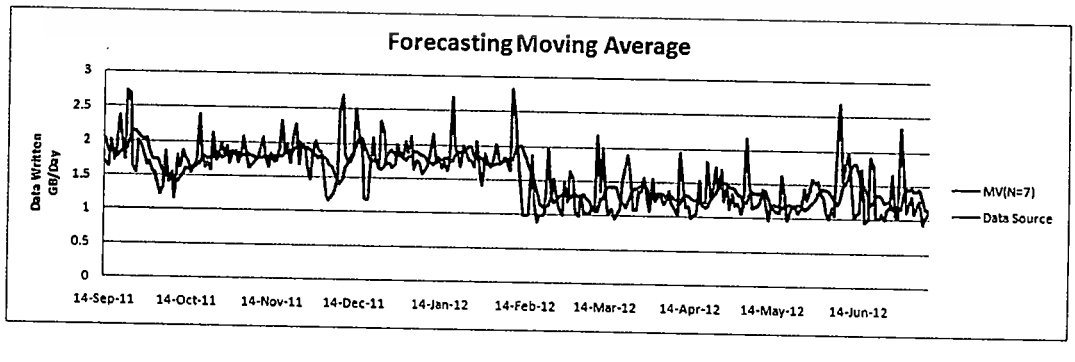
4.1 ผลการทดสอบ อนุกรมเวลา ตามทฤษฎี

โดยทั่วไปแล้ว ผู้พยากรณ์มักเลือกวิธีการพยากรณ์ที่ให้ผลการพยากรณ์ที่แม่นยำที่สุด ซึ่งเป็นเหตุให้มีการเปรียบเทียบว่าแบบจำลองใดเหมาะสมมากกว่ากัน จึงพิจารณาจากวิธีที่ทำให้ผลการพยากรณ์มีความคลาดเคลื่อนน้อยที่สุด โดยการหาผลต่างระหว่างค่าพยากรณ์ล่วงหน้ากับข้อมูลจริง ซึ่งเกณฑ์ที่ใช้ในการเปรียบเทียบวิธีการพยากรณ์ พิจารณาจาก ค่าเปอร์เซ็นต์ความคลาดเคลื่อนสัมบูรณ์เฉลี่ย (Mean Absolute Percentage Error: MAPE) วิธีการพยากรณ์ที่ให้ค่า MAPE ต่ำที่สุด แสดงว่า วิธีการพยากรณ์นั้นมีค่าแม่นยำที่สุด

4.1.1 การพยากรณ์ด้วยแบบจำลอง Moving Average Method จากผลการการศึกษาทำให้ทราบว่า เมื่อแทนค่าตัวแปร N=7 ด้วยวิธีของแบบจำลองการทำให้เรียบแบบง่ายด้วยวิธีค่าเฉลี่ยเคลื่อนที่ จะทำให้ได้ผลการการพยากรณ์ของวันที่ 7/17/2012 และได้ค่าพยากรณ์เป็น 1.1448 GB/Day เมื่อเทียบกับชุดข้อมูลต้นแบบ มีค่าเปอร์เซ็นต์ความคลาดเคลื่อนสัมบูรณ์เฉลี่ยน้อยที่สุด คือ MAPE=5.3501 ดังตารางผลที่ได้จากการวิเคราะห์ข้อมูล

ตารางที่ 4.1 ค่าที่ได้จากการพยากร์จากแบบจำลอง Moving Average Method ที่ N = 7

Forecasting Moving Average (MV)						
ค่าคงที่ N	6 Day	7 Day	8 Day	9 Day	10 Day	11 Day
MSE	0.1208	0.1175	0.1160	0.1152	0.1145	0.1142
MAPE	5.1756	5.0531	5.0773	5.1596	5.2181	5.2421
ค่าพยากรณ์วันที่ 7/17/2012	1.1133	1.1448	1.1539	1.1490	1.2179	1.3197



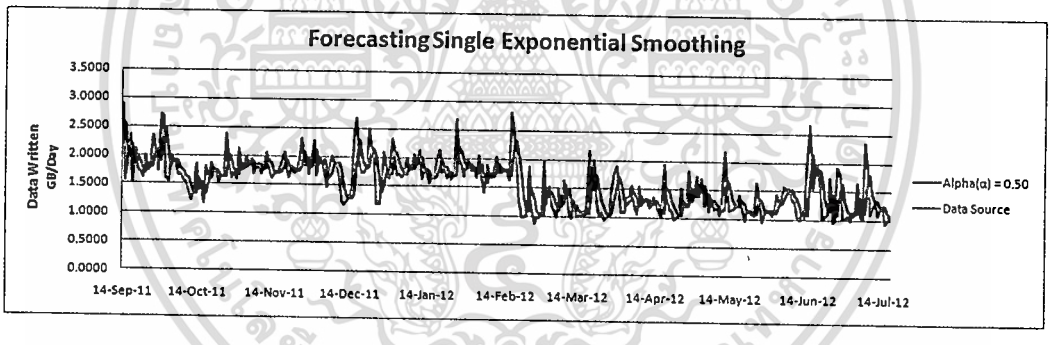
รูปที่ 4.1 การเปรียบเทียบของข้อมูลที่ N=7 กับชุดข้อมูลต้นแบบ

