Prediction and Optimization for World University Ranking of Prince of Songkla University

Pattira Jubjaimoh

A Thesis Submitted in Partial Fulfillment of the Requirements for the Degree of Master of Science in Mathematics and Statistics

Prince of Songkla University

2017

Prediction and Optimization for World University Ranking of Prince of Songkla University

Pattira Jubjaimoh

A Thesis Submitted in Partial Fulfillment of the Requirements for the Degree of Master of Science in Mathematics and Statistics

Prince of Songkla University

2017

Thesis Title Prediction and Optimization for World University Ranking of Prince of Songkla University.

Author Miss Pattira Jubjaimoh

Major Program Mathematics and Statistics

Major Advisor

....
(Dr. Klairung Samart)

Co-advisor

....
(Asst. Prof. Dr. Naratip Jansakul)

Examination Committee :

....Chairperson
(Dr. Watchareeporn Chaimongkol)

....Committee
(Dr. Klairung Samart)

....Committee
(Asst. Prof. Dr. Naratip Jansakul)

....Committee
(Dr. Thammarat Panityakul)

The Graduate School, Prince of Songkla University, has approved this thesis as partial fulfillment of the requirements for the Master of Science Degree in Mathematics and Statistics

....
(Assoc. Prof. Dr. Teerapol Srichana)
Dean of Graduate School

This is to certify that the work here submitted is the result of the candidate's own investigations. Due acknowledgement has been made of any assistance received.

.... Signature

(Dr. Klairung Samart)

Major Advisor

.... Signature

(Asst. Prof. Dr. Naratip Jansakul)

Co- advisor

.... Signature

(Miss Pattira Jubjaimoh)

Candidate

I hereby certify that this work has not been accepted in substance for any degree, and is not being currently submitted in candidature for any degree.

....Signature

(Miss Pattira Jubjaimoh)

Candidate

ชื่อวิทยานิพนธ์ การทำนายและการหาค่าที่เหมาะสมที่สุดสำหรับการจัดอันดับมหาวิทยาลัยโลกของมหาวิทยาลัยสงขลานครินทร์

ผู้เขียน นางสาวภัทรทิรา จับใจเหมาะ

สาขาวิชา คณิตศาสตร์และสถิติ

ปีการศึกษา 2559

บทคัดย่อ

การจัดอันดับมหาวิทยาลัยในระดับโลกสะท้อนถึงคุณภาพและศักยภาพของสถาบันการศึกษา ด้วยเหตุนี้ หลายๆสถาบันจึงประสงค์จะพัฒนาปรับปรุงการดำเนินงานด้านต่างๆ เพื่อให้สถาบันของตนอยู่ในอันดับที่ดีในการจัดอันดับ งานวิจัยนี้แบ่งออกเป็น 2 ส่วน ส่วนแรกมีวัตถุประสงค์เพื่อทำนายแนวโน้มที่เป็นไปได้ของคะแนนตัวชี้วัดตามวิธีการจัดอันดับมหาวิทยาลัยโลกในระบบ Times Higher Education (THE) และระบบ Quacquarelli Symonds Ranking (QS) ของมหาวิทยาลัยสงขลานครินทร์ โดยใช้วิธีปรับให้เรียบเอ็กซ์โพเนนเชียล (Exponential Smoothing) วิธีค่าเฉลี่ยเคลื่อนที่ (Moving Average) และตัวแบบ ARIMA เพื่อเปรียบเทียบหาวิธีการพยากรณ์ที่เหมาะสมสำหรับการทำนายผลการดำเนินงานด้านต่างๆ ของมหาวิทยาลัยสงขลานครินทร์ ข้อมูลที่ใช้ในการวิเคราะห์ ได้แก่ ข้อมูลจำนวนบุคลากร จำนวนนักศึกษาเต็มเวลา เป็นต้น ในปี 1994 ถึง 2014 การวิเคราะห์ข้อมูลเริ่มจากการจำแนกข้อมูลเป็น 2 ประเภทคือ ข้อมูลที่มีแนวโน้ม และข้อมูลที่ไม่มีแนวโน้ม สำหรับข้อมูลที่ไม่มีแนวโน้มใช้วิธีพยากรณ์ค่าเฉลี่ยเคลื่อนที่อย่างง่าย (Single Moving Average) และวิธีปรับให้เรียบเอ็กซ์โพเนนเชียลอย่างง่าย (Single Exponential Smoothing) สำหรับข้อมูลที่มีแนวโน้มจะพยากรณ์โดยวิธีค่าเฉลี่ยเคลื่อนที่สองครั้ง (Double Moving Average) และวิธีการปรับให้เรียบแบบโฮลท์ (Holt) ส่วนตัวแบบ ARIMA จะใช้ในการพยากรณ์สำหรับข้อมูลทั้งสองประเภท จากการเปรียบเทียบค่าความคลาดเคลื่อนกำลังสองเฉลี่ย (Mean Square Error) ของการพยากรณ์ด้วยวิธีดังกล่าว พบว่า วิธีค่าปรับให้เรียบอย่างง่าย (Single Exponential Smoothing) เป็นวิธีการพยากรณ์ที่เหมาะสมที่สุดสำหรับข้อมูลประเภทไม่มีแนวโน้ม ในขณะที่วิธีการปรับให้เรียบแบบโฮลท์ (Holt) เป็นวิธีพยากรณ์ที่เหมาะสำหรับข้อมูลประเภทมีแนวโน้ม

งานวิจัยส่วนที่สอง มีวัตถุประสงค์เพื่อหาค่าที่เหมาะสมสำหรับตัวชี้วัดอันได้แก่ สัดส่วนจำนวนบุคลากรกับนักศึกษาเต็มเวลา สัดส่วนของจำนวนการอ้างอิงผลงานวิจัยกับจำนวนบุคลากร สัดส่วนของบุคลากรที่เป็นต่างชาติ และสัดส่วนสัดส่วนของนักศึกษาต่างชาติ ซึ่งมีผลต่อการได้คะแนนรวมตามเกณฑ์การจัดอันดับของระบบ QS สูงที่สุด ตัวชี้วัดทั้ง 4ดังกล่าว เป็นตัวชี้วัดที่

สามารถควบคุมได้โดยสถาบันการศึกษา และเป็นตัวชี้วัดพื้นฐานที่ระบบจัดอันดับส่วนใหญ่ใช้ในกระบวนการจัดอันดับ งานวิจัยส่วนนี้สนใจศึกษาการหาค่าที่เหมาะสมที่สุดของปัญหาที่ลักษณะของฟังก์ชันจุดประสงค์ไม่เป็นฟังก์ชันเชิงเส้น โดยฟังก์ชันจุดประสงค์ถูกสร้างจากการทำนอร์มัลไลเซชัน (Normalization) และการถ่วงน้ำหนัก (Weighting) เงื่อนไขข้อจำกัดของปัญหาถูกกำหนดขึ้นด้วยขอบเขตที่แตกต่างกัน 3 กรณี และผลการวิจัยสรุปได้ว่า ค่าที่เหมาะสมสำหรับการให้ค่าสูงสุดของฟังก์ชันจุดประสงค์นั้นขึ้นอยู่กับบริบท สมรรถภาพ และนโยบายที่แตกต่างกันของแต่ละสถาบัน

Thesis Title	Prediction and Optimization for World University Ranking of Prince of Songkla University
Author	Miss Pattira Jubjaimoh
Major Program	Mathematics and Statistics
Academic Year	2016

ABSTRACT

World university rankings reflect quality of higher educational institutions. Therefore, there are high competitions among institutions to be in a higher rank. For that reason, most institutions are discussing about how to increase their global rank. Two subprojects were investigated in this thesis.

The first project was to forecast a possible tendency of indicator scores of the Times Higher Education (THE) and the Quacquarelli Symonds Rankings (QS) of Prince of Songkla University (PSU), Thailand. The exponential smoothing, the moving average and the ARIMA techniques have been compared to find out which technique was more appropriate to predict the performance of PSU based on the indicators of the THE and QS. The data such as the number of academic staff, the number of full-time students, etc., from 1994 to 2014 have been used in the analysis. The data were firstly classified into two classes relying on trend and no existing trend. For no-trend series, the single moving average (SMA) and the single exponential smoothing (SES) were chosen to predict the data tendency. For the other class, the double moving average (DMA) and Holt's method were applied. In addition, ARIMA was also used to forecast for both groups. According to the mean squared error (MSE), the SES is the most appropriate technique for the no-trend series, whereas the Holt’s method is suitable for the trend series.

The second project was to find optimal values for each of our studied indicators: faculty students ratio, citations per faculty, proportion of international faculty, and proportion of international students, that maximizes the overall score of QS Ranking. Those four indicators are commonly used in most university ranking systems and considered to be controllable. An approach of optimization using maximization of nonlinear programming problem in which the objective function was constructed from normalization and weighting was applied throughout this research. Three cases of constraints that are different in boundary determination were considered. The results from the analysis showed that the optimal values were varied depending on the constraints. The final decision for the optimal values is based on context, ability and policy of an individual educational institution.

Acknowledgements

I would like to express my sincere gratitude to my advisors, Dr. Klairung Samart and Asst. Prof. Dr. Naratip Jansakul, for their initial ideas, guidance, constant instruction, kindness and encouragement which enable me to carry out my study successfully.

My special appreciation is expressed to Dr. Thammarat Panityakul, Dr. Nattada Jibenja and Dr. Watchareeporn Chaimongkol for valuable comments and helpful suggestions.

I wish to thank all my teachers in the Department of Mathematics and Statistics, Prince of Songkla University for sharing their knowledge and support so that I can obtain this Master degree.

I gratefully appreciate my family for their love and encouragement. I would like to thank you me friends for their encouragement and suggestions.

Finally, I would like to thank everyone, who supported me, but I did not mention above.

Pattira Jubjaimoh

Contents

List of Tables

List of Figures

List of Abbreviations

ARIMA	Autoregressive Integrated Moving Average
C	Cyclical
DMA	Double Moving Average
ES	Exponential Smoothing
I	Irregular
LES	Linear Exponential Smoothing
LP	Linear Programming
MA	Moving Average
MAE	Mean Absolute Error
MAPE	Mean Absolute Percentage Error
MSE	Mean Square Error
NLP	Nonlinear Programming
PSU	Prince of Songkla University
QS	Quacquarelli Symonds Ranking
S	Seasonal
SES	Single Exponential Smoothing
SMA	Single Moving Average
T	Trend
THE	Times Higher Education

List of Publications

This thesis contains general summary (introduction, some background information and conclusion) and the following papers which are referred to in the text by their roman number.

Paper I Jubjaimoh, P., Samart, K., Jansakul, N., and Panityakul, T. A Comparison of Forecasting Techniques used to Predict World University Rankings Tendency of Prince of Songkla University. The International Conference on Applied Statistics 2016 "Solving Real World Problems with Statistics" (ICAS 2016), Phuket Graceland Resort & Spa, Phuket, Thailand, July 13-15, 2016, O221-O227.
Proceeding.

Paper II Jubjaimoh, P., Samart, K., Jansakul, N., and Jibenja, N. Optimization for Better World University Ranking.
Manuscript (Submitted).

