

# A MATLAB Toolbox for Misconceptions Analysis Based on S-P Chart, Grey Relational Analysis and ROC

**Tian-Wei Sheu, Phuoc-Hai Nguyen, Phung-Tuyen Nguyen, Duc-Hieu Pham, Ching-Pin Tsai and Masatake Nagai**

*Graduate Institute of Educational Measurement, National Taichung University of Education, Taichung, Taiwan;*

sheu@mail.ntcu.edu.tw, phuochai1979@gmail.com, tuyennguyenphung@gmail.com, hieuhagdthsp2@gmail.com, tsai.chenbin@msa.hinet.net, nagai@kamakuranet.ne.jp

## ABSTRACT

The purpose of this study is to design a MATLAB toolbox for misconceptions analysis based on Student-Problem Chart (S-P Chart), Grey Relational Analysis (GRA), and Receiver Operating Characteristic (ROC). Misconceptions analysis is actually necessary for teachers as well as students to improve teaching and learning process. Moreover, it is not only an objective research method, but also can evaluate the comprehensive academic achievement of students, and can select problems for the test. However, there is a difficult problem of the calculation for this method. Therefore, the design of a MATLAB toolbox for misconceptions analysis is necessary. This toolbox not only helps to process data quickly and accurately, but also displays the results on a graphical user interface visually. The experimental results show that this toolbox is actually useful for teachers in analyzing and evaluating the learning performance of students, and helps researchers to perform their educational researches more convenient.

**Keywords:** MATLAB toolbox, misconceptions analysis, S-P Chart, GRA, ROC.

## 1 INTRODUCTION

In the field of education, when teachers provide instruction on concepts in various subjects, they are teaching students who already have some pre-instructional knowledge about the topic. Knowledge of students can be erroneous, illogical or misinformed. These erroneous understandings are termed alternative conceptions or misconceptions. Misconceptions analysis is an integral and vital part of teaching and learning, providing feedback on progress through the assessment period to both students and teachers [1].

In recent years, many studies of misconceptions have been carried out in different countries and at all educational levels from elementary school to universities [2]. Thus, the study of methods to determine misconceptions is necessary for misconceptions analysis. In 2013, Sheu, T. W., et al., [3, 4] proposed a method for misconceptions analysis based on Student-Problem Chart, Grey Relational Analysis, and Receiver Operating Characteristic, that produced an analysis from a small amount of students and problems. It points out that which areas are the most important to address in the structure of concepts. Teachers can use this system as a reference when conducting remedial teaching. At this time, this method is becoming more relevant in education. This method provides teachers both effective and objective tool for research in teaching.

In 1969, Takahiro Sato proposed S-P Chart that could be used not only to display the diagnostic assessment of learning, but also give full play to improve the effectiveness of the curriculum. The S-P chart analysis could provide the two caution indices. One is the caution index for students (*CS*), and the other is the caution index for problems (*CP*). With these two indices, the researchers could select items with proper qualities and understand the cognition and learning condition of students [5]. Deng first introduced the grey system theory in 1982, and then it is applied to treat small amount insufficient information data. The data can be transformed from grey to white, and use to prediction and decision-making processes in the theory. Besides, the satisfactory prediction results with partially unknown parameters can be obtained. In grey system theory, there is a method of grey relational analysis which can treat uncertain, multiple, discrete, and incomplete information effectively. The main function of GRA is to calculate the discrete data and quantify the factors [6, 7]. Receiver operating characteristics graph is useful for organizing classifiers and visualizing their performance, it is commonly used in medical decision making. In recent years, it has been used increasingly in machine learning and data mining research [8]. The area under the ROC curve (*AUC*) is frequently used as a measure for the effectiveness of diagnostic markers. Its value can be interpreted as the probability that a classifier is able to distinguish a randomly chosen positive example from a randomly chosen negative example [9]. *AUC* is a measure for the diagnostic accuracy of a test and is often used to make comparisons between diagnostic tests or between observers [10].

In this paper, the researchers used a MATLAB toolbox based on Student-Problem Chart, Grey Relational Analysis, and Receiver Operating Characteristic to analyze misconceptions from the test results of the seventh grade mathematics for a class in Central Taiwan. This study is not only focused on the remedial instruction, but also noted how to enhance the quality of teaching strategies and how to choose students' learning misconceptions, and provide the teaching ideas to achieve the most effective teaching.

## 2 BASIC THEORIES

### 2.1 Student - Problem Chart Theory

The S-P Chart theory was first proposed by Takahiro Sato in 1969 [11]. Sato's method is to do analysis and sorting based on students' response pattern in the items. The analysis and sorting results can be used to get item quality and the students' diagnostic messages for teachers' effective learning [12]. The coefficient used for individual students and problems by S-P chart is caution index which is provided in caution index for student (CS) and caution index for item (CP). Caution index is mainly used as an index to judge students or questions whether are anomalies in response patterns. CS and CP are defined as follows [13, 14] (as shown in Table 1). Let item responses be

$$X = [x_{ij}], \quad x_{ij} = \begin{cases} 1, & \text{if answer is right} \\ 0, & \text{if answer is wrong} \end{cases} \quad (1)$$

and the number of student is  $S_i, i = 1, 2, \dots, m$ , the number of problem is  $P_j, j = 1, 2, \dots, n$ .