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

4.1.2 การพยากรณ์ด้วยแบบจำลอง (Single Exponential Smoothing) จากผลการแทนค่าสมการ จะเห็นได้ว่าผลการทดสอบที่แทนค่า ค่าถ่วงน้ำหนักที่ $0 \leq \alpha \leq 1$ โดยเริ่มจาก $\alpha = 0.1$ จนถึง 1 โดยเพิ่มจำนวนที่ละ 0.1 ผลการพยากรณ์ที่ได้ เมื่อเทียบกับชุดข้อมูลต้นแบบ จากผลการการศึกษาทำให้ทราบว่า เมื่อแทนค่าด้วยค่าถ่วงน้ำหนักของ$\alpha = 0.5$ ใน แบบจำลองการปรับเรียบแบบเอ็กซ์โปเนนเชียล จะทำให้ได้ผลการการพยากรณ์ของวันที่ 7/17/2012 และได้ค่าพยากรณ์เป็น 1.0977 GB/Day เมื่อเทียบกับชุดข้อมูลต้นแบบ มีค่าเปอร์เซ็นต์ความคลาดเคลื่อนสัมบูรณ์เฉลี่ย น้อยที่สุด คือ MAPE=4.4574 ดังตารางแสดงผลที่ได้จากการวิเคราะห์ข้อมูล

ตารางที่ 4.2 ค่าการพยากรณ์แบบจำลอง Single Exponential Smoothing ที่ $\alpha = 0.5$

Forecasting Single Exponential Smoothing							
ค่าคงที่	α =0.1	α =0.2	α =0.3	α =0.4	α =0.5	α =0.6	α =0.70
MSE	0.1114	0.1273	0.1286	0.1114	0.1132	0.1155	0.1185
MAPE	4.9186	4.9475	4.9519	4.4772	4.4574	4.4963	4.5821
ค่าพยากรณ์วันที่ 7/17/2012	1.2460	1.2525	1.2580	1.1241	1.0977	1.0858	1.0827



รูปที่ 4.2 การเปรียบเทียบของข้อมูลที่ Alpha=0.5 กับ ชุดข้อมูลต้นแบบ

จากผลการพยากรณ์ที่ได้จากแบบจำลองการทำให้เรียบแบบง่ายด้วยวิธีค่าเฉลี่ยเคลื่อนที่ และจากแบบจำลองการปรับเรียบแบบเอ็กซ์โปเนนเชียล พบว่าค่าของการพยากรณ์ที่ได้ของวันที่ 7/17/2012 ผลเปอร์เซ็นต์ความคลาดเคลื่อนสัมบูรณ์เฉลี่ยน้อยที่สุด คือ แบบจำลองการปรับเรียบแบบเอ็กซ์โปเนนเชียล

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

ตารางที่ 4.3 ผลลัพธ์ที่ได้จากการพยากรณ์ของ MV และ SES ของวันที่ 7/17/2012

แบบจำลอง	ค่าพยากรณ์ วันที่ 7/17/2012	MAPE
Moving Average (ค่าคงที่ N=7)	1.1448	5.0530
Single Exponential Smoothing (ค่าถ่วงน้ำหนัก α =0.5)	1.0977	4.4574

4.2 ผลการทดสอบ อนุกรมเวลา

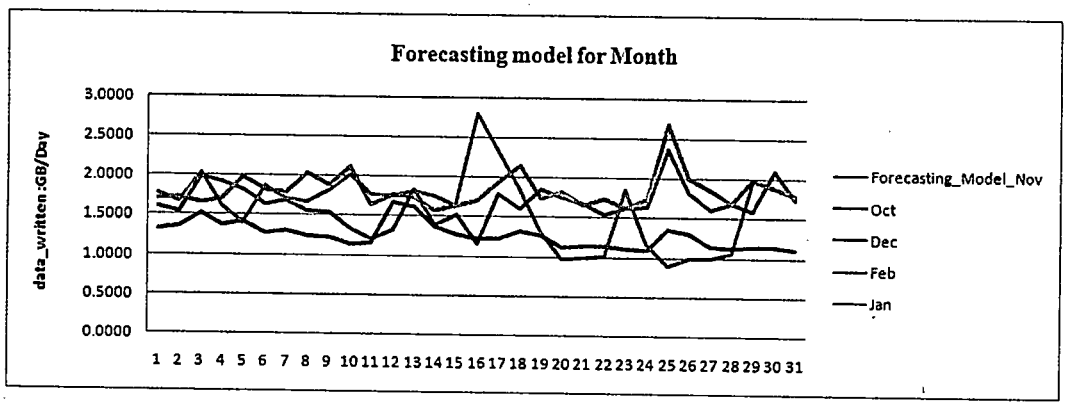
ทำการทดสอบความสามารถในการคัดเลือกชุดข้อมูลที่จะนำมาจัดทำเป็นโมเดลต้นแบบ ซึ่งทางผู้เขียนได้จัดทำเป็นโมเดลต้นแบบแบบรายเดือน ซึ่งใช้ทฤษฎีของคือ แบบจำลองการปรับเรียบแบบเอ็กซ์โปเนนเชียล ที่ใช้ค่าถ่วงน้ำหนัก $\alpha = 0.5$ และได้ผลการทดสอบในการพยากรณ์ของเดือนพฤศจิกายน และมีค่าเปอเซ็นต์ความคลาดเคลื่อนน้อยที่สุด ที่ 1.06%