Acceptance Letter for the Conference Proceedings

The International Conference on Applied Statistics 2016 (ICAS 2016)

June 24th, 2016

Dear Pattira Jubjaimoh, Klairung Samart, Naratip Jansakul, and Thammarat Panityakul,

On behalf of the ICAS 2016 conference organizing committee, we are pleased to inform you that your manuscript entitled “A Comparison of Forecasting Techniques Used to Predict World University Rankings Tendency of Prince of Songkla University” has met the international academic standard of blind peer review. Your full paper will also be published in the official conference proceedings

The conference will be held at Phuket Graceland Resort & Spa, Phuket, Thailand from Wednesday, July 13th to Friday, July 15th, 2016.

Thank you for participating in International Conference on Applied Statistics 2016 (ICAS 2016). We look forward to seeing you.

Sincerely Yours,

Samran Muntup

Associate Professor Samran Muntup
Head of Department of Mathematics and Statistics
Thammasat University
Conference Chairman of ICAS2016

Conference Venue
Phuket Graceland Resort & Spa,
Phuket, Thailand 83150
http://www.phuketgraceland.com/

ICAS 2016: Acceptance Letter

ICAS 2016

ศ. 3/6/2016 18:53

ถึง:PATTIRA JUBJAIMOH <5710220072@email.psu.ac.th>;

Submission Title: A Comparison of Forecasting Techniques used to Predict World University Rankings tendency of Prince of Songkla University
Submission Number: 88
Submission Authors: Pattira Jubjaimoh, Klairung Samart, Naratip Jansakul and Thammarat Panityakul

Dear Pattira Jubjaimoh,

We have recently sent an email with the attached acceptance letter of poster/oral presentation to all authors. If you did not receive an email yesterday, please immediately contact us.

All manuscripts are under the review process and we will soon notify you.

Please ignore this email if you already received an email from us yesterday.

Best Regards,
The ICAS 2016 organizing committee

Scientometrics

Optimization for Better World University Ranking

--Manuscript Draft--

Manuscript Number:	
Full Title:	Optimization for Better World University Ranking
Article Type:	Manuscript
Keywords:	world university ranking; indicator; z-score; nonlinear programming.
Corresponding Author:	Klairung Samart, Ph.D. Prince of Songkla University Hat Yai, Songkhla THAILAND
Corresponding Author Secondary Information:	
Corresponding Author's Institution:	Prince of Songkla University
Corresponding Author's Secondary Institution:	
First Author:	Pattira Jubjaimoh, B.Sc.
First Author Secondary Information:	
Order of Authors:	Pattira Jubjaimoh, B.Sc.
	Klairung Samart, Ph.D.
	Naratip Jansakul, Ph.D.
	Nattada Jibenja, Ph.D.
Order of Authors Secondary Information:	
Funding Information:	
Abstract:	World university rankings are very competitive among educational institutions. Many institutions realize the importance and require to step up to a better rank. The aim of this paper was to find the optimal values for each of our studied indicators: faculty students ratio, citations per faculty, proportion of international faculty, and proportion of international students, that maximizes the overall score of Quacquarelli Symonds World University Ranking. Those four indicators are commonly used in most university ranking systems and considered to be controllable. An approach of optimization using maximization of nonlinear programming problem in which the objective function was constructed from normalization and weighting was applied throughout this research. Three cases of constraints that are different in boundary determination were considered. The results from the analysis showed that the optimal values were varied depending on the constraints. The final decision for the optimal values is based on context, ability and policy of an individual educational institution.

1. Introduction

1.1 Background

Nowadays, the competition among educational institutions has increased rapidly which is obvious from the occurrence of many new world university ranking systems. The origin of world university ranking was the announcement of the American colleges annual ranking by the U.S. News and World Report which was the first educational institutions ranking in the world [21, 33]. Afterwards, the Shanghai Jiao Tong University (ARWU), China was founded as the first world university ranking that was the prototype for many recent systems [3] such as Accreditation Council of Taiwan (HEEACT), University Ranking by Academic Performance (URAP), LEIDEN, SCImago, WEBOMETRICS, Quacquarelli Symonds (QS) and Times Higher Education (THE), etc. These systems focus on different criteria which reflect institutions' quality and reputation in the domestic and international levels [6, 32, 35]. Therefore, they are the keys for an improvement to be in a higher position. As Prince of Songkla University (PSU) is one of the universities that has been ranked by many ranking systems and we want to improve our rank to a better position, hence this was a starting point to conduct this research.

1.2 Objective

Since the position of the universities ranking is emphasized by educational institutions as it reflects quality and potential of each educational institution, the objectives of this thesis are in the two subprojects.

Subproject I: To predict the indicator score trends for PSU based on the methodologies of QS and THE world ranking by using forecasting techniques.

Subproject II: To find the optimal values for the important indicators that maximize the overall score of an institution by using optimization technique on the normalization data process as the first step of most ranking systems.

2. Forecasting

Montgomery et al. [19] stated that forecasting is one of key components of making decision effectively. Since predicting is uncontrollable, so forecasting is a function that can be used for planning and controlling procedure management by using historical data. Planning is a long range forecasting to design and justify the equipment required to achieve the goal whereas controlling is a process which uses variables to predict the process ahead with restriction of the optimal time and actions. The importance of making decision relates with the three elements of time including the forecasting period, the forecasting horizon and the forecasting interval. Firstly, the forecasting period is the time basis unit for forecasting something. For example, when we want to predict the demand weekly then the period of forecasting is a week. Secondly, the forecasting horizon is the number of periods ahead which hold the forecast. For example, we predict demand by week for 10 weeks ahead, then the horizon is 10 weeks and the period is still a week. Lastly, the forecasting interval is the frequency of the new forecast and it is often the same with forecasting period. For example, the sales report is recorded monthly, then the interval can be both monthly and weekly. Moreover, the data using for forecasting is distinguished to two kinds as period data and point data. Period data are indicator of a variable over a period of time such as the average of annual rainfall. Point data are the value of the variable at the definite time points such as the rainfall at noon are point data. A time series is a time-ordered sequence of variable's observations. The time series analysis will use just the time series history data to predict the future values. The variable is observed as discrete time points. It is related with explanation of the phenomenon that occurs by sequence. The time forecasting needs the appropriate model to represent the observations that have several patterns. The characteristic patterns of time series include main four components as described by Chatfield [10];

Trend (T) is defined as long-term change. Nonetheless it may mean to such a long-term waving, so this trend means the number of available observations which are assessed as long-term or trend is comprised all cyclic components that have wave length more than the length of the time series of observations.

Seasonal (S) refers to the data which has many time series in the same data set such as the data of temperature that reports variation in period of a year.

Cyclical (C) is a part of seasonal which presents variation by fixed period due to any other physical motive.

Irregular (I) is the residual which is left after removing the trend and cyclic from data. Thus, it is one of irregular variation of data.

There are several methods for time series forecasting that are considered to use according to the different character of each data series namely differential in series components as T, S, C and I, pattern of data as a stationary or a non-stationary data and so on. **In Paper I**, the research compared some of time series forecasting techniques. The study analyzed the data of PSU which was based on THE and QS indicators and was collected in 1994 to 2014 as well as there was no seasonal effect in the series. There were 12 different indicators to analyse, namely doctoral degree awarded to academic staff, doctoral degree awarded to undergraduate degree awarded, academic staff to students, research income to academic staff, institute income to academic staff, research industry income to academic staff, normalize citation impact, international students to students, international academic staff to academic staff, papers to academic staff, paper with co-author to papers and citation to staff. The study started by classifying the data to trend and no trend by using the run chart and then compared the forecasting techniques, namely the moving average (MA), the exponential smoothing (ES) and the Box-Jenkins.

2.1 Run Chart

A run chart is a non-parametric testing which shows as a line chart of data plotted over time in some type of order. Plotting over time is a basic way to learn for patterns, trends, and variation in data. The principal objective of the run chart is to find out process of degradation or improvement, which will present like non-random patterns in the distribution of data points around the median line. Run chart consists of the horizontal axis, the vertical axis, the median line and line graph. The horizontal axis presents a time scale, whilst the vertical axis is a scale of observations. The median line is normally calculated and applied as the centerline of the chart which divides the data points to half, above and below the median. However, the extreme values in the data do not influence to the median. In brief, the run chart functions are roughly to specify the non-random variation that helps to understand and picture the change over time. There are several non-parametric tests for non-random variation that are available in run charts such as run test, shift test, trend test, etc [1, 24].

Hobai [12] stated that the non-parametric testing is less sensitive to extreme values and without assuming that the data comes after a specific distribution, thus it is an option to the linear trend and nonlinear trend models. For a single time series, the run test is recommended to use for randomness testing because this test is considered to be one of the easiest tests. The run test presents special causes variation which influences the time series, if any. The idea of the test is occurrence of the variation in processes that can be usual as a normal part in the environment or special variation which can come from external factor of the system. Its performance is based on the number of runs around the median or based on the number of runs up or down, which can determine the variations due to mixtures, clustering, oscillations or trends. The test for number of runs around the median is sensitive to mixtures and clustering behavior. Oscillations and trends are indicated by the test for number of runs up or down. Nowadays, the run test is available for many statistical softwares and the software Minitab 16 was applied throughout **Paper I.** However, this study focused specially on the approximate p-value for the trend determination (see Figure 2.1).

The approximate p-value for trend is the test based on the number of runs up or down. It is the number of increasing or decreasing runs which is sensitive to the oscillations and trends behavior. The p-value for trends formula is calculated based on the standard normal distribution as follows.

$$\text{p-value} = \mathrm{cdf}(Z)$$

where cdf(Z) is the cumulative probability to Z that can be calculated with the formula.

$$Z = \frac{r - \mathrm{E(r)}}{\sqrt{\dfrac{16N - 29}{90}}}$$

where r is the number of runs up or down.

N is the sum of x and y.

E(r) is the expected number of runs up and down with the formula

$$\mathrm{E(r)} = \frac{2(x+y)-1}{3}$$

where x is the number of points above center line

and y is the number of points below or equal to center line.

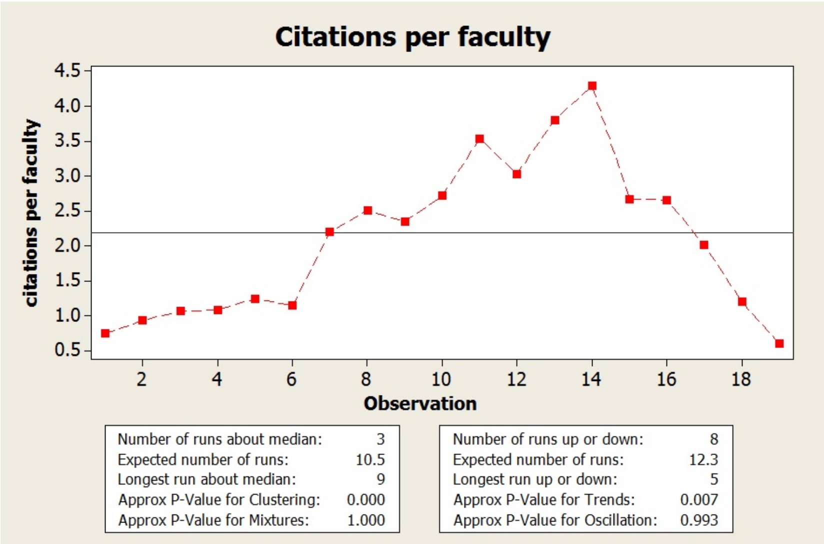

Figure 2.1: The example of run chart result by using Minitab16

All p-values considered in the paper come from the data processing by the Minitab16. If the p-value is less than the significant level of 0.05, it means that the null hypothesis is rejected or there is a trend in data. If it is greater than 0.05, it means to the contrary result.