Caution index for students (CS) are defined as follow:

$$CS_i = 1 - \frac{\sum_{j=1}^n (x_{ij})(x_{\bullet j}) - (x_{i\bullet})(\bar{x})}{\sum_{j=1}^n (x_{\bullet j}) - (x_{i\bullet})(\bar{x})} \quad (2)$$

where  $\bar{x} = \frac{1}{n} \sum_{j=1}^n x_{\bullet j}$  and  $N = x_{i\bullet} = \sum_{j=1}^n x_{ij}$

Table 1: S-P Chart

	Problem number $P_j, j = 1, 2, \dots, n$	Total score	CS
Student number $S_i, i = 1, 2, \dots, m$	$X = [x_{ij}]_{m \times n}$	High ↕ Low	$CS_i$
Number of correct answer	More ← $PP_j$	$\sum_{i=1}^m SS_i = \sum_{j=1}^n PP_j$	
CP	$CP_j$		

and caution index for problems (CP):

$$CP_j = 1 - \frac{\sum_{i=1}^m (x_{ij})(x_{i\bullet}) - (x_{\bullet j})(\bar{x}')}{M \sum_{i=1}^m (x_{i\bullet}) - (x_{\bullet j})(\bar{x}')} \tag{3}$$

where  $\bar{x}' = \frac{1}{m} \sum_{i=1}^m x_{i\bullet}$  and  $M = x_{\bullet j} = \sum_{i=1}^m x_{ij}$

## 2.2 Grey Relational Analysis

The GRA method can make the analysis more concrete and accurate, and the uncertain factors in the studies can also be analyzed [12]. The GRA procedures are shown as follows.

### 2.2.1 Establishing the raw data analysis

For setting up the GRA, there must be reference vector  $x_0$  and comparative vector  $x_i$ , and they are shown as follows.

$$x_0 = (x_0(1), x_0(2), \dots, x_0(k), \dots, x_0(m)) \tag{4}$$

where  $k = 1, 2, \dots, m$

$$\left. \begin{aligned} x_1 &= (x_1(1), x_1(2), \dots, x_1(k), \dots, x_1(m)) \\ x_2 &= (x_2(1), x_2(2), \dots, x_2(k), \dots, x_2(m)) \\ &\vdots \\ x_i &= (x_i(1), x_i(2), \dots, x_i(k), \dots, x_i(m)) \\ &\vdots \\ x_n &= (x_n(1), x_n(2), \dots, x_n(k), \dots, x_n(m)) \end{aligned} \right\} \tag{5}$$

where  $i = 1, 2, \dots, n$

### 2.2.2 Generations of grey relation

During the GRA process, researchers need to extract the available data by satisfying three rules, and they are: non-dimension, scaling and polarization. There are three methods to generate and standardize the data, and they are: larger-the-better, smaller-the-better, and nominal-the-better. In this paper, the researchers only use larger-the-better method. It is shown as follows.

$$x_i^*(k) = \frac{x_i(k) - \min_{\forall i} x_i(k)}{\max_{\forall i} x_i(k) - \min_{\forall i} x_i(k)} \tag{6}$$

where  $\max_{\forall i} x_i(k)$  means the maximum number in  $j$ , and  $\min_{\forall i} x_i(k)$  means the minimum number in  $j$ .

### 2.2.3 Calculations of Grey Relational

Based on Nagai's equation [15], the grey relation can be calculated in this paper when partial grey relation's reference vector is  $x_0$  and comparative vector is  $x_j$ . When  $\gamma_{0i}$  is close to 1, it means that  $x_0$  and  $x_j$  are highly related to each other. The equation of the localized grey relation is shown as follows.

$$\gamma_{0i} = \gamma(x_0(k), x_i(k)) = \frac{\bar{\Delta}_{\max} - \bar{\Delta}_{0i}}{\bar{\Delta}_{\max} - \bar{\Delta}_{\min}} \quad (7)$$

where  $\bar{\Delta}_{0i} = \|x_{0i}\|_{\rho} = \left( \sum_{k=1}^n [\Delta_{0i}(k)]^{\rho} \right)^{\frac{1}{\rho}}$ ,  $\bar{\Delta}_{\max}$  represents the maximum of  $\bar{\Delta}_{0i}$ , and  $\bar{\Delta}_{\min}$  represents the minimum of  $\bar{\Delta}_{0i}$ . When  $\rho \geq 1, 2, \dots, m$ , it means Minkowski's grey relation; when  $\rho = 2$ , it means Euclidean grey relation.

### 2.2.4 Grey relational grade

The decision-making is based on the comparison of the grey relation. Though ranking, we can find out an importance of the factors and then identify the maximum or minimum of impact factors. The factors become the associated principle of the system.

## 2.3 Receiver Operating Characteristic

Receiver operating characteristics are a powerful statistical modeling tool which can be used in decision making, particularly when setting threshold values for tests. Receiver operating characteristic curves are a graphical plot of sensitivity of a test and 1-specificity for a binary classifier system as the test threshold varies. In order to construct an ROC curve the researchers need to calculate the  $Se$  and  $Sp$  of the test for each possible cut point value, and they are shown as follows.

Table 2: The contingency table

	Actual Value	
Prediction Outcome	True Positive (a)	False Positive (b)
	False Negative (c)	True Negative (d)

$$\text{Sensitivity (Se)} = \frac{a}{a+b} \quad (8)$$

$$\text{Specificity (Sp)} = \frac{d}{c+d} \quad (9)$$

$$AUC = \frac{Se \cdot (1 - Sp)}{2} + \frac{(Se + 1) \cdot Sp}{2} \quad (10)$$

The area under the ROC curve ( $AUC$ ) is a summary statistic of diagnostic performance. The  $AUC$  could distinguish between non-predictive ( $AUC < 0.5$ ), less predictive ( $0.5 \leq AUC < 0.7$ ), moderately predictive ( $0.7 \leq AUC < 0.9$ ), highly predictive ( $0.9 \leq AUC < 1$ ) and perfect prediction ( $AUC = 1$ ) [16-17].