ตารางที่ 4.4 ผลลัพธ์ที่ได้จากการพยากรณ์ของ SES ที่ $\alpha = 0.5$

Month	MSE	MAPE
Oct,11	0.0710	2.05%
Nov,11	**0.0360**	**1.06%**
Dec,11	0.1356	5.32%
Jan,12	0.0481	1.28%
Feb,12	0.0395	7.27%
Mar,12	0.0745	5.73%
Apr,12	0.0976	2.96%
May,12	0.0045	3.19%
Jun,12	0.0376	7.96%

จากการผลการทดสอบที่ได้จัดทำเป็นโมเดลต้นแบบแบบรายเดือน มาใช้ในการทำนายหาอัตราการใช้งานในอนาคต

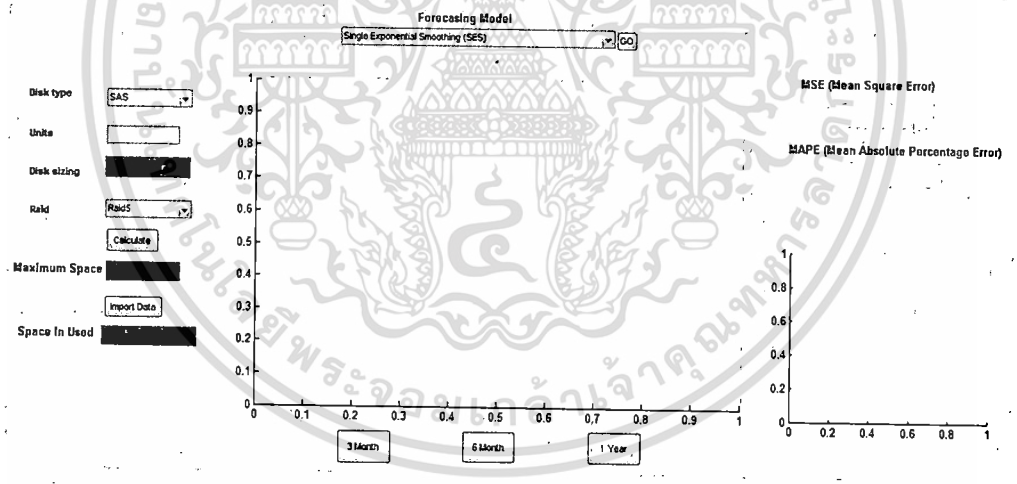
เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้



รูปที่ 4.3 การเปรียบเทียบของข้อมูลรายเดือน กับ ชุดข้อมูลต้นแบบ

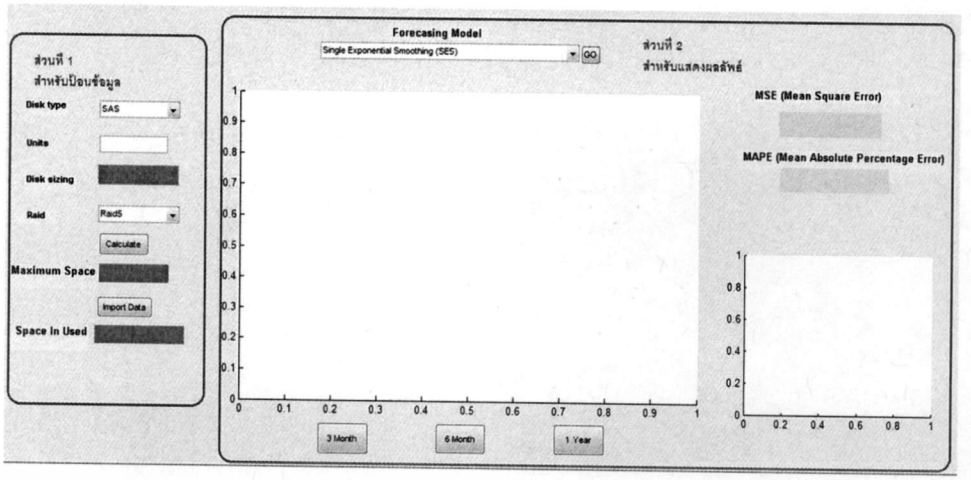
4.3 ตัวอย่างหน้าต่างโปรแกรม และการใช้งาน

