SEM, GM and Other Statistic Analyses Concerning Index Values Extracted from Outdoor Agricultural Workers Data

Shinji Kawakura
Department of Information Technology and Human Factors, National Institute of Advanced Industrial Science and Technology, Tsukuba, Japan
Email: s.kawakura@aist.go.jp

Ryosuke Shibasaki
Center for Spatial Information Science, The University of Tokyo, Meguro, Japan
Email: shiba@csis.u-tokyo.ac.jp

***Abstract*—The authors have been grouping and developing various applied sensing systems to solve difficulties in achieving agricultural technical teaching. For these purposes, the authors developed wearable systems containing the latest sensing modules, and measured acceleration data. Their basic data and data concerning subjects' fatigue and subjects' intensity of tasks were collected. The authors analyzed those data using statistics. Because thought that past methods and systems were insufficient to analyze human motion related to users' fundamental information, their feeling of fatigue and of intensity about agricultural works. The authors have obtained and observed time-line hoe acceleration data when subjects have been cultivating using a common hoe. Subsequently, the authors have computed those data using basic statistical methods, as well as SEM (Structural Equation Models) and GM (Graphical Modeling) analysis based on multiple regressions. In future, the members hope to make useful suggestions to the field of agricultural informatics using those data.**

***Index Terms*—multiple regression analysis, SEM and GM analysis, correlation ratio, wearable sensing system, worker impression**

I. Introduction

Human dynamics, electronics, mechanical engineering, IT and statistical techniques of recent years have been used to conduct interventions in various agricultural fields. These flexible combinations have been thought of as beneficial and useful. On the other hand, according to past researches, there were diverse problems related to the serious shortage of young, beginning (novice) agricultural workers and of methods concerning transmitting of skills from experienced workers to inexperienced workers and numerical indexes for directing inexperienced workers.

However, in past studies, there have been insufficient researches and instructional approaches relating to practical, concrete user suggestions for improving, sophisticating physical activities in outer farmlands, especially for inexperienced workers [1]-[6]. Therefore, under these circumstances, our research efforts have been caring for such agricultural workers by utilizing various academic techniques and electronic systems. What is more, many researchers have been realizing them by various computational statistical methods of late years to provide users with numerical, visual and oral feedbacks and various suggestions after worker tasks [7]-[16].

In light of the lessons learned from them, the main theme of this research was to utilize modern statistical analysis (SEM, GM and calculations of correlation ratio) to detect and to show both superficial and hidden relationships and effects concerning tilling movements; it is one of the universal, typical agricultural physical tasks. The authors thought that "flesh-and-blood" workers' movements were extremely important and meaningful; however these mathematical methods would be instrumental in advancing the evolution and popularization of agricultural teaching. In the future, diverse statistical skills with current IT services and electric mechanics (e.g. microcomputers, small sensors and analyzing software) could be indispensable for agricultural farmlands' managers and labors. Furthermore, they must be combined uniformly and flexibly for analyzing records and extracting useful insights and information appropriately.

II. Materials and Methods

The authors reviewed past accomplishments in these fields (academic papers, industrial goods, and patents) and discussed them with researchers, agricultural workers and farmland managers. Next, the authors selected statistical methods to build the hierarchical logics and to analyze data. These methods are standard calculations for average values, variances, standard deviations, simple correlations, correlation ratio, multiple regression, graphs from SEM (Structural Equation Models) and GM (Graphical Modeling) based path analysis, and so on.

Manuscript received March 15, 2017; revised September 18, 2017.

doi: 10.18178/joaat.4.4.296-307

A. Equation

1) Simple correlation

The authors calculated the items in the matrix concerning the values of simple, general correlations (r).

2) Correlation ratio

The authors selected and calculated the square vales of correlation ratio (η^2). Correlation ratio is one of measures of the relationship between the statistical dispersions in individual categories and the dispersion across the whole sample. This measure is thought as the ratio of two standard deviations representing these types of variations. The context here is the same as that of the intra-class correlation coefficient; the value is the square of the correlation ratio.