### ROC Analysis Method

The analysis steps of ROC as follows:

**Step 1:** Establishing the response result of students

$$\text{Let } X = [x_{ij}]_{m \times n}, x_{ij} = \begin{cases} 1, & \text{if answer is right} \\ 0, & \text{if answer is wrong} \end{cases} \quad (11)$$

be the response result of students, analysis results of S-P Chart are shown in Table 1.

**Step 2:** The set of positive and negative class labels of students and problems are defined in definitions 1.

#### Definition 1: The prediction matrix for students

The state setting of positive "1" and negative "0" in prediction matrix  $[d_{ij}]_{m \times n}$  is determined by the following formula:

$$D_S = [d_{ij}]_{m \times n}, d_{ij} = \begin{cases} 1, & \text{if } j \leq SS_i \\ 0, & \text{if } j > SS_i \end{cases} \quad (12)$$

where  $SS_i$  is the score of  $i$ -th student in S-P chart,  $D_S = [d_{ij}]_{m \times n}$  is the matrix that predicts states of positive and negative of students.

#### Definition 2: The prediction matrix for problems

The state setting of positive "1" and negative "0" in prediction matrix  $[d'_{ij}]_{m \times n}$  is determined by the following formula:

$$D_P = [d'_{ij}]_{m \times n}, d'_{ij} = \begin{cases} 1, & \text{if } i \leq PP_j \\ 0, & \text{if } i > PP_j \end{cases} \quad (13)$$

where  $PP_j$  is the number of correct answer of  $j$ -th problem in S-P chart,  $D_P = [d'_{ij}]_{m \times n}$  is the matrix that predicts states of positive and negative of problems.

**Step 3:** The researchers calculated AUC values based on the response result of students  $[x_{ij}]_{m \times n}$  with the prediction matrix  $[d_{ij}]_{m \times n}$  and  $[d'_{ij}]_{m \times n}$ . When the student answered all problems correctly for the whole test, his (her) AUC value is set of "1", in contrast, when he (she) did not answer any problem correctly, his (her) AUC value is set of "0". Similarly, when the problem is answered correctly by all students in the group of student surveyed, its AUC value is set of "1", in contrast, when it is not answered correctly by any student, its AUC value is set of "0". According to the AUC values, the researchers analyzing items through the epsilon ( $\epsilon$ ) threshold was equal to 0.7 in this study. Because the AUC could distinguish between non-predictive ( $AUC < 0.5$ ), less predictive ( $0.5 \leq AUC < 0.7$ ), moderately predictive ( $0.7 \leq AUC < 0.9$ ), highly predictive ( $0.9 \leq AUC < 1$ ) and perfect prediction ( $AUC = 1$ ) [18].

## 2.4 The Method of Analyzing Misconceptions

The synthesis of the results of the S-P Chart, GRA and ROC allows to analyze the misconceptions based on the given conditions as follows:

### Definition 3: Misconceptions of students

Misconceptions of students ( $M_S$ ) that are defined by the following formula:

$$M_S = \left\{ x_i \left| \begin{array}{l} \gamma_S > 0.5, CS_i \geq 0, AUC_i < \epsilon \\ \gamma_S \leq 0.5, 0 < CS_i \leq 0.75, AUC_i < \epsilon \end{array} \right. \right\} \quad (14)$$

where  $i = 1, 2, \dots, m$ .

In misconceptions of students ( $M_S$ ), the students have Gamma value ( $\gamma_S$ ) that are greater than 0.5 from S-P Chart, caution index for students ( $CS_i$ ) that are greater than or equal to 0, and AUC values ( $AUC_i$ ) that are less than the epsilon ( $\epsilon$ ) threshold; and the students have Gamma value ( $\gamma_S$ ) that are less than or equal to 0.5 from S-P Chart, caution index for students ( $CS_i$ ) that are greater than 0 and less than or equal to 0.75, and AUC values ( $AUC_i$ ) that are less than the epsilon ( $\epsilon$ ) threshold [19-22].

### Definition 4: Misconceptions of problems

Misconceptions of problems ( $M_P$ ) that are defined by the following formula:

$$M_P = \left\{ x_j \left| \begin{array}{l} \gamma_P > 0.5, CP_j \geq 0, AUC_j < \epsilon \\ \gamma_P \leq 0.5, 0 < CP_j \leq 0.75, AUC_j < \epsilon \end{array} \right. \right\} \quad (15)$$

where  $j = 1, 2, \dots, n$ .

In misconceptions of problems ( $M_P$ ), the problems have Gamma value ( $\gamma_P$ ) that are greater than 0.5 from S-P Chart, caution index for problems ( $CP_j$ ) that are greater than or equal to 0, and AUC values ( $AUC_j$ ) that are less than the epsilon ( $\epsilon$ ) threshold; and the problems have Gamma value ( $\gamma_P$ ) that are less than or equal to 0.5 from S-P Chart, caution index for students ( $CP_j$ ) that are greater than 0 and less than or equal to 0.75, and AUC values ( $AUC_j$ ) that are less than the epsilon ( $\epsilon$ ) threshold [19-22].

### 3 BUILDING TOOLBOX

#### 3.1 Software Specifications and Requirements

- Windows 7, Windows 8 or upgrade versions.
- Screen resolution 1280×800.
- MATLAB R2012b version or upgrade versions.

This paper presents a sample program that is developed by MATLAB including many scientific functions due to supply an experiment environment on the computer, and then a reliable program can be developed. The program for misconceptions analysis has been developed by MATLAB R2012b software. Using this program depends on the user's knowledge about MATLAB.

#### 3.2 The Program for Misconceptions Analysis

The program for misconceptions analysis has 6 basic steps.

**Step 1:** Input data.

Data is the response matrix of students that is numerical and written in \*.csv file or \*.xlsx file.