ที่ใช้คำนวนหาอัตราการใช้งานสตอเรจในอนาคตที่ได้จากโมเดลต้นแบบ ซึ่งส่วนประกอบการตัวโปรแกรม จะประกอบไปด้วย 2 ส่วน คือส่วนที่หนึ่งใช้สำหรับการป้อนข้อมูลพื้นฐานของการใช้งานสตอเรจ สำหรับส่วนที่สองใช้สำหรับแสดงผลลัพธ์ของการประมวลผล



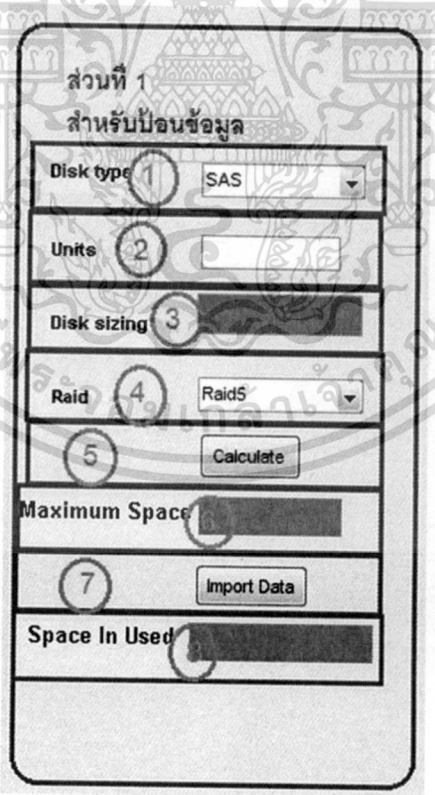
รูปที่ 4.4 โปรแกรมการพยากรณ์โดยใช้วิธีของอนุกรมเวลา

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้



รูปที่ 4.5 ส่วนองค์ประกอบของโปรแกรม

4.3.1 ใช้สำหรับการป้อนข้อมูลพื้นฐานของการใช้งานสตอเรจ ประกอบไปด้วยส่วนประกอบของการนำข้อมูลเข้าระบบเพื่อใช้ในการคำนวณหาพื้นที่ของการใช้งานทั้งหมดของสตอเรจที่จะสามารถรองรับการใช้งานได้