3) Basic multiple regression

The authors presumed criterion (dependent) variables (Y_1, Y_2, … Y_n) from explanatory (independent) variables (X_1, X_2, … X_n) linearly. The equation concerning them is equation (1). (X: explanatory variable, independent variable, Y: criterion variable, dependent variable)

$$Y_1 = A_1X_{1i} + A_2X_{2i} + \ldots + A_nX_{ni} + e_i \quad (1)$$

(i = 1, 2, … n, A: partial regression coefficient, e_i: error or residual error)

When n = 1, the equation (1) means $Y_i = A_1X_{1i} + e_i$; in short, it is a single regression. To minimize the value of e_i^2, the authors must calculate and set the values of A_1, A_2, … A_n. The authors can explicitly define Ai (see equation (2)) for the equation of this multiple regression.

$$A_i = Si_1\, S_{1y} + S_{i2}\, S_{2y} + \ldots + S_{in}\, S_{ny} \quad (2)$$

(A_i: Cohen regression coefficient of i in the equation for this multiple regression, S_{in}: Some of products of deviations concerning explanatory variables concerning explanatory variables (X_1, X_2, … X_n), S_{ny}: Sum products of deviations concerning Explanatory variables X_{nj} and criterion variable Y, these elements corresponding to the line i and the column n from the determinant of a n×n square matrix)

B. Software

The authors executed multiple regressions successively for the following path analyses by SEM and GM methods by the following programs [17]-[26].

1) Excel GM1.9b.xlsm, Excel SEM1.8b.xlsm, Excel GM1.9β.xlsm (spreadsheets distributed with the book "Excel de Manabu Kyoubunsan Kouzou to Graphical Modeling", Ohmsha (Tokyo, Japan))

2) R (Version 3.1.3, the copyright belongs to "The R Foundation for Statistical Computing")

The precise descriptions concerning 1) and 2) are below.

1) SEM is popular in many disciplines.

The authors selected the Multiple Indicators and Multiple Causes (MIMIC) model and Partial Least Squares (PLS) model, because in this statistical field, these are fundamental and valid methods. The MIMIC model is a special case of a longitudinal SEM. The influences of formative indicators on unobservable latent variables are assessed through their influence and impact on the reflective indicators. These devised statistical models provide the first comprehensive assessment in a framework that simultaneously assesses multiple dimensions. The PLS model is especially suited for situations when data is not normally distributed. The PLS model approach to SEM can offer an alternative to covariance based on SEM. PLS path modelling is referred to as a valid and significant modeling technique with minimal demands regarding measurement scales, sample sizes and residual distributions..

The authors selected various path diagrams made of obtained variables, latent variables, outcome variables, path-coefficient, error variables (these are non-standardized estimates), both side arrows, and one side arrows. In both the MIMIC model and the PLS model, the authors chose "non-standardized estimates." The reason for this choice was to observe the differences concerning variances on vectors directing to each factor, because the value of at least one variance on vectors is 1.

In the GM model, the authors chose "standardized estimates," and the variances of all obtained valuables were 1 for the uniformity and fairness concerning each statistical calculation. Given that the values of six included factors were widely dispersed, mutual effects between these six factors were extremely complex. Additionally, this study was relatively experimental. That was why the authors thought that in each model, for certain, the authors should set the variances of all obtained valuables into 1.

Concerning "non-standardized estimates," the authors can write the multiple regression model as equation (3), while the multiple regression model for "standardized estimates" takes the form of equation (4). The variable Y is the aforementioned outcome variable, and X_1 and X_2 are obtained valuables.

$$Y = a + b_1{*}X_1 + b_2{*}X_2 \quad (3)$$

$$Y = b_1{*}X_1 + b_2{*}X_2 \quad (4)$$

2) R is a language and environment for statistical computing and graphics. It is a GNU project that is similar to the S language and environment that was developed at Bell Laboratories (formerly AT&T, now Lucent Technologies) by John Chambers and colleagues. R could be considered as a different implementation of S. R provides access to a wide variety of statistical (linear and nonlinear modelling, classical statistical tests, time-series analysis, classification, and clustering) and graphical techniques, and is highly extensible.

One of R's strengths is the ease with which well-designed publication-quality plots can be produced, including mathematical symbols and formulae, where needed. Great care has been taken over the defaults for the minor design choices in graphics, but the user retains full control. R is available as Free Software under the terms of the Free Software Foundation's GNU General Public License in source code form. It compiles and runs on a wide variety of UNIX platforms and similar systems (including FreeBSD and Linux), Windows OS, and Mac OS.

In this study, this sequential procedure for using R was described below.

1) Reading and inputting CSV-style raw data
2) Specifying the figure of SEM style path diagram model
3) Calculating this correlation matrix
4) Directing the pattern of SEM model registered in one of the default SEM packages in the R install folder
5) Calculating partial regression coefficients
6) Setting various factors of the path diagram
7) Outputting one DAT-style file to draw the figure of this path diagram
8) Opening free software Rgraphviz, and drawing the aforementioned path diagram

C. Measurement

1) Targeted task

The authors categorized various agricultural tasks from various points of view (e.g., their intensities, generalities). Specifically, the authors researched 1) Work in a sitting position (e.g., cropping (picking) onions), 2) Work in a semi-crouching position (e.g., cultivating (digging, tilling) with a hoe), 3) Work in a standing position (e.g., cutting branches with shears), 4) Work alternating between a sitting and a standing position (e.g., suspending onions on beams in an outhouse). Next, the authors selected the action of 2) "Cultivating with a hoe" as the targeted task, because the action is common around the world, repetitive, and involves intense full-body movement. As shown in Fig. 1, the time line of one set was composed of previous settings, waiting time, trial time, waiting time and post handling time.