**Step 2:** Calculate student total, problem total, and Cronbach's Alpha value.

**Step 3:** S-P Chart.

Calculate students' total scores, correct answer of problems, and correct ratio. Next, arrange in order of students' total scores, and correct answer of problems. Then, calculate caution index for Student (*CS*), and Caution Index for Problem (*CP*), and design the table for results.

**Step 4:** Grey Relational Analysis (GRA).

Calculate GRA for students, and GRA for problems. Next, arrange in order of GRA for students, and GRA for problems. Then, drawing GRA graphs for students and problems.

**Step 5:** Receiver Operating Characteristic (ROC).

Calculating Sensitivity (*Se*), and Specificity (*Sp*). Next, calculate the area under the ROC curve (*AUC*). Then, draw ROC graphs for students and problems.

**Step 6:** Continue or exit program.

If new data is selected, the program will continue and back to step 1, or else the program will be ended. The main algorithms of the program are described in figure 1. The user can save the results and the graphs to \*.csv file or \*.xlsx file and jpeg image formatted.

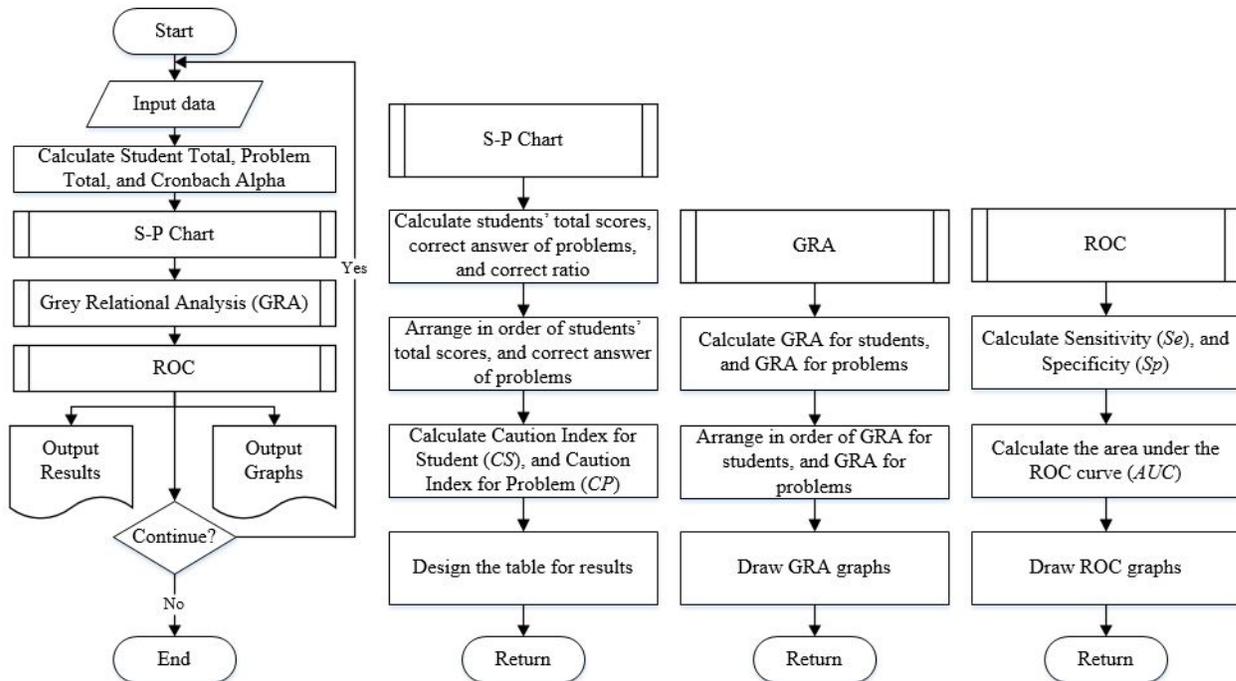


Figure 1: The flowchart for misconceptions analysis

## 4 EXPERIMENTAL RESULTS AND DISCUSSION

### 4.1 Working with the MATLAB Toolbox

The graphical user interface starts up when file “SPChart\_GRA\_ROC.m” is opened, the graphical user interface of program is displayed on the screen of the computer (Fig. 2 and Fig. 3). Firstly, the user needs to click the button “Input Data” to input the data. Data is the response matrix of students that is numerical and written in \*.csv file or \*.xlsx file. Secondly, the user clicks the button “Graphs for Students” or “Graphs for Problems”, and clicks the button “OK”, the program will process data, and then program displays the results and the graphs. In this panel, the user can clicks the button “Save Results” to save the results with \*.csv file or \*.xlsx file formatted, and clicks the button “Save Graphs” to save the graphs with \*.JPG file formatted. Finally, if new data is selected, the program will continue, or else if the user clicks the button “Exit” the program will be ended.