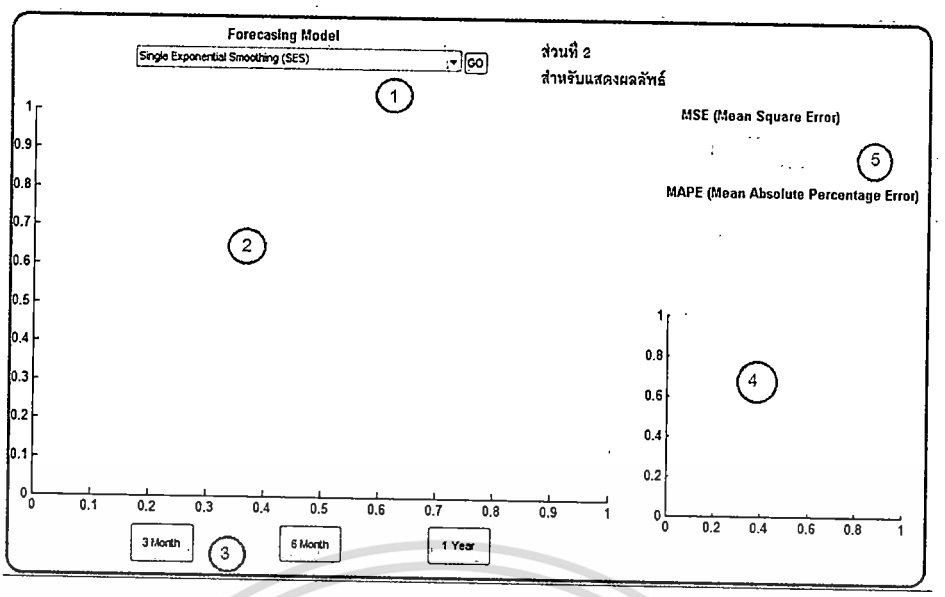
รูปที่ 4.6 การนำข้อมูลเข้า

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

1) Disk type จะใช้ในการตรวจสอบว่า ฮาร์ดดิสที่นำมาใช้งาน เป็นลักษณะประเภทใด และมีความสามารถในการเข้าถึงข้อมูล ซึ่งแต่ละประเภทก็จะให้ผลการตอบสนองต่อการใช้งานที่ต่างกัน
2) Units จะทำหน้าที่รับข้อมูลเข้าระบบว่ามีจำนวนของ ฮาร์ดดิส กี่ลูกที่ใช้งาน เพื่อที่จะได้นำข้อมูลที่ได้ไปทำการคำนวณหาพื้นที่ทั้งหมด ที่จะสามารถใช้งานได้
3) Disk sizing จะทำหน้าที่รับข้อมูลว่าขนาดความจุของของฮาร์ดดิส มีความจุเท่าไหร่ เพื่อที่จะได้นำข้อมูลที่ได้ไปทำการคำนวณหาพื้นที่ทั้งหมด ที่จะสามารถใช้งานได้ต่อไป
4) Raid จะส่วนประกอบที่จะสามารถรับข้อมูลจากการเลือกใช้งาน ของยอมรับความเสี่ยงที่อาจเกิดขึ้นกับข้อมูล และการยอมรับเรื่องความสามารถของ Raid แต่ละชนิด เพื่อให้สอดคล้องกับการใช้งานของแต่ละระบบที่จะนำไปใช้งาน
5) Calculate จะเป็นการนำข้อมูลที่ได้เลือกไว้ทั้งแต่ ข้อ 1) ถึงข้อ 4) มาทำการคำนวณหาผลลัพธ์
6) Maximum space จะเป็นการแสดงผลการคำนวณหาค่าของพื้นที่ ที่สามารถนำไปใช้งานได้
7) Import data จะเป็นการนำเข้าข้อมูลของโมเดลที่จะใช้ในการพยากรณ์ของแต่ละช่วงเวลาที่จะนำมาวิเคราะห์ความเป็นไปได้ของอัตราการใช้งานในอนาคต
8) Space in used จะเป็นการแสดงผลของพื้นที่คงเหลือที่จะสามารถใช้งานได้ในอนาคต

4.3.2 ใช้สำหรับแสดงผลลัพธ์ของการประมวลผล ประกอบไปด้วยส่วนประกอบของการแสดงผลลัพธ์ที่ได้จากวิเคราะห์จากทฤษฎีของการพยากรณ์ของอนุกรมเวลา เพื่อใช้ในการคำนวณหาพื้นที่ของการใช้งานทั้งหมดของสตอเรจที่จะสามารถรองรับการใช้งานในอนาคตได้

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้



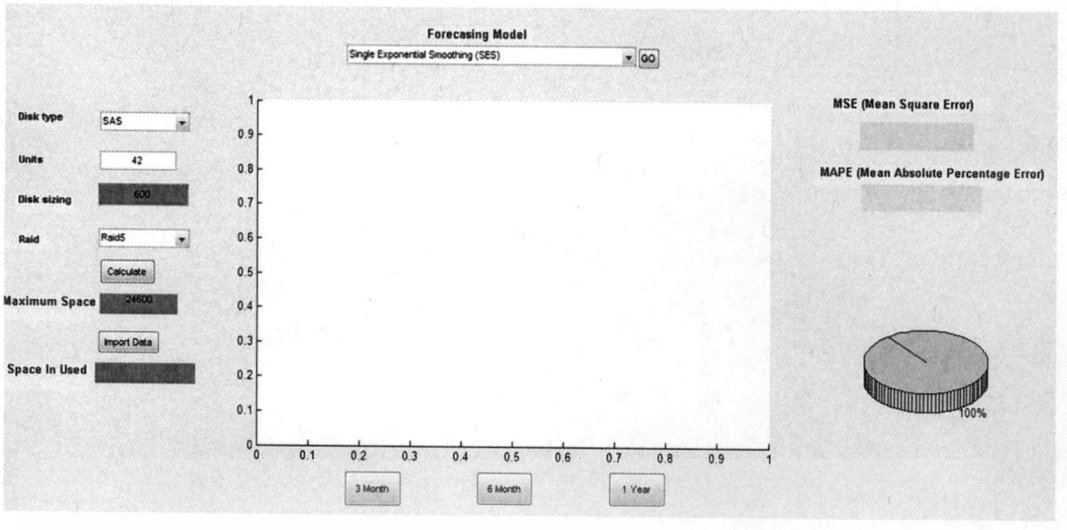
รูปที่ 4.7 การแสดงผลของการพยากรณ์

1) Forecasting Model จะใช้สำหรับการเปรียบเทียบความสามารถของการพยากรณ์ว่าโมเดลต้นแบบกับค่าที่พยากรณ์นั้นมีแนวโน้มการพยากรณ์เป็นอย่างไร
2) Graph 1 ใช้สำหรับการแสดงผลลัพธ์เป็นกราฟ เพื่อง่ายต่อการวิเคราะห์ข้อมูลและการนำไปใช้งาน เปรียบเทียบและวิเคราะห์แนวโน้มของการใช้งานสตอเรจได้
3) Time Line Forecasting ชุดเวลาต้นแบบที่จะใช้ในทำการพยากรณ์ล่วงหน้า ว่าในระยะเวลาดังกล่าวจะใช้พื้นที่ในการใช้งานเท่าไหร่
4) Graph 2 ใช้สำหรับการแสดงผลลัพธ์เป็นกราฟ เพื่อง่ายต่อการวิเคราะห์ข้อมูลและการนำไปใช้งาน เปรียบเทียบและจำนวนพื้นที่ทั้งหมดกับพื้นที่ ที่จะให้งานได้ ในอนาคต
5) MSE and MAPE จะใช้สำหรับการแสดงตัวเลขค่าความคลาดเคลื่อนของการพยากรณ์เมื่อเทียบกับชุดข้อมูลต้นแบบ