2.2 Forecasting Techniques

Different forecasting techniques were applied to analyse the data where the trend was determined by the run chart. The 3 main techniques with the 5 sub-techniques were used in **Paper I.** It consisted of the moving average where the simple moving average (SMA) and the double moving average (DMA) were applied for the no-trend and trend series data, respectively. The exponential smoothing where the single exponential smoothing (SES) and the linear exponential smoothing (LES or Holt's method) were used for the no-trend and trend series data, respectively and the Box-Jenkins technique where the autoregressive integrated moving average model (ARIMA) was applied for both type of data. All these techniques have been used to predict the tendency of PSU performances based on indicators of the THE and QS via the Minitab l6 to look ahead for planning the improvement of world university ranking.

2.3 Accuracy Measuring

2.3.1 Mean Square Error (MSE)

MSE is an accepted technique in general for evaluating exponential smoothing and other methods [31]. The equation is:

$$\text{MSE} = \frac{1}{n}\sum_{t=1}^{n}(e_t)^2 = \frac{1}{n}\sum_{t=1}^{n}(Y_t - F_t)^2$$

where $e_t = Y_t - F_t$ is the residual or the actual value (Y_t) minus the forecast value (F_t) for time t

and n is the total number of the time periods.

2.3.2 Root Mean Square Error (RMSE)

RMSE measures error in terms of units that are equal to the original value. It is used mostly in environmental research studies such as meteorology, climate and air quality [9]. The equation is the square root of MSE as:

$$\text{RMSE} = \sqrt{\frac{1}{n}\sum_{t=1}^{n}(e_t)^2} = \sqrt{\frac{1}{n}\sum_{t=1}^{n}(Y_t - F_t)^2}$$

where Y_t is the actual value for time t

F_t is the forecast value for time t

and n is the total number of the time periods.

2.3.3 Mean Absolute Error (MAE)

MAE is also a useful method which is widely used to measure the accuracy in model. Chai and Draxler [9] provided the MAE formula as:

$$\text{MAE} = \frac{1}{n}\sum_{t=1}^{n} |e_t|$$

where $e_t = Y_t - F_t$ is the residual or the actual value (Y_t) minus the forecast value (F_t) for time t

and n is the total number of the time periods.

2.3.4 Mean Absolute Percentage Error (MAPE)

MAPE is similar to MAE in with the percentage term [31]. Its equation is

$$\text{MAPE} = \frac{1}{n}\sum_{t=1}^{n} \frac{|e_t|}{y_t} 100\%$$

where $e_t = Y_t - F_t$ is the residual or the actual value (Y_t) minus the forecast value (F_t) for time t

and n is the total number of the time periods.

Comparing for the most proper forecasting technique, needs to examine the errors or bias of the information as well as to compare accuracy values for the fitted models. Although, there are several measurements to evaluate the model accuracy but there is no any guarantee which measurement is the best one to define the most appropriate forecasting method [23]. However, the most significant evaluation of the quality of forecasting models is considering the smallest value of a goodness of fit indicator. **In Paper I**, the MSE was used to compare the most appropriate forecasting models. Nonetheless, the MAPE is recommended since it is more appropriate for comparison when data series have different appearances.

2.4 Forecasting Results

After using the run chart in Minitab 16 to classify the trend, we found that 5 indicators, namely institute income to academic staff, international academic staff to academic staff, papers to academic staff, paper with co-author to papers and citations to staff, provided the approximate p-values for trend less that 0.05 which implies that trend variation occurred in the series. The rest of 7 indicators from 12 indicators are the no-trend series because the p-values were greater than 0.05. Then, the data were fitted with forecasting models in the same software. The results of forecast for trend series are shown in Figure 2.2. Together with Table 2.1 which presents MAPEs comparisons, the results show that the DMA method provided the smallest MAPE for the academic staff (inter) to staff (total) and the paper with co-author to paper (total). The Holt's method gave a minimum value for the citations to staff and the institute income to staff. The papers (total) to academic staff presented smaller values by using ARIMA(0,2,1) with no constant when compared with the results of the other methods.

Table 2.1: Comparisons of forecasting values and MAPE of the trend series by using DMA, Holt's and ARIMA methods

Indicators	Methods	Actual value	Forecast value			MAPE
		2014	2015	2016	2017	
1. Academic staff(inter) / staff	DMA	0.033521985	0.0376777	0.037677	0.037677	5.416
	Holt		0.0386287	0.039372	0.0401153	11.758
	ARIMA					
2. Papers(total) / staff	DMA	0.27601219	0.260585	0.260585	0.260585	20.092
	Holt		0.295615	0.310122	0.324629	19.247
	ARIMA(0,2,1)		0.290104	0.304196	0.318288	15.273
3. Paper with co-author / papers (total)	DMA	0.356466877	0.370493	0.370493	0.370493	8.154
	Holt		0.317698	0.296698	0.275697	9.741
	ARIMA					

4. Citations / staff	DMA	0.592076622	1.88825	1.88825	1.88825	27.817
	Holt		-0.71119	-1.6981	-2.68512	19.942
	ARIMA(0,2,1)		0.21537	-0.16134	-0.53805	22.414
5. Institute income / staff	DMA	462412.9299	441303	441303	441303	8
	Holt		474306	489525	504744	2
	ARIMA					

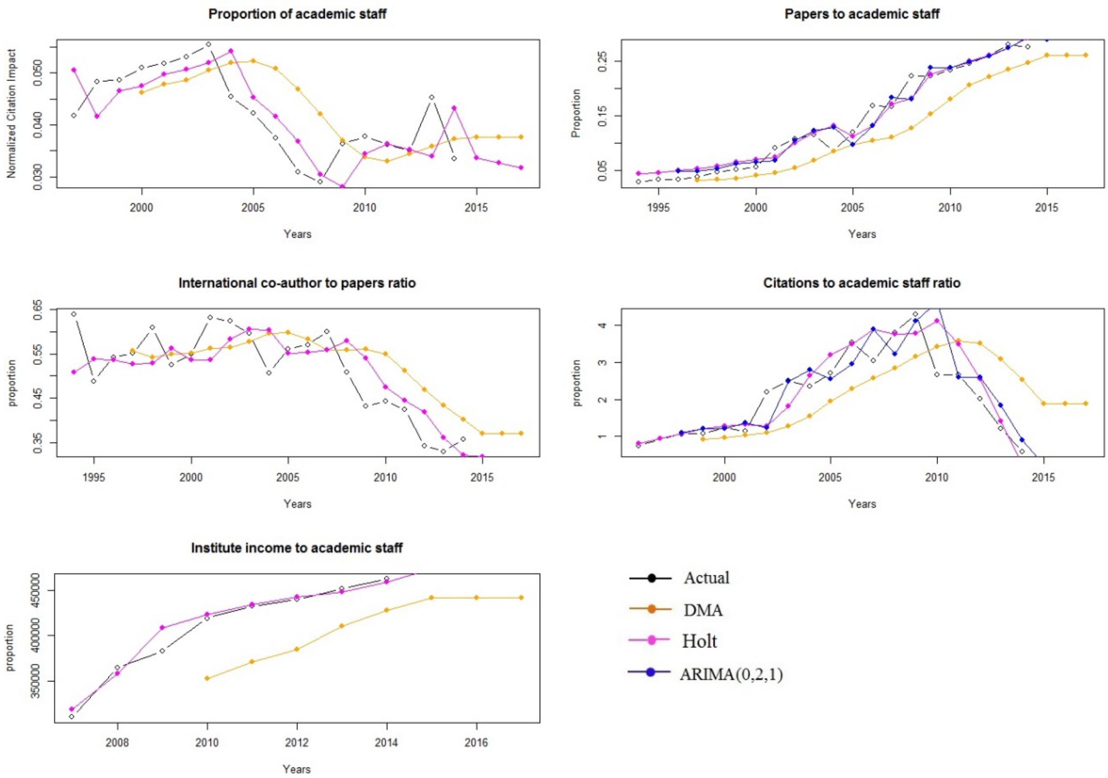

Figure 2.2: Comparisons of forecasting tendency of the trend series by using DMA, Holt's and ARIMA methods.

When consider the results of the no-trend series by using SMA, SES and ARIMA(1,0,0) with no constant from Table 5 in **Paper I,** the academic staff to students, research income to academic staff, research industry income to academic staff and normalized citation impact provided minimum MSEs in SES method when compared with the no-trend results from the other two methods. On the other hand, when consider the results based on the method that provided the smallest MAPE, Table 2.2 shows that the SMA method provided minimum MAPE for the research income industry to staff and the SES method provided minimum MAPE for the research income to staff and normalized citation impact and students (inter) to students (total). Lastly, ARIMA(1,0,0) with no constant provided the smallest MAPE for the doctoral degree awarded to staff, doctoral degree awarded to undergraduate awarded and academic staff to students (total). Figure 2.3 shows that most indicators appear the increasing of the forecast values in 2015 when consider based on the method that gave a minimum value of MAPE.

Table 2.2: Comparisons of forecasting values and MAPE of the no-trend series by using SMA, SES and ARIMA methods

Indicators	Methods	Actual value	Forecast value			MAPE
		2014	2015	2016	2017	
1. Doctoral degree awarded / staff	SMA	0.002612103	0.0028621	0.0028621	0.0028621	137.797
	SES		0.0028660	0.0028660	0.0028660	52.131
	ARIMA(1,0,0)		0.0024165	0.0022356	0.0020682	16.557
2. Doctoral awarded /undergraduate awarded	SMA	0.000742666	0.0007888	0.0007888	0.0007888	209.078
	SES		0.0007887	0.0007887	0.0007887	95.369
	ARIMA(1,0,0)		0.0006801	0.0006228	0.0005704	59.114
3. Academic staff / students (total)	SMA	5.575968655	5.52814	5.52814	5.52814	14.045
	SES		5.58779	5.58779	5.58779	11.866
	ARIMA(1,0,0)		5.60327	5.63071	5.65828	11.776
4. Research income / staff	SMA	109101.4253	117680	117680	117680	16
	SES		110303	110303	110303	14
	ARIMA(1,0,0)		109306	109512	109718	14.056
5. Normalized citation impact	SMA	0.67	1.19667	1.19667	1.19667	33.379
	SES		1.09201	1.09201	1.09201	31.377
	ARIMA					
6. Research income industry / staff	SMA	2597.009578	6611.67	6611.67	6611.67	62
	SES		4948.01	4948.01	4948.01	723
	ARIMA(1,0,0)		1940.37	1449.75	1083.19	280.445
7. Students (inter) / students (total)	SMA	0.021627108	0.0225822	0.0225822	0.0225822	46.786
	SES		0.0227512	0.0227512	0.0227512	46.279
	ARIMA(1,0,0)		0.0216524	0.0216778	0.0217032	52.894