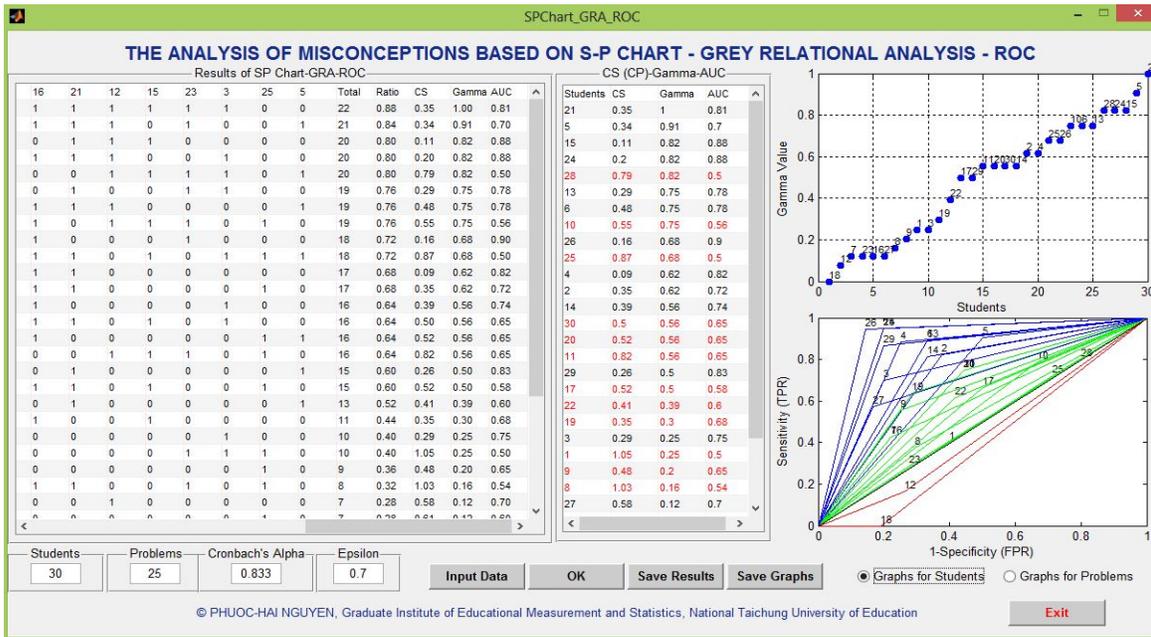


Figure 2: The graphical user interface of the toolbox for the results and the graphs of students

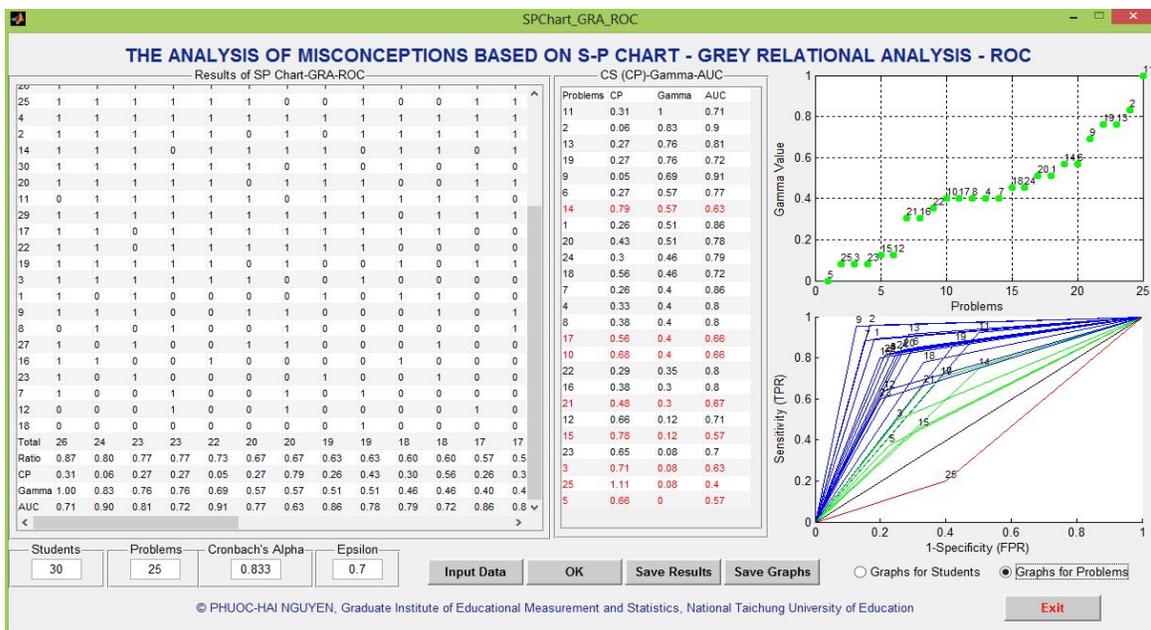


Figure 3: The graphical user interface of the toolbox for the results and the graphs of problems

## 4.2 Application Example

In this paper, the researchers used a MATLAB toolbox based on Student-Problem Chart, Grey Relational Analysis, and Receiver Operating Characteristic to analyze misconceptions from the test results of the seventh grade mathematics for a class in Central Taiwan. The number of students was 30 students and the number of items was 25 problems. This study carried out a reliability test for the answers by using Cronbach's Alpha coefficient. The results showed that

Cronbach's Alpha value of the test was  $\alpha = 0.833$ , which meant the reliability of the information was quite good. The results of the test is shown in table 3.

According to the results from figure 2 and table 3 showed that 17 students having *AUC* values are less than 0.7 ( $S_{28}, S_{10}, S_{25}, S_{30}, S_{20}, S_{11}, S_{17}, S_{22}, S_{19}, S_1, S_9, S_8, S_{16}, S_{23}, S_7, S_{12}$ , and  $S_{18}$ ). However, six students having *Gamma* values are less than 0.5 and *CS* values are greater than 0.75 ( $S_1, S_8, S_{23}, S_7, S_{12}$ , and  $S_{18}$ ). The results from figure 3 and table 3 showed that eight problems having *AUC* values are less than 0.7 ( $P_{14}, P_{17}, P_{10}, P_{21}, P_{15}, P_3, P_{25}$ , and  $P_5$ ). However, two problems having *Gamma* value are less than 0.5 and *CP* value is greater than 0.75 ( $P_{15}$  and  $P_{25}$ ). The researchers determined misconceptions based on the given conditions of *Gamma* value, *CS* value, *CP* value, and *AUC* value of students and problems for this class. The result of analyzing misconceptions showed that 11 students have misconceptions ( $S_{28}, S_{10}, S_{25}, S_{30}, S_{20}, S_{11}, S_{17}, S_{22}, S_{19}, S_9$ , and  $S_{16}$ ), and six problems have misconceptions in this class ( $P_{14}, P_{17}, P_{10}, P_{21}, P_3$ , and  $P_5$ ). The results from figure 2, figure 3, figure 4, and table 3 showed that this toolbox not only helps to process data accurately, but also displays the results on a graphical user interface visually.