The authors thought that this duration, the numbers of tasks, and the whole experiment set were appropriate and well balanced considering various farmland managers' opinions. Each inexperienced subject participated in three trial sets. Each trial comprised 30 swings; the length of each one trial was about one minute. Sets of three trials were conducted successively for the different subjects, on the same day. There was an interval of a few minutes between sets. Before the trials, the authors interviewed the subjects about their daily lives.

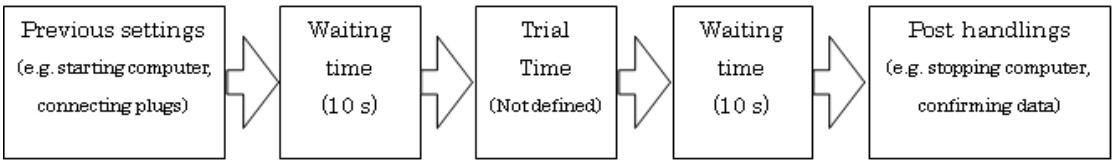

Figure 1. Time line concerning each trial.

2) Indicators and survey sheet

After basic trials, the authors defined ordinary indicators concerning acceleration data in the vertical direction: the maximum value, the minimum value, the Standard Deviation (S.D.), and the Direct Current (DC) component. Such indicators have been used in other studies [5], [6], [11], [12], [15], [27], [28]. In particular, the S.D. and DC components were mentioned as the most significant indicators. For the analysis of acceleration data, the authors used the 95th-percentile values to eliminate outliers derive from various methodological settings and conditions.

Concerning survey sheets in this study, the authors surveyed and chose existing indicators in the field of sports science and physiology. In particular, the authors paid attention to the methods used to obtain numeric data concerning the sentiment and impression of subjects (in past studies, however, there were almost no such fusions of agriculture, statistics, sports science and physiology). Such methods have been used by various researchers, such as the Japan Association of Industrial Health and other health organizations. These researchers (and their methods used) are considered reliable.

Firstly, the authors chose two specific scales: The Visual Analogue Scale (VAS) and The Borg RPE (Rating of Perceived Exertion) Scale to measure worker fatigue and feelings regarding the difficulty of the task. The VAS scale is a psychometric response scale used in questionnaires. It is an instrument for measuring subjective characteristics or attitudes. When responding to a VAS item, respondents specify their level of agreement with a statement by indicating a position along a continuous line between two end-points. In this study, the range of scores used was 0-100. This continuous (or "analog") aspect of the scale differentiated it from discrete scales. The scale RPE measures perceived exertion, in sports and particularly in exercise testing. In medicine, this scale is used to document a patient's exertion during a test, and sports coaches use the scale to assess the intensity of training and competition. The original scale rated exertion on a scale of 6-20. In addition, the authors also asked about subjects' age, prior career, height, weight, fitness habits.

3) System

Past studies have insufficiently addressed subtle changes related to the physical motions of agricultural workers. From the results of past researches [5], [6], [9], [11], [12], [15], [28], the authors designed support systems for agricultural workers, such as a Wearable Sensing system (WS) that included TSND121 multi-sensors (ATR-Promotions Inc., Japan, see Fig. 2 and Fig. 3). The authors used this WS to obtain time series of acceleration data in this study and analyzed the data using original programs written in Visual Studio 2010's Visual Basic 2010.

The sampling rate for obtaining acceleration data was 10 Hz. The measurement time window was 500 ms. the authors decided these values in light of past achievements' data [5], [6], [9], [11], [12], [15], [28]. That module was connected to a small and light (1100 g)

laptop PC (VersaPro VY10A, NEC Inc., Japan) by a USB cable. A specific "sensor server" (distributed by ATR-Promotions Inc., Japan, free system) to mediate between the laptop PC and the TSND121 module was installed and running during these measurements (port number: 10000 or 11000, connection name: localhost, TCP port: System.Net.Sockets.TcpClient). Concerning experienced subjects, the average time span of each trial was around 60 seconds.

Figure 2. Hoe connected to sensors and laptop PC, one subject equipped with measuring modules and knapsack with laptop PC connecting to various modules.

Figure 3. Subject equipped with the sensing system.