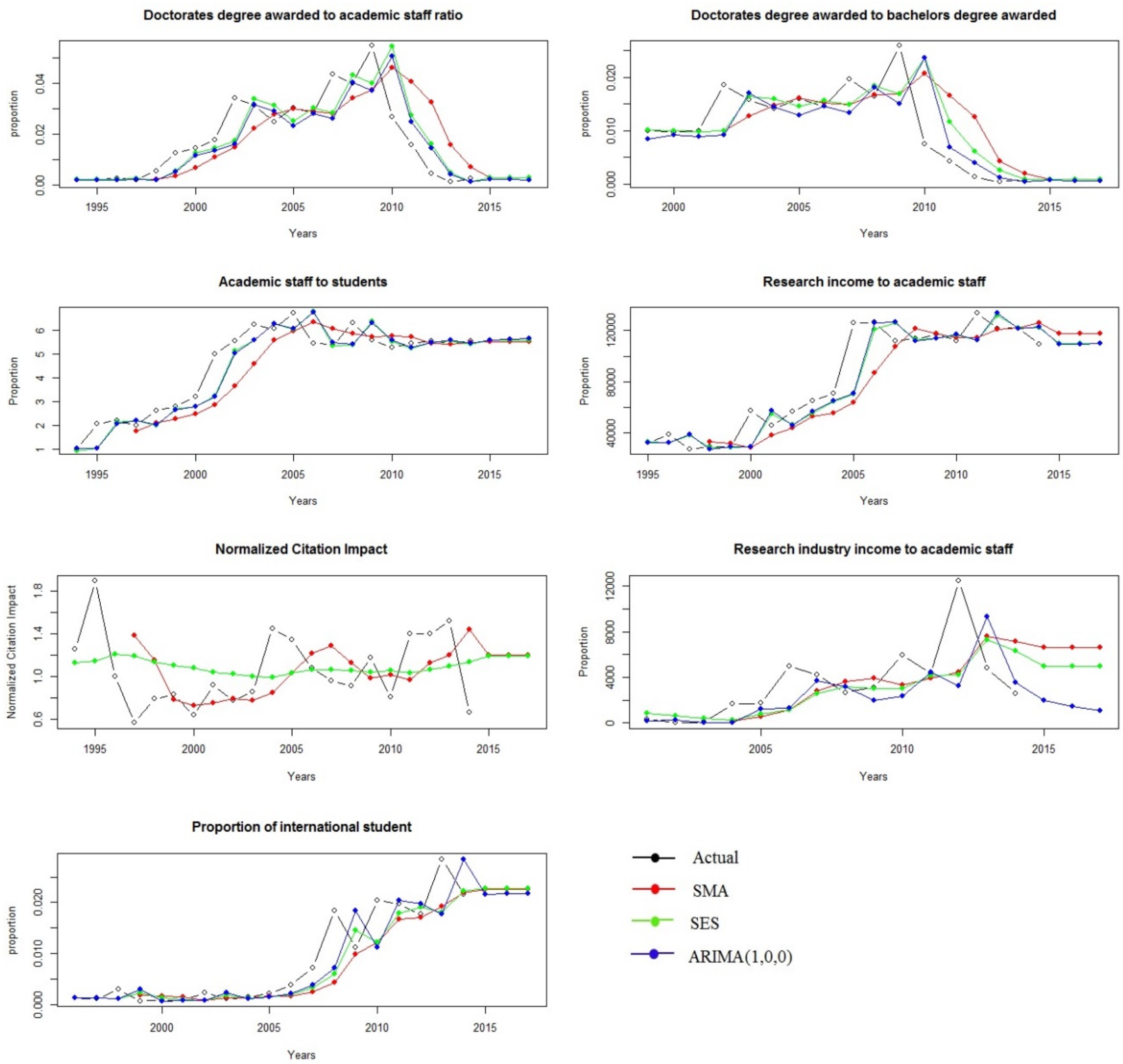

Figure 2.3: Comparisons of forecasting tendency of the no-trend series by using SMA, SES and ARIMA methods.

Furthermore, the ARIMA models for the academic staff (inter) to staff (total), paper with co-author to paper (total) and institute income to staff of trend series in Table 2.1 are ARIMA(0,1,0) with no constant which is random walk with no constant process. This phenomenon happens when taking the first difference of a non-stationary series with mean and variance increasing through the time. From Table 2.2, the normalized citation impact provided ARIMA(0,0,0) which is white noise process that occurs with stationary series. Both random walk and white noise are the processes that all autocorrelations are not significant. Therefore, these indicators disappear the forecast values for ARIMA model.

3. Optimization

Optimization is minimizing or maximizing some functions over the feasible set. The advantage of an optimization leads to develop in calculation and analyzing together with suitable modeling in software applications. The project demand or the project problem is designed into the function which calls objective function by a user. The objective function can be equations or inequations depending on user's requirement. Linear programming was a prototype in modern optimization for objective function and constraints of the problem. Therefore, the problem structure has more diversity in linear programming. Some problems are contrasted by the nonlinear programming and this becomes the beginning of several optimization techniques for various problem types [17].

Luenberger and Ye [16] mentioned that the problems are typically divided into 2 categories: linear programming (LP) and nonlinear programming (NLP) which are divided into subtypes of constrained problems and unconstrained problems. LP is a technique of mathematical programming problem which is the simplest and most widely applied to reach the best result for minimizing or maximizing. The objective problem is to optimize a linear function of variables with subject to one or more linear constraints. On the other hand, NLP technique is a mathematical programming in which the objective function and constraints are nonlinear and it is more complex than linear functions. Moreover, there are many methods that can be used to solve an NLP problem such as penalty and barrier, gradient projection, branch and bound, modified newton, nonlinear simplex or quadratic, and so on. **In Paper II,** we applied the NLP technique to find optimal values for the important indicators that maximize the overall score of an institution by using an optimization technique on the normalization data process as the first step of most ranking systems.

Since the development in computer and technology has been driven over time, thus the several softwares were created for solving the NLP problems. In fact, there is no fixed method to solve the NLP. However the suitable method will be chosen by the software and the Maple l8 was applied in the analysis throughout **Paper II.**

In Paper II, the study started from setting the objective function based on four indicators of the QS ranking which are considered to be controllable. They are faculty students ratio, citations per faculty, proportion of international faculty and proportion of international students. The constraints were constructed in the three cases which considered boundaries of each variable following context, ability and a policy of PSU. The first constraint, assigned all inequality constraints. Secondly, defined constraints to both equality and inequality. The last one, designed to be the boundary of standardization. The optimization results provided diversity of optimal values of each variable which depended on flexibility of boundaries setting.

4. Conclusions

The two subprojects were studied to find out the way for planning organization management to improve the performances of PSU that relate with ranking criteria to exist in higher rank among world university ranking. **In Paper I,** we studied the tendency of performances of PSU based on THE and QS indicators by using the five classical forecasting techniques and compared the most appropriate method. The results showed that most variables increased over time. The Holt's method provided a minimum value of MSE for the trend series and the SES for no-trend series. Since both of them are smoothing technique. Therefore, we can imply that the suitable technique to predict the data in this study was the exponential smoothing. Whereas considering the MAPE to identify the appropriate forecasting technique, we conclude that the smoothing technique is quite proper technique and most indicators increase in 2015, except the citations to staff for trend series and for no-trend are doctoral degree awarded to staff and doctoral degree awarded to staff. **In Paper II**, optimization was carried out to find the optimal values of variables that are in common of indicators for world university rankings, over the problem function which was created relying on calculation of world ranking score. The results showed that when the boundaries of constraints were changed, the optimal values of each variable were also changed. Thus, we can only achieve the approximate optimal values from this study to accomplish the better ranking.

5. Suggestions

5.1 For trend testing, the unit root test; the tool for test a stationary, is recommended to use instead of using the run chart.

5.2 For an effective of prediction, each indicator should be forecasted base on the same size of time series.

5.3 The forecast values which apply to the next, should not differ much from the previous.

5.4 Comparing forecasting models should compare by using the MAPE for easier understanding.

5.5 Since the factors which effect to the higher rank of PSU are related to people, researches, budget as well as policies and contexts of the university therefore it will be better if we fix the values of them into the constrain in Paper II too.

6. References

[1] Anhøj, J., & Olesen, A. V. (2014). Run Charts Revisited: A Simulation Study of Run Chart Rules for Detection of Non- Random Variation in Health Care Processes. *PLoS One, 9(11),* 1-13.

[2] Baldigara, T. (2013). Forecasting Tourism Demand in Croatia: A Comparison of Different Extrapolative Methods. *Journal of Business Administration Research, 2(1),* 84-92.

[3] Baty, P. (2014). The Times Higher Education World University Rankings, 2004-2012. *ETHICS IN SCIENCE AND ENVIRONMENTAL POLITICS, 13(2),* 125-130.

[4] Becerra, V. M. (2004). Solving Optimal Control Problems With State Constraints Using Nonlinear Programming and Simulation Tools. *IEEE TRANSACTIONS ON EDUCATION, 47(3),* 377-384.

[5] Bougnol, M. L. & Dulá, J. H. (2013). A mathematical model to optimize decisions to impact multi-attribute rankings. *Scientometrics, 95(2),* 785-796.

[6] Çakır, M. P., Acartürk, C., Alaşehir, O., & Çilingir, C. (2015). A comparative analysis of global and national university ranking systems. *Scientometrics, 103(3),* 813-848.

[7] Carey, J. J., & Delaney, M. F. (2010). T-Scores and Z-Scores. *Clinical Reviews in Bone and Mineral Metabolism , 8(3),* 113-121.

[8] Buela-Casal, G., Gutiérrez-Martínez, O., Bermúdez-Sánchez, M. P., & Vadillo-Muñoz, O. (2007). Comparative study of international academic rankings of universities. *Scientometrics, 71(3),* 349-365.

[9] Chai, T., & Draxler, R. R. (2014). Root mean square error (RMSE) or mean absolute error (MAE)?-Arguments against avoiding RMSE in the literature. *Geoscientific Model Development, 7(3),* 1247-1250.

[10] Chatfield, C. (l989). *THE ANALYSIS OF TIME SERIES AN: INTRODUCTION.* NewYork, NT: Chapman and Hall.

[11] Gill, P. E., Murray, W., & Wright, M. H. (1984). *Practical Optimization.* New York: Academi Press.

[12] Hobai, R. (2009). ANALYSIS OF AIR TEMPERATURE TENDENCY IN THE UPPER BASIN OF BARLAD RIVER. *Carpathian Journal of Earth and Environmental Sciences, 4(2),* 75-88.

[13] Huang, M. H. (2011). A Comparison of Three Major Academic Rankings for World Universities: From a Research Evaluation Perspective. *Journal of Library and Information Studies, 9(1),* 1-25.

[14] Khor, K. A., & Yu, L. G. (2016). Influence of international co-authorship on the research citation impact of young universities. *Scientometrics, 107(3),* 1095-1110.

[15] Khosrowjerdi, M., & Kashani, Z. S. (2013). Asian top universities in six world university ranking systems. *Webology, 10(2),* 1-9.

[16] Luenberger, D. G., & Ye, Y. (1984). *Linear and nonlinear programming. (2nd ed.).* Reading, MA: Addison-Wesley.

[17] Malekzadeh, P., & Gore, S. D. (2012). A stochastic Approach to General Nonlinear Programming Problems. *Communications in Statistics- Simulation and Computation, 41(10),* 1989-1997.

[18] Marginson, S. (2014). University Rankings and Social Science. *European Journal of Education, 49(1),* 45-59.

[19] Montgomery, D. C., Jennings, C. L., & Kulahci, M. (2015). *Introduction to time series analysis and forecasting. (2nd ed.).* John Wiley & Sons.

[20] Morse, R., Krivian, A., & Jackwin, A. (2016). How U.S. News Calculated the Best Global Universities Ranking. http://www.usnews.com/education/best-global-universities!articles/methodology. Accessed 24 October 2016.