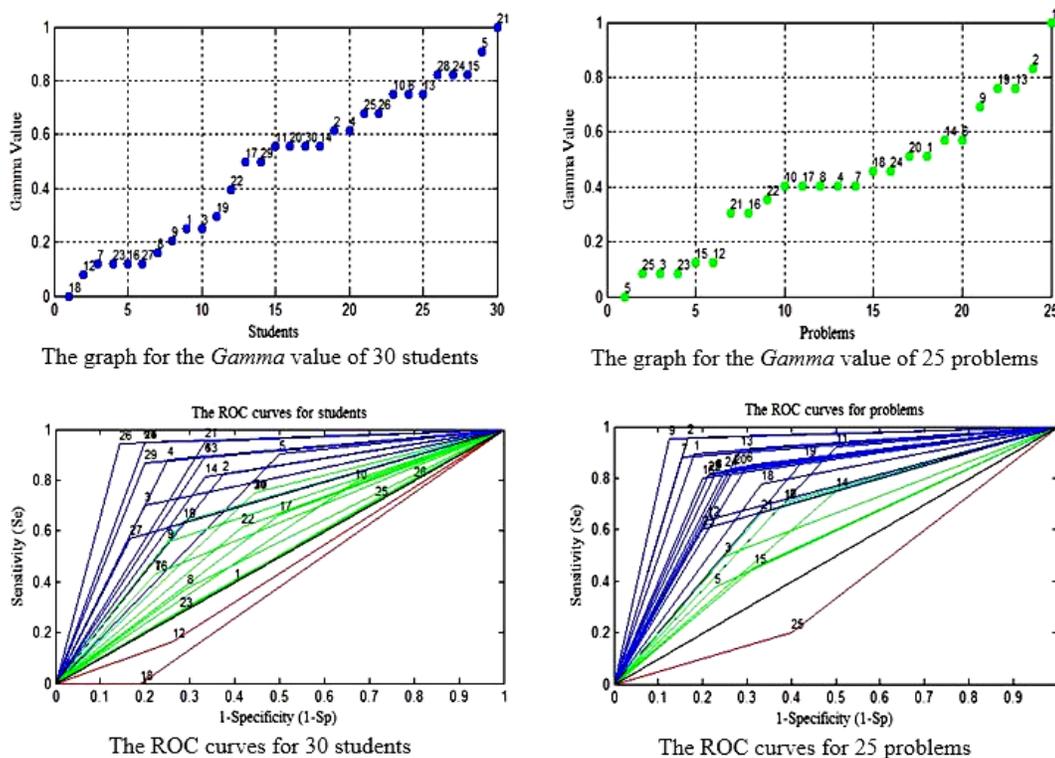


Figure 4: The graphs of application example

**Table 3: The results of application example**

S \ P	11	2	13	19	9	6	14	1	20	24	18	7	4	8	17	10	22	16	21	12	15	23	3	25	5	Total	Ratio	CS	GS	AUC
21	1	1	1	1	1	1	1	1	1	1	0	1	1	1	1	1	1	1	1	1	1	1	1	0	0	22	0.88	0.35	1.00	0.81
5	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	0	1	1	1	1	0	1	0	0	1	21	0.84	0.34	0.91	0.70
15	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	0	1	1	1	1	0	0	0	0	20	0.80	0.11	0.82	0.88
24	1	1	1	1	1	1	1	1	1	1	1	1	1	0	1	1	1	1	1	1	0	0	1	0	0	20	0.80	0.20	0.82	0.88
28	1	1	1	1	1	1	0	1	1	1	1	1	1	1	0	1	1	0	0	1	1	1	1	0	1	20	0.80	0.79	0.82	0.50
13	1	1	1	1	1	1	1	1	1	1	1	0	1	1	1	1	0	1	0	0	1	1	0	0	0	19	0.76	0.29	0.75	0.78
6	1	1	1	1	1	1	0	1	0	1	1	1	1	1	1	1	1	1	1	1	0	0	0	0	1	19	0.76	0.48	0.75	0.78
10	1	1	1	1	1	1	1	1	1	1	1	1	0	1	1	0	0	1	0	1	1	1	0	1	0	19	0.76	0.55	0.75	0.56
26	1	1	1	1	1	1	1	1	1	1	1	1	1	1	0	1	1	1	0	0	0	1	0	0	0	18	0.72	0.16	0.68	0.90
25	1	1	1	1	1	1	0	0	1	0	0	1	1	0	1	1	1	1	1	0	1	0	1	1	1	18	0.72	0.87	0.68	0.50
4	1	1	1	1	1	1	1	1	1	1	1	1	1	1	0	0	1	1	1	1	0	0	0	0	0	17	0.68	0.09	0.62	0.82
2	1	1	1	1	1	0	1	0	1	1	1	1	1	1	0	1	1	1	1	0	0	0	0	0	1	17	0.68	0.35	0.62	0.72
14	1	1	1	0	1	1	1	1	0	1	1	0	1	1	1	1	1	1	0	0	0	0	0	1	0	16	0.64	0.39	0.56	0.74
30	1	1	1	1	1	1	0	1	0	1	0	1	0	1	1	1	0	1	1	0	1	0	1	0	0	16	0.64	0.50	0.56	0.65
20	1	1	1	1	1	0	1	1	1	0	0	1	1	1	0	1	1	1	0	0	0	0	0	0	1	16	0.64	0.52	0.56	0.65
11	0	1	1	1	1	1	0	1	1	1	1	1	0	1	1	0	0	0	0	1	1	1	0	1	0	16	0.64	0.82	0.56	0.65
29	1	1	1	1	1	1	1	1	1	0	1	1	1	0	1	0	0	0	1	0	0	0	0	0	1	15	0.60	0.26	0.50	0.83
17	1	1	0	1	1	1	1	1	1	1	0	0	0	0	0	0	1	1	1	0	1	0	0	0	1	15	0.60	0.52	0.50	0.58
22	1	1	0	1	1	1	1	1	1	0	0	0	0	1	1	1	0	0	1	0	0	0	0	0	1	13	0.52	0.41	0.39	0.60
19	1	1	1	1	1	0	1	0	0	1	0	1	1	0	0	0	0	1	0	0	1	0	0	0	0	11	0.44	0.35	0.30	0.68
3	1	1	1	1	1	1	0	0	1	0	0	0	0	0	1	1	0	0	0	0	0	0	1	0	0	10	0.40	0.29	0.25	0.75
1	1	0	1	0	0	0	0	1	0	1	1	0	0	0	1	0	1	0	0	0	0	1	1	1	0	10	0.40	1.05	0.25	0.50
9	1	1	1	0	0	1	1	0	0	0	1	0	1	0	1	0	0	0	0	0	0	0	0	1	0	9	0.36	0.48	0.20	0.65
8	0	1	0	1	0	0	1	0	0	0	0	0	1	0	0	0	0	1	1	0	0	1	0	1	0	8	0.32	1.03	0.16	0.54
27	1	0	1	0	0	1	1	0	0	0	1	0	0	0	1	0	0	0	0	1	0	0	0	0	0	7	0.28	0.58	0.12	0.70
16	1	1	0	0	1	0	0	0	0	1	0	0	0	1	0	1	0	0	0	0	0	0	0	0	1	7	0.28	0.61	0.12	0.60
23	1	0	1	0	0	0	0	1	0	0	1	0	0	0	0	1	0	0	0	1	1	0	0	0	0	7	0.28	0.83	0.12	0.50
7	1	0	0	1	0	0	1	0	1	0	0	0	0	0	1	0	0	0	0	0	0	1	0	0	1	7	0.28	0.88	0.12	0.60