4.4 ผลการวิเคราะห์ด้วยโปรแกรมพยากรณ์

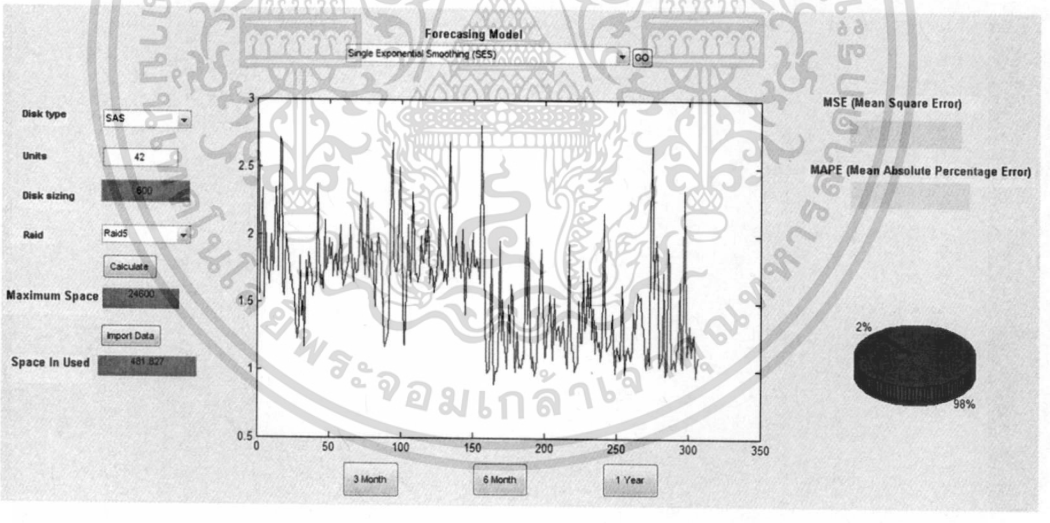
แสดงการนำเข้าข้อมูลจากการบันทึกข้อมูลพื้นฐานของการใช้งานสตอเรจลงไป และทำการคำนวณหาจำนวนพื้นที่สามารถของรับการใช้งานได้

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้



รูปที่ 4.8 ผลลัพธ์ที่ได้จากการคำนวณ

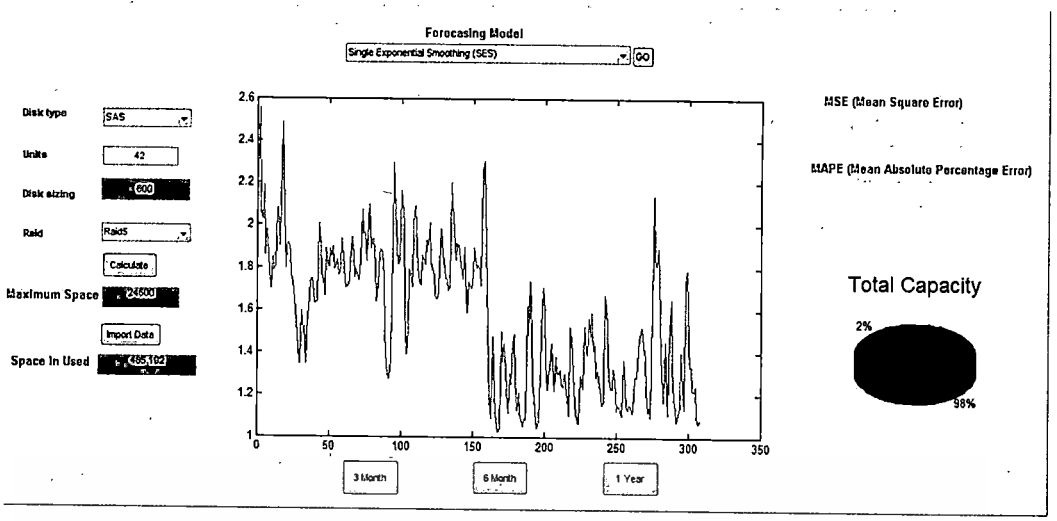
การนำเข้าข้อมูลของการชุดข้อมูลต้นแบบที่นำมาทำเป็นโมเดลต้นแบบ ซึ่งโมเดลนี้จะได้มาจากการแปลงข้อมูลในตอนต้น เพื่อที่จะนำไปใช้คำนวณ พยากรณ์ต่อไป