[21] Moskovkin, V. M., Golikov, N. A., Peresypkin, A. P., & Serkina, O. V. (2015). Aggregate ranking of the world's leading universities. *Webology, 12(1),* 1-10.

[22] Nazim, A., & Afthanorhan, A. (2014). A comparison between single exponential smoothing (SES), double exponential smoothing (DES), holts (brown) and adaptive response rate exponential smoothing (ARRES) techniques in forecasting Malaysia population. *Global Journal of Mathematical Analysis, 2(4),* 276-280.

[23] Ostertagova, E., & Ostertag, O. (2012). FORECASTING USING SIMPLE EXPONENTIAL SMOOTHING METHOD. *Acta Electrotechnica et Informatica, 12(3),* 62-66.

[24] Perla, R. J., Provost, L. P., & Murray, S. K. (2011). The run chart: a simple analytical tool for learning from variation in healthcare processes. *BMJ Quality Safety, 20(1).* 46-51.

[25] Rayanakorn, M. (2011). Global Ranking of Higher Education Institutions. *Chiangmai J. Sci., 38(2),* 1-4.

[26] Riddle, W. R., & DonLevy, S. C. (2010). Generating expected growth curves and Z-scores for premature infants. *Joumal of Perinatology, 30(11),* 741-750.

[27] Robinson-García, N., Torres-Salinas, D., López-Cózar, E. D., & Herrera, F. (2014). An insight into the importance of national university rankings in an international context: The case of the I-UGR Rankings of Spanish universit*ies. Scientometrics, 101(2),* 1309-1324.

[28] Rockafellar, R. T. (1999). Extended Nonlinear Programming. *Nonlinear Optimization and Applications 2,* 381-399.

[29] Sahu, P. K., & Kumar, R. (2013). Demand Forecasting For Sales of Milk Product (Paneer) In Chhattisgarh. *International Journal of Inventive Engineering and Sciences (IJIES), 1(9),* 10-13.

[30] Sanguanrungsirikul, D., Chiewanantavanich, H., & Sangkasem, M. (2015). A comparative Study to Determine Optimal Models for Forecasting the Number of Patients Having Epidemiological-Surveillance Diseases in Bangkok. *KMUTT Research And Development Journal, 38(1),* 35-55.

[31] Singh, A., P., Gaur, M., K., KumarKasdekar, D., & Agrawal, S. (2015). A Study of Time Series Model for Forecasting of Boot in Shoe Industry. *International Journal of Hybrid Information Technology, 8(8),* 143-152.

[32] Shehatta, I., & Mahmood, K. (2016). Correlation among top 100 universities in the major six global rankings: policy implications. *Scientometrics, 109(2),* 1231-1254.

[33] Smith, P. (2012). Ranking and the Globalization of Higher Education. *Silpakorn University Journal of Social Sciences, Humanities, and Arts. 12(2),* 35-70.

[34] Taesombut, S. (2006). *Quantitative Forecasting. (1st ed.)* Thailand, TH: Kasetsart University Printing.

[35] Tijssen, R. J., Yegros-Yegros, A. & Winnink, J. J. (2016). University-industry R&D linkage metrics: validity and applicability in world university rankings. *Scientometrics, 109(2),* 677-696.

[36] Usher, A., & Savino, M. (2007). A Global Survey of University Ranking and League Tables. *Higher Education in Europe, 32(1),* 5-15.

[37] Yaffee, R. A., & McGee, M. (2000). *An introduction to time series analysis and forecasting: with applications of SAS@and SPSS@.* Academic Press.

Paper I

A Comparison of Forecasting Techniques used to Predict World University Rankings Tendency of Prince of Songkla University

Jubjaimoh, P., Samart, K., Jansakul, N., and Panityakul, T.

Proceeding

A Comparison of Forecasting Techniques used to Predict World University Rankings Tendency of Prince of Songkla University

Pattira Jubjaimoh[1], Klairung Samart[2*], Naratip Jansakul[3]and Thammarat Panityakul[4]
[1]*Department of Mathematics and Statistics, Prince of Songkla University, Songkhla, 90110,Thailand, 5710220072@email.psu.ac.th*
[2]*Department of Mathematics and Statistics, Prince of Songkla University, Songkhla, 90110,Thailand, klairung.s@psu.ac.th*
[3]*Department of Mathematics and Statistics, Prince of Songkla University, Songkhla, 90110,Thailand, naratip.j@psu.ac.th*
[4] *Department of Mathematics and Statistics, Prince of Songkla University, Songkhla, 90110,Thailand, thammarat.p@psu.ac.th*
(* corresponding author)

Abstract

World university rankings reflect quality and potential of educational institutes. Therefore, most institutes need to improve their performance for being higher rank on league tables. The objective of this study is to forecast a possible tendency of indicator scores of the Times Higher Education (THE) and the Quacquarelli Symonds Rankings (QS) of Prince of Songkla University (PSU), Thailand. The exponential smoothing, the moving average and the ARIMA techniques have been compared to find out which technique is more appropriate to predict the performance of PSU based on the indicators of the THE and QS. The data such as the number of academic staff, the number of full-time students, etc., from 1994 to 2014 have been used in the analysis. The data were firstly classified into two classes relying on trend and no existing trend. For no-trend series, the single moving average (SMA) and the single exponential smoothing (SES) were chosen to predict the data tendency. For the other class, the double moving average (DMA) and Holt's method were applied. In addition, ARIMA was also used to forecast for both groups. According to the mean squared error (MSE), the SES is the most appropriate technique for the no-trend series, whereas the Holt's is suitable for the trend series.

Keywords: time series, ARIMA, Quacquarelli Symonds Rankings, Times Higher Education, world university ranking

1. Introduction

1.1 World University Ranking

Nowadays the position on the university ranking league tables is emphasized by educational institutions as it reflects quality and potential of each educational institution. The university rankings have been built for fifteen years where the first ranking was announced in 1983 by the US News and World Report publishing for the American colleges annual rankings [7, 9]. The global ranking system was established in 2003 by Shanghai Jiao Tong University, China. The system was operated focusing on the performance of the university researches. Such systems became the prototype of several current systems [1]. Appearance of Shanghai ranking system affected the development of an educational market to grow up rapidly [4]. The benchmarking tools are developed and the educational institute competition is expanded as well as racing to make a higher academic quality together with increasing the amount of academic research and so on. Due to the worldwide spread of the educational competition, most universities need to increase rank by magnifying researches, building a reputation and recruiting outstanding students for leading them to be on top or be an acceptable position of world ranking table [8]. Therefore, these lead to the cooperation of scholars, colleges, universities, academic institutes and government, to be in the higher place of educational institution on league tables to compare with the others and can be considered that the university is the educational leader and acceptable in world class [3]. Today, the global ranking systems rely on internationally accessible bibliometric databases and on reputation surveys to develop league tables at a global level. They analyzed in terms of the types of criteria which are similar among the performance indicators employed [2]. Cakir et al. [2], mentioned to the eight examples of famous world university ranking systems namely Academic Ranking of World Universities (ARWU), Higher Education Evaluation and Accreditation Council of Taiwan (HEEACT), LEIDEN, SCImago, Quacquarelli Symonds (QS), Times Higher Education (THE), University Ranking by Academic Performance (URAP) and WEBOMETRICS. Their common methodologies are related with research and citations [5, 9] whereas THE and QS also consider teaching performance and international outlook as additional indicators. Khosrow-jerdi and Kashani [6] studied on similarity and justness of the 200 Asian universities ranked by world ranking systems which operate on similar indicators direction. The similarity is monitored in indicators method and the justness is distinguished by the correlations of each university's position on each system's league table. At the end, the finding showed that there were three pairs that are highly correlated, namely, QS-WEBOMETRICS, ARWU-HEEACT and QS-THE. These indicated that the ranking systems are parallel. As QS and THE are the most popular ranking systems, this research will focus on predicting world university rankings of Prince of Songkla University based on THE and QS.

1.2 Times Higher Education (THE)

The recent THE ranking uses 13 performance indicators which are branched from grouping for the five core areas: teaching, research, citations, international outlook and industry income [2]. The different weighting for each indicator is shown in Table 1.

Table 1: Indicators of Times Higher Education

Indicators	Weight
Teaching	30%
- Teaching reputation	15%
- Doctorates awarded to academic staff ratio	6%
- Students-to-staff ratio	4.5%
- Doctorate-to-bachelor's ratio	2.25%
- Institutional income / academic staff	2.25%
Research	30%
- Research reputation	18%
- Research income/ academic staff	6%
- Papers / academic and research staff	6%
Citations	30%
- Normalized Citation Impact	
International outlook	7.5%
- International-to-domestic staff ratio	2.5%
- International-to-domestic students ratio	2.5%
- International co-author / papers	2.5%
Industry income	2.5%
- Research income from industry / academic staff	

Marginson [8] had mentioned that the comprehensiveness of THE is its main strength. Comprehensiveness in Marginson's sense means to the thirteen separated indicators from five areas which are weighted, scaled as well as crushed till the ending unitary score. Teaching, research and citations are the three areas that hold the big percentage as 30% for each and followed by international outlook and industry income which hold the smallest proportion from the five. For sub-indicators, reputational surveys as teaching and research hold the two biggest ratios that constitute 15% and 18%, respectively.

1.3 Quacquarelli Symonds Rankings (QS)

QS was first published in 2004. Since 2010, it has been using the Scopus database for research methodology [11]. Over 700 from the total of 2,000 universities worldwide are ranked. The first 400 were ranked individually, whereas the remaining were ranked by grouping [18]. Moreover, the purpose of QS focuses on comprehensive performances beside research. It consists of six indicators for world ranking methodologies as shown in Table 2.

Table 2: Indicators of Quacquarelli Symonds Rankings

Indicators	Weight
Academic reputation	40%
Staff-to students ratio	20%
Citations per faculty	20%
Employer reputation	10%
Proportion of international students	5%
Proportion of international staff	5%

QS's methodology works mostly on reputation of institutes. Its indicators were originated with five areas; reputation, staff-to-students ratio, citations per faculty, proportion of international students and proportion of international staff. For reputation, it comes from two parts. The first one is academic reputation which counts 40%. The second one is employer reputation which is based on graduate employers' survey worldwide weighted for 10%. For staff-to-students ratio, it holds on potential of teaching by monitoring the proportion of staff and student which is weighted for 20%. For citations per faculty, it indicates the research potential with the proportion between citations and academic faculty (full time) which is weighted for 20%. Lastly, it is an international image with the two indicators as proportion of international academic staff which is weighted for 5% and proportion of international students which is weighted for 5%.

Since Prince of Songkla University (PSU) also requires improving its rank to be in the world class, therefore this research is conducted. This research will focus on looking at the indicator score trends for PSU, based on the methodologies of QS and THE world ranking by using forecasting techniques. The exponential smoothing (SMA and DMA), the moving average (SES and HOLT's method) and the ARIMA techniques have been compared in the form of predicting the performance of PSU on the basis of the indicators of the THE and QS. The data such as the number of academic staff, the number of full-time students, the number of citations, etc., from 1994 to 2014 have been used in the analysis.