12	0	0	0	1	0	0	1	0	0	0	0	1	0	0	0	0	1	0	1	0	0	0	1	0	0	6	0.24	1.01	0.08	0.45
18	0	0	0	0	0	0	0	0	1	0	0	0	0	0	0	1	0	0	0	1	1	0	0	0	0	4	0.16	1.34	0.00	0.40
<b>Total</b>	<b>26</b>	<b>24</b>	<b>23</b>	<b>23</b>	<b>22</b>	<b>20</b>	<b>20</b>	<b>19</b>	<b>19</b>	<b>18</b>	<b>18</b>	<b>17</b>	<b>17</b>	<b>17</b>	<b>17</b>	<b>17</b>	<b>16</b>	<b>15</b>	<b>15</b>	<b>11</b>	<b>11</b>	<b>10</b>	<b>10</b>	<b>10</b>	<b>8</b>	<b>423</b>				
<b>Ratio</b>	0.870	0.800	0.770	0.770	0.730	0.670	0.670	0.630	0.630	0.600	0.600	0.570	0.570	0.570	0.570	0.570	0.530	0.500	0.500	0.370	0.370	0.330	0.330	0.330	0.27					
<b>CP</b>	0.310	0.060	0.270	0.270	0.050	0.270	0.790	0.260	0.430	0.300	0.560	0.260	0.330	0.380	0.560	0.680	0.290	0.380	0.480	0.660	0.780	0.650	0.711	1.10	0.66					
<b>GP</b>	1.000	0.830	0.760	0.690	0.570	0.570	0.510	0.510	0.460	0.460	0.400	0.400	0.400	0.400	0.350	0.300	0.300	0.120	0.120	0.080	0.080	0.00	0.00	0.00	0.00					
<b>AUC</b>	0.710	0.900	0.810	0.720	0.910	0.770	0.630	0.860	0.780	0.790	0.720	0.860	0.800	0.800	0.660	0.660	0.800	0.800	0.670	0.710	0.570	0.700	0.630	0.400	0.57					

S = 30  
P = 25

## 5 CONCLUSION

Based on the findings from this study, some viewpoints and suggestions are discussed as follows.

- This study has successfully developed a MATLAB toolbox for misconceptions analysis based on S-P Chart, GRA, and ROC. This toolbox has many advantages such as: easy for use, time-saving, accurate and clearly visual output. Especially, the user can save the results and the graphs to \*.csv file or \*.xlsx file and jpeg image formatted.
- This toolbox is not only an objective research method, but also helps teachers to diagnose the situation exploration of student learning. This will be essential tool to innovate teaching methods and training quality management in education.
- This toolbox is actually useful for teachers as well as students to improve teaching and learning process. Teachers can use it as a reference when conducting remedial teaching. At this time, this toolbox is becoming more relevant in education. This method provides teachers both effective and objective tool for research in teaching.

In conclusion, this toolbox is actually useful for teachers in analyzing and evaluating the learning performance of students, selecting problems in the teaching process, and helps researcher convenience to conduct their educational researches.