รูปที่ 4.9 ผลลัพธ์ที่ได้จากการนำเข้าของชุดข้อมูลต้นแบบ

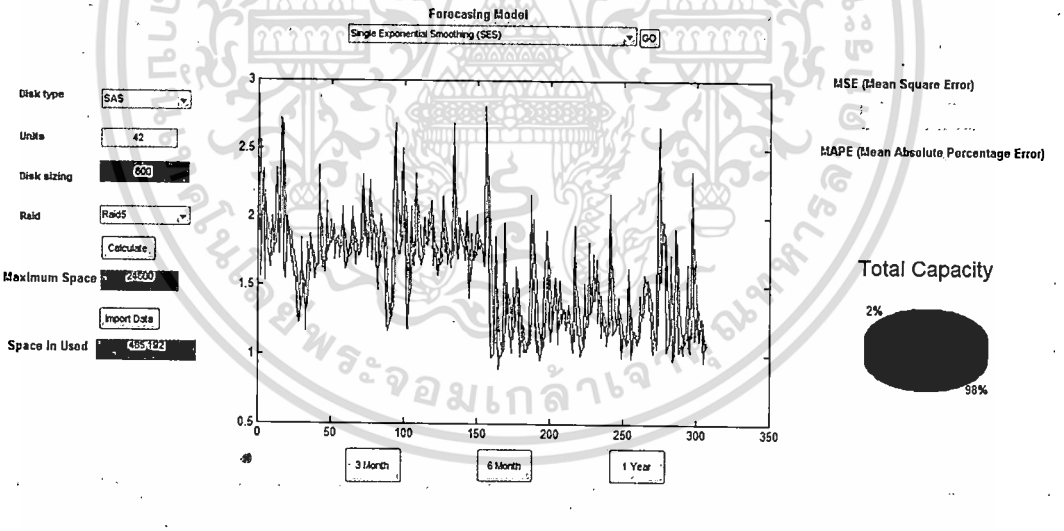
การนำเข้าข้อมูลของการชุดข้อมูลโมเดลต้นแบบ ซึ่งโมเดลนี้จะได้มาจากการพยากรณ์ของแบบจำลองการปรับเรียบแบบเอ็กซ์โปเนนเชียลที่มีค่าคงที่ Alpha = 0.5 เพื่อที่จะนำไปใช้คำนวณพยากรณ์ต่อไป

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้



รูปที่ 4.10 ผลลัพธ์ที่ได้จากการพยากรณ์ของข้อมูลต้นแบบ

ผลการเปรียบเทียบโมเดลต้นแบบ กับ ผลการพยากรณ์ได้ ของแบบจำลองการปรับเรียบแบบเอ็กซ์โปเนนเชียล ที่มีค่าคงที่ Alpha ที่ 0.5 เพื่อที่จะนำไปใช้คำนวณ พยากรณ์ต่อไป



รูปที่ 4.11 ผลการเปรียบเทียบโมเดลต้นแบบ กับ ชุดข้อมูลต้นแบบ

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

บทที่ 5

สรุปผลการทดลอง

จากการศึกษาข้อมูลอัตราการใช้งานของสตอเรจ มาทำการวิเคราะห์ด้วยทฤษฎีของอนุกรมเวลา โดยมีการเปรียบเทียบคุณสมบัติต่างๆ ว่าจะนำทฤษฎีอนุกรมเวลาใดมาทำการวิเคราะห์ข้อมูลนั้น เมื่อเปรียบเทียบกับลักษณะของข้อมูลต้นแบบ เป็นลักษณะข้อมูลแบบคงที่ ซึ่งเป็นชุดข้อมูลที่ไม่ปัจจัยภายนอกใดที่สามารถทำให้ข้อมูลเปลี่ยนแปลงได้ ทำให้ผู้เขียนได้นำ ทฤษฎีการวิเคราะห์ข้อมูลอนุกรมเวลาของแบบจำลองการทำให้เรียบแบบง่ายด้วยวิธีค่าเฉลี่ยเคลื่อนที่ และจากแบบจำลอง แบบจำลองการปรับเรียบแบบเอ็กซ์โปเนนเชียล มาใช้ในการพยากรณ์หาอัตราการใช้งานของสตอเรจในอนาคต

5.1 สรุปผลการศึกษา

จากการศึกษาทฤษฎีของอนุกรมเวลาและวิเคราะห์ข้อมูลการใช้งานของสตอเรจ และได้นำมาจัดทำเป็นโปรแกรมที่จะใช้ในการพยากรณ์หาความสามารถของสตอเรจที่จะรองรับการใช้งานในอนาคต ซึ่งผลการศึกษาจากทดสอบกับเครื่องแม่ข่ายที่มีการเชื่อมต่อใช้งานกับสตอเรจ และใช้ข้อมูลการใช้งานย้อยหลังหนึ่งปีมาเป็นชุดข้อมูลต้นแบบเปรียบเทียบหาเปอร์เซ็นความคลาดเคลื่อน ผลที่ได้มีค่าความคลาดเคลื่อนน้อยมากเมื่อเทียบกับปริมาณการใช้งานพื้นที่ของสตอเรจ ซึ่งสามารถนำผลการพยากรณ์ที่ได้ มาใช้เป็นข้อมูลสนับสนุนในการตัดสินใจของฝ่ายบริหาร เรื่องการจัดสรรงบประมาณ การวางแผน และการบริหารจัดการ การใช้งานสตอเรจในอนาคตได้อย่างมีประสิทธิภาพมากที่สุด