2. Research Methodology

2.1 Data

Data used in the analysis are the number of academic staff (total, domestic and international), the number of full-time students (total, domestic and international), the number of doctorate degree awarded, the number of bachelor degree awarded, the amount of institute income, the amount of research income, the amount of research income from industry, the number of citations, the number of papers and the number of co-foreigner author papers. The data are yearly collected in 1994 to 2014 and there is no seasonal pattern in the series. In the part of analysis, there are different twelve indicators that we use for study in this research based on THE and QS indicators. Those are doctoral degree awarded to academic staff, doctoral degree awarded to under degree awarded, academic staff to students, research income to academic staff, institute income to academic staff, research industry income to academic staff, normalize citation impact, international students to students, international academic staff to academic staff, papers to academic staff, paper with co-author to papers and citation to staff.

2.2 Forecasting techniques

Yaffe and McGee [17] mentioned that averaging techniques and exponential smoothing are used to make a forecast value. Since the Minitab 16 is used to analyse data in this research and a seasonal component does not appear in the data series, so the appropriate techniques include the two of averaging techniques as the simple moving average and the double moving average. Moreover, there are also two exponential techniques as the simple exponential smoothing and the Holt's linear exponential smoothing. Furthermore, Taesombut [16] also stated that these techniques can be classified as trend and no trend series by using "run test".

The software Minitab 16 is applied throughout this research. Firstly, we start from separating trend and no trend series by using the option "run chart" in Minitab 16 and determine which indicator has trend or no trend for further analysis.

2.2.1 Run chart

Perla, et. al. [19] described about run chart that it is a graphical display of data plotted in order. It works by calculating the median and applied as the centerline. Therefore, the median is not affected by severe values in the data. Run chart helps to conceive the impact of varied interventions and tests of over time hanging. Thus, we can determine to the trend series if $p<0.05$ and no-trend if $p > 0.05$ [16].

2.2.2 Forecasting Techniques for no-trend series

2.2.2.1 Single Moving Average method

Sahu and Kumar [13] mentioned about the moving average method or SMA that this technique involves calculating the average of observations and then using that average like the predictor for the next period. The moving average technique is highly dependent on the number of periods which are selected for requiring to average. The SMA model is shown as follows:

$$F_{t+1} = \frac{(Y_t + Y_{t-1} + Y_{t-2} + \ldots + Y_{t-n+1})}{n} \qquad (1)$$

where :

F_{t+1} = the forecast value at time t+1

Y_t = the actual value at the period t

n = the number of items in the moving average

2.2.2.2 Single Exponential Smoothing method:

The single exponential smoothing technique (SES) uses a weighted moving average of historical data to be a forecasting basic process [13]. It is the simplest technique of exponential smoothing techniques which is suitable for no trend series and no seasonal patterns [14]. The condition of single exponential smoothing technique depends on the optimal value. The perfect optimal value should be between 0 and 1 and prefers to be range from 0.1 to 0.3 as mentioned in Sahu and Kumar [13]. The equation for the simple exponential smoothing method is [29]:

$$F_{t+1} = \alpha Y_t + (1-\alpha) F_t \qquad (2)$$

where:

F_{t+1} = the forecast value for time period t+1

α = the smoothing constant (should be in range 0 to 1)

Y_t = the actual value for time period t

F_t = the forecast value of variable Y for time t

2.2.3 Forecasting Techniques for trend series

2.2.3.1 Double Moving Average method

According to Baldigara [12], the double moving average method or DMA is used for a variation data as linear trend series. Forecasting by a double moving average needs to assign two averages, computing and after that calculating the second moving average [13]. At the time that both simple and double moving averages are available, a slope and a trend coefficient are calculated by these averages. Then the average has changed over p periods to forecast another one or more periods into the future. All equations for DMA and the model for forecasting h periods ahead are:

$$M_t = F_{t+1} = \frac{(Y_t + Y_{t-1} + Y_{t-2} + \ldots + Y_{t-n+1})}{n} \qquad (3)$$

$$M_t^{'} = \frac{(M_t + M_{t-1} + M_{t-2} + \ldots + M_{t-n+1})}{n} \qquad (4)$$

$$S_t = 2M_t - M_t^{'} \qquad (5)$$

$$T_t = (\frac{2}{n-1})(M_t - M_t^{'}) \qquad (6)$$

$$F_{t+h} = S_t + T_t h \qquad (7)$$

where :

Y_t = the actual series value at the time period t

S_t = intercept or estimate level value at time t

T_t = slope coefficient or estimate of trend at time t

t = some of time period

h = the number of time periods ahead to forecast

n = the number of period of double moving average

M_t = the single moving average

$M_t^{'}$ = the double moving average

2.2.3.2 Linear Exponential Smoothing method

Nazim and Afthanorhan [14] stated that the linear exponential smoothing technique (LES) *or HOLT's method* is developed from the double exponential smoothing (DES) which it is not only using different smoothing constant to smooth the trend and the slope directly but it also gives more adaptability in selecting the rates at which trend and slopes are followed. The model of the Holts method requires three equations as follows:

$$T_t = \beta(S_t - S_{t-1}) + (1-\beta)T_{t-1} \quad (9)$$

$$S_t = \alpha y_t + (1-\alpha)(S_{t-1} + T_{t-1}) \quad (10)$$

$$F_{t+h} = S_t + T_t(h) \quad (11)$$

where: $0 < \alpha < 1$, $0 < \beta < 1$ and

S_t = the estimate level value at the time t

T_t = the trend estimates at the time t

F_{t+h} = the forecast for h-step-ahead period

2.2.4 The Box-Jenkins (ARIMA)

The Box-Jenkins (ARIMA) is a method to defer the series to stationary and combine the moving average with the autoregressive parameters to provide a forecasting model. The highlight of Box-Jenkins is a greater flexibility and power modeling. The model does not just explain the process of the series generating but it is also a basis for predicting. There are a few limitations of Box-Jenkins model. It can work better than the other forecasting techniques when it has enough data at least 50 observations [17]. Furthermore, appearing of missing values among series is one of limitations because these may bring outliers to the series. Therefore, the assumption of Box-Jenkins includes the observations lengths and the complete time series. Moreover, other assumptions are a stationary in mean, variance and autocovariance of series which should be constant for time lags. If the series is non-stationary, it needs to be transformed to be a stationary series first [17]. These are the basic processes of the Box-Jenkins ARIMA (p, d, q) model. The AR (p) and MA (q) values receive from considering correlograms of PACF and ACF respectively. In case of the value of d comes from integrating the model, if a series is I(0), it can assume that it is a stationary series and ARIMA (p,0,q) model is produced. A combination of this process is sometimes called ARMA (p, q) model. Another case, if series need a first differencing to be stationary which gives the value d=1 then it produces ARIMA (p,1,q) model. Hence, the ARIMA must consist of AR (p) and MA (q) components which is important to determine the pattern of ARIMA model. Yaffe and McGee [17] presented the actual difference equation of ARIMA model as:

$$\begin{aligned} Y_{t+h} &= \varphi_1 Y_{t+h-1} + \varphi_2 Y_{t+h-2} + \ldots + \varphi_{p+q} Y_{t+h-p-d} \\ &- \theta_1 \varepsilon_{t+h-1} - \theta_2 \varepsilon_{t+h-2} - \ldots - \theta_q \varepsilon_{t+h-q} - \varepsilon_{t+q} \end{aligned} \quad (8)$$

where:

Y_{t+h} = the forecast value at time t+h

ε_{t+q} = forecast error value at time t+q

θ = the coefficient from moving average process

φ = the coefficient from autoregressive process

2.2.4.1 Autoregressive process:

Yaffe and McGee [17] presented the autoregressive model for AR(1) as

$$Y_t = \varphi_1 Y_{t-1} + \varepsilon_t \quad (12)$$

where:

Y_t = the function of some portion of Y_{t-1} plus an error term ε_t

φ_1 = the portion of the previous rating carried over the rating at time t

2.2.4.2 Moving average process:

Yaffe and McGee [17] also presented the moving average model for MA(1) as

$$Y_t = \varepsilon_t - \theta_1 \varepsilon_{t-1} \quad (13)$$

where;

Y_t = the mean centered series (or Y_t = y_t - μ where; y_t is the original series and μ is the mean of series)

ε_t = the error term at time t

ε_{t-1} = the previous error term

θ_1 = the moving average coefficient

2.3 Accuracy measuring

Ostertagova and Ostertag [15] mentioned that each model's accuracy should be evaluated to examine the errors or bias of the information as well as to compare for the fitted forecast models. Nowadays, no one can tell which measurement is the best one to determine the most appropriate forecasting method. However, the most significant evaluation of the quality of forecasting models is a goodness of fit indicator which considers from a minimum value. Examples of common indicators to measure accuracy include mean absolute error (MAE), mean squared error (MSE), root mean square error (RMSE) and mean absolute percentage error (MAPE). In this research we will compare forecasting models by using the MSE. The equation is given by

$$MSE = \frac{1}{n}\sum_{t=1}^{n}(e_t)^2 = \frac{1}{n}\sum_{t=1}^{n}(Y_t - F_t)^2 \quad (15)$$

where:

$e_t = Y_t - F_t$ =the residual or the actual value minus the forecast value

n =the total number of the time periods

3. Research Result and Discussion

3.1 The result of run chart

Table 3 shows the results of run chart to consider trend existing. For doctoral degree awarded to academic staff, doctoral degree awarded to undergraduate degree awarded, academic staff to students, research income to academic staff, research industry income to academic staff, normalize citation impact and international students to students, the p-values are greater than the significance level of 0.05. This implies that there is no indication of special cause of trend variation. On the other hand, institute income to academic staff, international academic staff to academic staff, papers to academic staff, paper with international co-author to papers and citation to staff, are indicators that the p-values are less than 0.05. This implies that trend variation exists in series.

3.2 The result of forecasting

3.2.1 The result of forecasting for trend series

After using the run chart in Minitab 16 to classify the trend and fitting data with forecasting models, the results of forecast and MSE for trend series are shown in Table 4.

Table 3: Result of run chart for indicators

Indicators	**Number of runs up or down**	**Expected number of runs**	**p-value**
1. Doctoral degree awarded / staff	12	13.7	0.183
2. Doctoral awarded /undergraduate awarded	10	10.3	0.417
3. Academic staff / students (total)	11	13.7	0.074
4. Research income / staff	7	9	0.087
5. Institute income / staff	1	5	0
6. Research industry income / staff	7	9	0.087
7. Normalized citation impact	14	13.7	0.572
8. Students (inter) / students (total)	13	12.3	0.649
9. Academic staff (inter) /staff (total)	6	11.7	0
10. Papers (total) / academic staff	8	13.7	0.001
11. Paper with co-author / papers (total)	10	13.7	0.024
12. Citation / staff	8	12.3	0.007

Table 4: A comparison of forecasting values and MSE of the trend series by using DMA, HOLT's and ARIMA methods

Indicators	Actual value	Methods	Forecast values			MSE
	2014		2015	2016	2017	
1. Academic staff (inter) /staff (total)	0.033521985	DMA	0.0376777	0.0376777	0.0376777	0
		HOLT	0.0335788	0.032606	0.0316333	0
		ARIMA	0.334834	0.0334449	0.0334065	0.000031709
2. Papers (total) / academic staff	0.27601219	DMA	0.260585	0.260585	0.260585	0.0009
		HOLT	0.295615	0.310122	0.324629	0.0004
		ARIMA	0.278636	0.281285	0.283958	0.00049076
3. Paper with co-author / papers (total)	0.356466877	DMA	0.370493	0.370493	0.370493	0.00206
		HOLT	0.317698	0.296698	0.275697	0.00363
		ARIMA	0.376480	0.393863	0.408962	0.0038084
4. Citation / staff	0.592076622	DMA	1.88825	1.88825	1.88825	0.4485
		HOLT	-0.71119	-1.6981	-2.68512	0.3113
		ARIMA	0.57402	0.55650	0.53953	0.39819
5. Institute income / staff	462412.9299	DMA	441303	441303	441303	1331700778
		HOLT	474306	489525	504744	103002016
		ARIMA				