## REFERENCES

- [1]. Bambico, T., Mathematical Strengths, Difficulties and Misconceptions of Teachers: Analysis of Their Performance in an Achievement Test. *Journal of International Development and Cooperation*, 2002. 9(1): p. 41-60.
- [1]. Çelikten, O., S. İpekçioğlu, H. Ertepinar, and Ö. Geban, The effect of the conceptual change oriented instruction through cooperative learning on 4th grade students' understanding of earth and sky concepts. *Science Education International*, 2012. 23(1): p. 84-96.
- [2]. Sheu, T. W., P. H. Nguyen, P. T. Nguyen, D. H. Pham, C. P. Tsai, and M. Nagai, Using S-P Chart, Grey Relational Analysis, and Receiver Operating Characteristic to Analyze Misconceptions. *International Conference on Grey System Theory and Kansei Engineering Conference*, 2013. p. 25-32.
- [3]. Sheu, T. W. P. H. Nguyen, P. T. Nguyen, D. H. Pham, C. P. Tsai, and M. Nagai, The Analysis of Misconceptions Based on S-P Chart, Grey Relational Analysis, and Receiver Operating Characteristic. *International Journal of Kansei Information*, 2014. 5(1): p. 1-12.
- [4]. Lin, Y. H., and S. M. Chen, "The Investigation of S-P Chart Analysis on the Test Evaluations of Equality Axiom Concepts for Sixth Graders," *Proceedings of the 2nd WSEAS/IASME International Conference on Educational Technologies*, Bucharest, Romania, 2006.
- [5]. Deng, J. Introduction to Grey System Theory. *Journal of Grey System*, 1989. 1: p. 1-24.
- [6]. Liang, J. C., T. W. Sheu, J. W. Tzeng, B. T. Wang, and M. Nagai, 5W1H, GRA and GSM in the evaluation and identify on optimal design of bike lamps. *Journal of Convergence Information Technology (JCIT)*, 2011. 6(12): p. 266-274.
- [7]. Fawcett, T., An introduction to ROC analysis. *Pattern Recognition Letters*, 2006. 27: p. 861-874.
- [8]. Faraggi, D., and B. Reiser, Estimation of the area under the ROC curve. *Statistics In Medicine*, 2001. 21: p. 3093-3106.
- [9]. Nguyen, P. H., T. W. Sheu, P. T. Nguyen, D. H. Pham, C. P. Tsai, and M. Nagai, Based on WPOT Combines with ROC Concepts Structure Analysis and Cognitive Diagnostic. *Kansei Information of Creative Contest*, Chien Kuo Technology University, Taichung, Taiwan, 2013. p. 21-29.
- [10]. Sheu, T. W., T. L. Chen, J. W. Tzeng, C. P. Tsai, and M. Nagai, Study on the Conception of Learning Problems of Students by Combining the Misconception Domain and Structural Analysis Methods. *Journal of Computers*, 2013. 8(5): p. 1255-1266.

- [11]. Sheu, T. W., P. H. Nguyen, P. T. Nguyen, and D. H. Pham, A Matlab Toolbox for AHP and LGRA-AHP to Analyze and Evaluate Factors in Making the Decision. *International Journal of Kansei Information*, 2013. 4(3): p. 149-158.
- [12]. Sheu, T. W., D. H. Pham, P. T. Nguyen, and P. H. Nguyen, A Matlab Toolbox for Student-Problem Chart and Grey Student-Problem Chart and Its Application. *International Journal of Kansei Information*, 2013. 4(2): p. 75-86.
- [13]. Sheu, T. W., T. L. Chen, J. W. Tzeng, C. P. Tsai, H. J. Chiang, C. L. Chang, and M. Nagai, Applying Misconception Domain and Structural Analysis to Explore the Effects of the Remedial Teaching. *Journal of Grey System*, 2013. 16(1): p. 17-34.
- [14]. Tzeng, J. W., T. W. Sheu, J. C. Liang, B. T. Wang, and M. Nagai, Study on GSP Chart of the Important Certification Indicators of the E-Learning Courses in Taiwan. *International Journal of Digital Content Technology and Its Application*, 2012. 6(9): p. 9-16.
- [15]. Hanley, J. A., and B. J. McNeil, The Meaning and Use of the Area under Receive Operating Characteristic (ROC) Curve. *European Journal of Radiology*, 1982. 143: p. 29-36.
- [16]. Arian, R., V. Erkel, M. Peter, and T. Pattynama, Receiver operating characteristic (ROC) analysis: Basic principles and applications in radiology. *European Journal of Radiology*, 1998. 27: p. 88-94.
- [17]. Faraggi, D., and B. Reiser, Estimation of the area under the ROC curve. *Statistics In Medicine*, 2001. 21: p. 3093-3106.
- [18]. Sheu, T. W., J. W. Tzeng, C. P. Tsai, and , and T. L. Chen, Applying Problem-Concept Chart Combined with Structural Analysis to Investigate the Learning Misconcept-Simple Equation with One Variable for Example. *Journal of Grey System*, 2012. 15(1): p. 56-66,
- [19]. Sheu, T. W., C. P. Tsai, J. W. Tzeng, T. L. Chen, and M. Nagai, An Algorithm of the Misconception Order. *Applied Mechanics and Materials*, 2013. 284-287: p. 3010-3014.
- [20]. Sheu, T. W., T. L. Chen, C. P. Tsai, J. W. Tzeng, C. P. Deng, and M. Nagai, Analysis of Students' Misconception Based on Rough Set Theory. *Journal of Intelligent Learning Systems and Applications*, 2013. 5:p. 67-83.
- [21]. Sheu, T. W., C. P. Tsai, J. W. Tzeng, T. L. Chen, H. J. Chiang, W. L. Liu, and M. Nagai, Using Rough Set and Grey Structural Model to Investigate the Structure of the Misconceptions' Domain. *Journal of Grey system*, 2012. 15(4): p. 205-220.