5.2 ประโยชน์ที่ได้รับจากการศึกษาและพัฒนาระบบ

1. ทำให้เข้าใจทฎษฎี และหลักการวิเคราะห์ข้อมูลของอนุกรมเวลา และเทคโนโลยีของสตอเรจ จากการศึกษาการทำงานและวิธีการวิเคราะห์ผล มากขึ้น รวมถึงทำให้ทราบวิธีการวางแผนการดำเนินการ

2. ทำให้ได้โปรแกรม ที่ใช้ในการพยากรณ์พื้นที่ ที่จะสามารถใช้งานของสตอเรจในอนาคต และสามารถนำผลจากการพยากรณ์ที่ได้ไปใช้ประโยชน์ ด้านอื่นต่อไปได้

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

5.3 ข้อเสนอแนะ

เพื่อปรับปรุงโปรแกรมประยุกต์นี้ในอนาคต ผู้เขียนมีความคิดเห็นว่า ในการเพิ่มความน่าเชื่อถือของโปรแกรม จำเป็นต้องใช้ข้อมูลที่เกิดขึ้นจริงในอดีตมาพยากรณ์ และ เป็นข้อมูลที่มีความสมบูรณ์ อีกทั้งเป็นข้อมูลที่ไม่นานมากจนเกินไปเพราะจะทำให้ ผลลัพธ์ที่ใกล้เคียงกับความเป็นจริงมากที่สุด การที่จะทำให้ค่าของการพยากรณ์แม่นยิ่งขึ้นนั้นจำเป็นต้องอาศัยประสบการณ์และความชำนาญของผู้พยากรณ์ในเรื่องนั้น ๆด้วย จึงจะทำให้ค่าของการพยากรณ์นั้นแม่นยำมากยิ่งขึ้นด้วย

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

บรรณานุกรม

มุกดา แม้นมินทร์. 2549, **อนุกรมเวลาและการพยากรณ์**, กรุงเทพฯ: สำนักพิมพ์ประกายพรึก

ดร.วิชิต หล่อจีรชุณห์กุล, ดรจิราวัลย์ จิตรถเวช. **เทคนิคการพยากรณ์**, กรุงเทพฯ: สำนักงานกิจการโรงพิมพ์ องค์การสงเคราะห์ทหารผ่านศึก

ศุวิล ชมชัยยา. 2552, 10 กุมภาพันธ์, **Storage Area Network (SAN) อีกหนึ่งทางเลือกสำหรับองค์กรในวันนี้และวันหน้า,** Online:

http://www.mvt.co.th/viewarticle.php?cid=3&nid=126&page=2

อดิศัย โทวิชา. 2550, **อนุกรมเวลาและเลขดัชนี**, มหาวิทยาลัยราชภัฏสวนสุนันทา

อนันต์ แก้วร่วมวงศ์, เกียรติ อินทรสุริยวงษ์ และ ภูรินท์ ผสม.

Storage IT MANAGER FRIENDBOOK, หน้า 90-105

Bruce Mackenzie-Low February. 2554, 20 กุมภาพันธ์, **Exploring Windows Storage Technologies: DAS, NAS and SAN-Based Solutions,** Online:

http://www.petri.co.il/das- nas-san-storage-technologies.htm

George Ou. 2550, 16เมษายน **,An introduction to enterprise data storage,** Online:

http://www.techrepublic.com/article/an-introduction-to-enterprise-data-storage/6176456

Nate Derby. 2551, 10กรกฎาคม, **Time Series Forecasting Methods**, Online:

http://www.sas.com/offices/NA/canada/downloads/presentations/Victoria2008/Time.pdf

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

ประวัติผู้เขียน

ชื่อ-นามสกุล	นายมงคล จันทร์เทศ
วัน เดือน ปี เกิด	4 ตุลาคม 2524
ที่อยู่	13/84 หมู่ 3 ต.บางรักพัฒนา อ.บางบัวทอง จ.นนทบุรี 11110
ประวัติการศึกษา	วิทยาศาสตร์บัณฑิต มหาวิทยาลัยนเรศวร
ประสบการณ์ทำงาน	
พ.ศ. 2548 - 2555	บริษัทโปรเฟสชั่นนอลวัน
พ.ศ. 2555 – ปัจจุบัน	บริษัทไอบีเอ็ม

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้