30

Table 5: A comparison of forecasting values and MSE of the no-trend series by using SMA, SES and ARIMA methods

Indicators	Actual value 2014	Methods	Forecast values 2015	2016	2017	MSE
1. Doctoral degree awarded / staff	0.002612103	SMA	0.0028621	0.0028621	0.0028621	0
		SES	0.002866	0.002866	0.002866	0.0001
		ARIMA	0.0024165	0.0022356	0.0020682	0.00009236
2. Doctoral awarded /undergraduate awarded	0.000742666	SMA	0.0007888	0.0007888	0.0007888	0
		SES	0.0007887	0.0007887	0.0007887	0
		ARIMA	0.0006801	0.0006228	0.0005704	0.00036374
3. Academic staff / students (total)	5.575968655	SMA	5.52814	5.52814	5.52814	0.816
		SES	5.58779	5.58779	5.58779	0.447
		ARIMA	5.60327	5.63071	5.65828	0.46753
4. Research income / staff	109101.4253	SMA	117680	117680	117680	444236844
		SES	110303	110303	110303	26955219
		ARIMA	109306	109512	109718	285691554
5. Normalized citation impact	0.67	SMA	1.19667	1.19667	1.19667	0.1389
		SES	1.09201	1.09201	1.09201	0.1213
		ARIMA				
6. Research income industry / staff	2597.009578	SMA	6611.67	6611.67	6611.67	11054103
		SES	4948.01	4948.01	4948.01	8446234
		ARIMA	1940.37	1449.75	1083.19	10619015
7. Students (inter) / students (total)	0.021627108	SMA	0.0225822	0.0225822	0.0225822	0
		SES	0.0227512	0.0227512	0.0227512	0
		ARIMA	0.0216524	0.0216778	0.0217032	0.000025528

From Table 4, we see that the HOLT's method gives the smallest MSE in four indicators; the papers to academic staff, paper with co-author to papers, citation to staff and institute income to academic staff. When consider the results based on using HOLT's method, papers to academic staff and institute income to academic staff appear to increase from 2014 but the paper with co-author to papers and citation to staff decrease in 2015. For the international academic staff to academic staff, the MSE of DMA and HOLT's methods are equal but DMA method gives a forecast value higher than actual value in 2014.

3.2.1 The result of forecasting for no-trend series

From Table 5, for academic staff to students, research income to academic staff, research industry income to academic staff and normalize citation impact, SES method provides minimum MSEs when compared with the other two methods and the forecast values are higher than the actual values in 2014. Whereas, the results of doctoral degree awarded to under degree awarded and international students to students show an equal and smaller MSE value when compared with the result of ARIMA model. The tendency of them is increasing as well. Moreover, the doctoral degree awarded to academic staff is one of indicators of this group in which ARIMA (1, 0, 0) model gives the smallest MSE but a forecast value is less than a value in 2014.

From Table 4 and 5, there are no results of institute income to academic staff and normalized citation impact by ARIMA model. According to Yaffee, et. al. [17], ARIMA model can work better for data that are available for more than 50 observations. However, in our data there are only 7 observations available for institute income to academic staff.

4. Conclusion

In this study, the tendency of most performances of PSU based on THE and QS indicators is increasing over time and there are some indicators that are decreasing but not in a big deal. Since this study aims to compare five methods of forecasting, the data were classified into two categories; trend series and no-trend series. Methods used for a group of trend series include SMA, SES and ARIMA. The output shows that SES method gives a minimum value of MSE for most indicators in this type. Whereas, another group which is no-trend series show results that HOLT's method provides a minimum value of MSE. Both SES and HOLT's are methods of smoothing technique. Thus, the exponential smoothing is the suitable technique to predict the data in this study. Furthermore, most of indicators appear an increasing value in 2015, except the paper with co-author to papers and citation to staff indicators.

5. Acknowledgement

We would like to thank Prince of Songkla University for providing some data and information for this study. We would also like to thank Department of Mathematics and Statistics for all facilities used in this research.

6. References

[1] Baty, P. (2014). *The Times Higher Education World University Ranking, 2004-2012*. Ethics Sci Environ Polit, 13, 125-130.

[2] Cakir, P., M., Acarturk, C., Alasehir, O. and Cilingir, C. (2015). *A comparative analysis of global*

and national university ranking systems. Scientometrics, 103, 813-848.

[3] Casal, G., B., Martinez, O., G., Sanchez, M., P., B., and Munoz, O., V. (2006). *Comparative study of international academic rankings of universities*. Scientometrics, 1-17.

[4] Garcia, N., R., Salinas, D., T., Cozar, E., M., L., Herrera, F. (2014). *An insight into the importance of national university rankings in an international context: The case of the I-UGR Rankings of Spanish universities*. Scientometrics, 101(2), 1309-1324.

[5] Hung, M., H. (2011). *A Comparison of Three Major Academic Rankings for World Universities: From a Research Evaluation Perspective*. Journal of Library and Information Studies, 9(1), 1-25.

[6] Khosrowjerdi, M., and Kashani, Z.S. (2013). *Asian top universities in six world university ranking systems*. Webology, 10(2), 1-9.

[7] Smith, P. (2012*). Ranking and the Globalization of Higher Education*. Journal of Social, 12(2), 35-69.

[8] Marginson, S. (2014). *University Rankings and Social Science.* European Journal of Education, 49(1), 45-59.

[9] Moskovkin, V., M., Golikov, N., A., Peresypkin, A., P., and Serkina, O., V. (2015). *Aggregate ranking of the world's leading universities*. Webology, 12(1), 1-10.

[10] Usher, A. and Savino, M. (2007). *A Global Survey of University Ranking and League Tables*. Higher Education in Europe, 32(1), 5-15.

[11] Rayanakorn, M. (2011). *Global Ranking of Higher Education Institutes*. Chimai J. Sci., 38(2), 1-4.

[12] Baldigara (2013). *Forecasting Tourism Demand in Croatia: A Comparison of Different Extrapolative Methods*. Sciedu Press, 2(1), 84-92.

[13] Sahu, P, K., and Kumar, R. (2013). *Demand Forecasting For Sales of Milk Product (Paneer) In Chhattisgarh*. International Journal of Inventive Engineering and Sciences (IJIES), 1(9), 10-13.

[14] Nazim, A., and Afthanorhan, A. (2014). *A comparison between single exponential smoothing (SES), double exponential smoothing (DES), holts (brown) and adaptive response rate exponential smoothing (ARRES) techniques in forecasting Malaysia population*. Global Journal of Mathematical Analysis, 2 (4), 276-280.

[15] Ostertagova, E., and Ostertag, O. (2012). *FORECASTING USING SIMPLE EXPONENTIAL SMOOTHING METHOD*. Acta Electrotechnica et Informatica, 12(3), 62-66.

[16] Taesombut, S. (2006). *Quantitative Forecasting*. Thailand, TH: Kasetsart University Printing.

[17] Yaffee, R. A., & McGee, M. (2000). *An introduction to time series analysis and forecasting: with applications of SAS® and SPSS®*. Academic Press.

[18] Marginson, S. (2014). *University Rankings and Social Science*. European Journal of Education, 49(1), 45-59.

[19] Sanguanrungsirikul, D., Chiewanantavanich, H., Sangkasem, M. (2015). *A comparative Study to Determine Optimal Models for Forecasting the Number of Patients Having Epidemiological-Surveillance Diseases in Bangkok*. KMUTT Research And Development Journal, 38(1), 35-55.

Paper II

Optimization for Better World University Ranking

Jubjaimoh, P., Samart, K., Jansakul, N., and Jibenja, N.

Submitted Manuscript

Optimization for Better World University Ranking

Pattira Jubjaimoh[1], Klairung Samart[1], Naratip Jansakul[1] and Nattada Jibenja[1]

Abstract World university rankings are very competitive among educational institutions. Many institutions realize the importance and require to step up to a better rank. The aim of this paper was to find the optimal values for each of our studied indicators: faculty students ratio, citations per faculty, proportion of international faculty, and proportion of international students, that maximizes the overall score of Quacquarelli Symonds World University Ranking. Those four indicators are commonly used in most university ranking systems and considered to be controllable. An approach of optimization using maximization of nonlinear programming problem in which the objective function was constructed from normalization and weighting was applied throughout this research. Three cases of constraints that are different in boundary determination were considered. The results from the analysis showed that the optimal values were varied depending on the constraints. The final decision for the optimal values is based on context, ability and policy of an individual educational institution.

Keywords World University Ranking · Indicator · Z-score · Nonlinear Programming

Acknowledgement

We would like to thank Prince of Songkla University for providing some data and information for this study. We would also like to thank Department of Mathematics and Statistics for all facilities used in this research.

Pattira Jubjaimoh
5710220072@email.psu.ac.th

✉ Klairung Samart
klairung.s@psu.ac.th

Naratip Jansakul
naratip.j@psu.ac.th

Nattada Jibenja
natthada.ye@psu.ac.th

[1] Department of Mathematics and Statistics, Faculty of Science, Prince of Songkla University, Songkhla, 90110,Thailand

34

Introduction

Nowadays, there are more than 17,000 higher academic institutions that are involved in the global university ranking systems. The most famous ranking systems are US News & World Report Best Global University Ranking (USNWR), Academic Ranking of World Universities (ARWU), University Ranking by Academic Performance (URAP), Quacquarelli Symonds World University Ranking (QS), Times Higher Education World University Ranking (THE), etc. Hence, the culture of competition among institutions has been increased remarkably. Also, the world university rankings obtain more interest from many groups of people such as parents, scholars, educational institutions, government, business sector, media and so on. Therefore, most universities or institutions are discussing about how to improve their rank for being acceptable as a world class institution. The discussion is mostly concentrated on how to improve indicator scores which indicate the institution's ability in different aspects. Therefore, potential or quality as well as a reputation of educational institutions are the keys to climb for a higher rank (Shehatta and Mahmood 2016; Tijssen et al. 2016).

For ranking methodology, there are different weights and criteria indicators for each ranking system. The USNWR focuses on research performance and reputation scores. ARWU uses the awards to measure academic research achievement. URAP focuses on scientific research efficiency only. QS and THE emphasize on the teaching, research as well as international performance. Different systems bring in different indicators. The common indicators are divided into teaching efficiency, research ability, basis of education, financial outcome, reputation, input staff and resources as well as collaboration activities such as research collaboration and international collaboration (Khor and Yu 2016; Shehatta and Mahmood 2016). However, the major criteria of famous academic ranking systems are teaching quality, research quality and internalization (Huang 2011; Marginson 2014; Cakir et al. 2015; Shehatta and Mahmood 2016). In order to climb in university rankings, Bougnol and Dulá (2013) recommended that concentrating on raising the score of few indicators is better than more. Therefore, this research focuses on the four indicators, namely faculty students ratio, citations per faculty, proportion of international faculty and proportion of international students. These indicators are the common ones and they are significant among most world university ranking systems. On the contrary, the reputation scores get values from the opinion surveys and those are getting out of control.

With the main research question of how a university or an institution can improve scores for better global university ranking, this study aims to find the optimal values for the important indicators that maximize the overall score of an institution by using an optimization technique on the normalization process since the first step of most ranking systems is to normalize (using z-scores) the data. However, the final process which produces an exactly final score of each system can be different. Therefore, this paper does not go deeply into this detail.

Z-scores

Z-score is derived from a normal distribution or Gaussian distribution. The z-score is generated by subtracting the population mean (μ) from the study value (x) and divided by the population standard deviation (σ). The feature of z-score is

$$\text{Z-score} = \frac{x - \mu}{\sigma}. \tag{1}$$

The Z-score indicates how many the population standard deviations a data point is from the population mean of that transformation. Given this, the study value is above the mean if z is positive and under the mean if z is negative (Carey and Delaney 2010; Riddle and DonLevy 2010).

For the global ranking scores, they are calculated by using the weight together with the z-scores for each indicator. Since a z-score is a standardized score, it is an equal comparison between different types of information. Hence, that transformation of the data is necessary when combining various data into a single ranking. However, being highly skewed of some indicators can be managed by using the logarithmic transformation to rescale the original values for a more normalized and uniform spread across other indicators. In normalization, a z-score of each indicator is calculated in order to standardize the different data types to be normal scale. Then, a cumulative probability function is obtained to evaluate where a particular institution's indicator locates within that function. To our knowledge, the overall score of QS and USNWR is computed by summing up the weighting value for each indicator of each institution and compare score of each institution with the highest score in the range of 0 and 100 to rank all institutions on league table ranking (Morse et al. 2016). As the z-score is the key function in calculating ranking scores, it was set to be the objective function for optimization problem in this research.

Optimization

Optimization processes are increasingly interesting since many organizations need an application to help for internal management decision making. The optimization as the process for solving complex problems consists of two main components namely, objective function and constraints. The objective function involves minimizing or maximizing problem and it is a set of decision variable values that are assumed to be optimal. The other components are the constraints or the limitations of the problem. It is a set of variables that are acceptable in this setting (Gill et al. 1981).

The optimization problems are divided into 2 types: linear programming (LP) and nonlinear programming (NLP) which consists of constrained problems and unconstrained problems (Luenberger and Ye 1984). LP is a technique of mathematical programming problem which is the simplest and most widely applied to reach the best result for minimizing or maximizing. The objective problem is to optimize a linear function of variables with subject to one or more linear constraints.

NLP technique is however a mathematical programming in which the objective function and constraints are nonlinear and solving the problem is more difficult than all linear functions. In general, there is no confirmation to the best NLP solution. Moreover, there are many methods that can be used to solve an NLP problem such as penalty and barrier, gradient projection, branch and bound, modified newton, nonlinear simplex or quadratic, and so on.

The maximization form of NLP problem is shown as follows (Malekzadeh and Gore 2012).

$$\begin{aligned} &\text{Maximize} && f(x) \\ &\text{subject to} && g_i(x) \le b_i, \qquad i = 1, 2, \ldots, m \end{aligned} \tag{3}$$

where x is the decision variables vector for n-dimensions,

$f(x)$ is the objective function to be maximized,

and $g_i(x)$ is the inequality constraint.

In this form, the functions are assumed to be continuously differentiable in $\Re^n$. For the minimization problem, it can be defined in a similar way (Malekzadeh and Gore 2012). In many cases of NLP problem, the set of optimal values are large and solving algorithm is complicated. Therefore, the analytic instruments like computer software, are developed and widespread. All results were analyzed by using the software Maple 18 throughout this paper and the software automatically chose an appropriate method to solve the optimization problem.

Methodology and Results Discussion

As mentioned above, an education market increasingly focuses on being a higher rank on the league tables. This is a starting point of this study with the problem of what optimal value of each indicator that maximizes the overall score. The research emphasized on four common indicators; faculty students ratio, citations per faculty, proportion of international faculty and proportion of international students which appear among world university ranking systems and are considered to be controllable. Based on some literature review, the process of overall score is originated by z-scores of each indicator. Then, multiply that score with the weight of each indicator and finally obtain the overall score. Therefore, the mean (μ) and the standard deviation or sd (σ) are also in our optimization process. However, these two values are not available for all ranking systems. Among those, we found that QS is the only system that provides this information. Hence, we used the QS's criteria in our analysis as a case study.

Table 1 Indicators, weights, means and standard deviations of QS World Rankings 2016

Indicators	**Weight**	**Mean (μ)**	**Standard deviation (σ)**
Academic reputation	40%	77.39	52.89
Faculty students ratio	20%	0.10	0.04
Citations per faculty	20%	37.55	29.70
Employer reputation	10%	18.20	11.11
Proportion of international faculty	5%	0.18	0.12
Proportion of international students	5%	0.16	0.10

The optimization problem form for this study is as follows.

$$\text{Maximize } 20\cdot P\left(\mathrm{X}<\left(\frac{\frac{x_1}{x_2}-0.10}{0.04}\right)\right)+20\cdot P\left(\mathrm{X}<\left(\frac{\frac{x_3}{x_1}-37.55}{29.70}\right)\right)+5\cdot P\left(\mathrm{X}<\left(\frac{\frac{x_4}{x_1}-0.18}{0.12}\right)\right)+5\cdot P\left(\mathrm{X}<\left(\frac{\frac{x_5}{x_2}-0.16}{0.10}\right)\right) \quad (4)$$

where X_1 is the number of faculty,

X_2 is the number of students,

X_3 is the number of citations,

X_4 is the number of the international faculty,

and X_5 is the number of international students.

Here, the objective function (4) is an NLP problem which is constructed from normalization and weighting for each indicator as shown in Table 1. In total, it consists of four terms to maximize the cumulative probability functions of faculty students ratio, citations per faculty, proportion of international faculty, and proportion of international students, respectively. The three cases of constraints considered are presented in Table 2. Case 1 defines all inequality constraints to set the maximal z-scores to be three and the boundary of each variable follows Prince of Songkla University context. Case 2 defines both equality and inequality constraints. The last one sets the boundary of standardization to be between -3 and 3.

38 **Table 2** The lists of constraints and results

Case	Constraints	X1	X2	X3	X4	X5
1	$\frac{x_1}{x_2} \le 0.22, \frac{x_3}{x_1} \le 126.65, \frac{x_4}{x_1} \le 0.54, \frac{x_5}{x_2} \le 0.46$ $x_1 \ge 2,000, x_2 \le 40,000, x_3 \ge 0, x_4 \ge 100, x_5 \ge 285$	3753.97	17063.5	475440	2027.14	7849.21
2	$\frac{x_1}{x_2} = 0.22, \frac{x_3}{x_1} = 126.65, \frac{x_4}{x_1} = 0.54, \frac{x_5}{x_2} = 0.46$ $x_1 \ge 2,000, x_2 \le 40,000, x_3 \ge 0, x_4 \ge 100, x_5 \ge 285$	2000	9090.91	253300	1080	4181.82
3	$-3 \le \frac{\frac{x_1}{x_2} - 0.10}{0.04} \le 3, -3 \le \frac{\frac{x_3}{x_1} - 37.55}{29.70} \le 3$ $-3 \le \frac{\frac{x_4}{x_1} - 0.18}{0.12} \le 3, -3 \le \frac{\frac{x_5}{x_2} - 0.16}{0.10} \le 3$	1372.6	6239.09	173839	741.2	2863.07

From Table 2, we found that different conditions give different outcomes for making decision. Each case contains boundaries which cannot precisely indicate. In fact, the boundaries are based on context, ability including a policy of an individual educational institution which can be different. Thus, the results depend on flexibility of boundaries setting. We cannot identify the best values for all decision variables. However, this research can give an idea of how each variable should be and this can lead to a plan making for improving the ranking of an individual educational institution.

Conclusions

Calculation of world university ranking score is complicated and can be different for each ranking system. Some are involved with scores of other institutions especially those who reach the top score and that information is generally not released. However, the initial process of most ranking systems is to standardize the data into z-scores. In this study, optimization was carried out to find the optimal values of variables that are in common of indicators for world university rankings namely faculty students ratio, citations per faculty, proportion of international faculty and proportion of international students. The results showed that the optimal value of each variable was changed depending on the boundaries which were based on context, ability including the policy of an individual educational institution. Nevertheless, the rough estimation is what we obtained from this research for getting the concept of the optimal value of each variable to achieve a better world university ranking.

References

Bougnol, M. L. & Dulá, J. H. (2013). A mathematical model to optimize decisions to impact multi-attribute rankings. *Scientometrics, 95*(2), 785-796.

Cakir, M. P., Acarturk, C., Alasehir, O. & Cilingir, C. (2015). A comparative analysis of global and national university ranking systems. *Scientometrics*, *103*(3), 813-848.

Carey, J. J., & Delaney, M. F. (2010). T-Scores and Z-Scores. *Clinical Reviews in Bone and Mineral Metabolism* , *8*(3), 113-121.

Gill, P. E., Murray, W., & Wright, M. H. (1984). *Practical Optimization*. New York: Academic Press Inc.

Huang, M. H. (2011). A Comparison of Three Major Academic Rankings for World Universities: From a Research Evaluation Perspective. *Journal of Library and Information Studies, 9*(1), 1-25.

Khor, K. A. & Yu, L. G. (2016). Influence of international co-authorship on the research citation impact of young universities. *Scientometrics, 107*(3), 1095-1110.

Luenberger, D. G. & Ye, Y. (1984*). Linear and nonlinear programming*. (2nd ed.). Reading, MA: Addison-Wesley.

Malekzadeh, P. & Gore, S. D. (2012). A stochastic Approach to General Nonlinear Programming Problems. *Communications in Statistics - Simulation and Computation*, *41*(10), 1989-1997.

Marginson, S. (2014). University Rankings and Social Science. *European Journal of Education, 49*(1), 45-59.

Morse, R., Krivian, A. & Jackwin, A. (2016). *How U.S. News Calculated the Best Global Universities Ranking*. http://www.usnews.com/ education/best-global-universities/ articles/methodology. Accessed 24 October 2016.

Riddle, W. R. & DonLevy, S. C. (2010). Generating expected growth curves and Z-scores for premature infants. *Journal of Perinatology, 30*(11), 741-750.

Shehatta, I. & Mahmood, K. (2016). Correlation among top 100 universities in the major six global rankings: policy implications. *Scientometrics, 109*(2), 1231-1254.

Tijssen, R. J., Yegros-Yegros, A. & Winnink, J. J. (2016). University-industry R&D linkage metrics: validity and applicability in world university rankings. *Scientometrics, 109*(2), 677-696.

VITAE

Name Miss Pattira Jubjaimoh

Student ID 5710220072

Educational Attainment

Degree	Name of Institution	Year of Graduation
B.Sc. (Mathematics)	Thammasat University	2009

Scholarship Awards during Enrolment

Teaching Assistantship from Faculty of Science, Prince of Songkla University, 2014-2016.

List of Publication and Proceeding Publications

1. Jubjaimoh, P., Samart, K., Jansakul, N., and Panityakul, T. A Comparison of Forecasting Techniques used to Predict World University Rankings Tendency of Prince of Songkla University, *The International Conference on Applied Statistics 2016 "Solving Real World Problems with Statistics" (ICAS 2016).*, Phuket Graceland Resort & Spa, Phuket, Thailand, July 13-15, 2016, O221-O227.
2. Jubjaimoh, P., Samart, K., Jansakul, N., and Jibenja, N. Optimization for Better World University Ranking. Submitted Manuscript.