The authors surveyed various kinds of typical Japanese hoes before choosing one hoe which tip was rectangular and common in Kantou Area, Japan. After comparing the specification with other hoes, the authors decided that the form, size, and weight of the hoe were typical of hoes used in Japanese agriculture. The specifications were: length = 1000 mm, weight = 1250 g, head length = 290 mm, head width = 180 mm, handle major axis = 30 mm, handle minor axis = 25 mm. The head is made of steel and the handle is made of wood. The authors defined the X, Y and Z directions of the 3-axis acceleration sensor module as shown in Fig. 4, setting direction X as the "vertical acceleration direction."

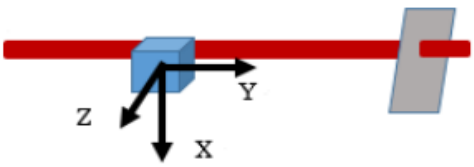

Figure 4. Three-axis direction of the acceleration sensor module attached on a hoe.

Additionally, the authors obtained visual data (video data) for the subjects in a static position from a distance of 3.5 meters. The main aim was to confirm their motions visually later. The authors selected one common (nonspecific) digital video camera, the CANON 410f ixy. The sampling rate was 30 fps on a solid frame. The authors did not selected a high-speed, specific camera and high sampling rate, because the authors have been thinking the later, wide popularization of those methods for common farmlands.

4) Subjects

The authors selected 12 subjects: six in the "inexperienced" group and six in the "experienced" group (see Table I). The authors gathered male experienced and inexperienced workers without any remarkable mental or physical characteristics. In short, they were of average body size for Japanese farmers.

The main criteria of not having any serious diseases, peculiar habits, and specific prior careers (especially in sports and martial arts) were confirmed by preliminary surveillances (hearings). The length of experienced workers' careers were 25-60 years, which were recognized as generally sufficient by many of the farmland managers who the authors interviewed. Experienced workers' Standard Deviation (S.D.) values concerning each index are higher than those of inexperienced workers.

III. Results

In this study, the authors obtained significant statistical data. The authors could contribute to real agricultural managers and workers by showing these data both in real time and after their works.

A. SEM Model

1) Contrast between MIMIC model and PLS model

These path models and their path coefficients below (see Fig. 5 and Fig. 6) were non-standardized MIMIC models and PLS mod els. The MIMIC models contained three phases. The 1st step was the set of "Age", "Career", "Stature" and "Weight". The 2nd step was "Tendency". The 3rd step was "VAS" and "RPE", or "SdAcceHoe1st" and "DcAcceHoe1st", or "SdAcceHoe3rd" and "DcAcceHoe3rd". The PLS models contained four phases. The 1st step was the set of "Age", "Career", "Stature" and "Weight". The 2nd step was "Tendency". The 3rd step was "Working Style". The 4th step was "VAS" and "RPE", or "SdAcceHoe1st" and "DcAcceHoe1st", or "SdAcceHoe3rd" and "DcAcceHoe3rd". The authors defined "Tendency" as the first mediate factor and "Working Style" as the second mediate factor, in short, two phases. The "Tendency" was effected by "Age", "Career", "Stature" and "weight", and effected "Working Style". The "Working Style" was effected by "Tendency", and effected the last two factors (e.g. the set of "VAS" and "RPE").

TABLE I. BASIC INFORMATION RELATED TO SUBJECTS (RANGE, AVERAGE (AVE), STANDARD DEVIATION (S.D.), AVERAGE, Z VALUE (Z), COEFFICIENT OF VARIATION (CV) AND DESCRIPTION)

Index	Inexperienced Group (N=6).	Experienced Group (N = 6)
Experience (year)	None	25–60 (Average (Ave) 40.8, Standard Deviation (S.D.) 14.3, Average Z value (Z) -1.73* 10^{-16}, Coefficient of Variation (CV) 2.61)
Age (year)	23–26 (Ave, 24.2, S.D. 1.07, Z -1.15 * 10^{-15}, CV 4.42)	58–74 (Ave 66, S.D. 8.50, Z 0* 10^{-15}, CV 34.9)
Stature (cm)	170–180 (Ave 174.8, S.D. 3.34, Z -2.85* 10^{-15}, CV 1.91)	160–173 (Ave 162.5, S.D. 5.5, Z 0* 10^{-15}, CV 3.39)
Weight (kg)	58–28 (Ave 67.3, S.D. 7.91, Z 5.92* 10^{-16}, CV 11.7)	55–85 (Ave 69.7, S.D. 9.96, Z -1.73* 10^{-16}, CV 14.3)
Dominant hand	Right (all subjects)	Right (N=5), Left (N=1)
Past serious physical disorder	None	None (Some had acute low back pain)
Fitness habits	None or tennis or badminton (once or twice per week)	Walking

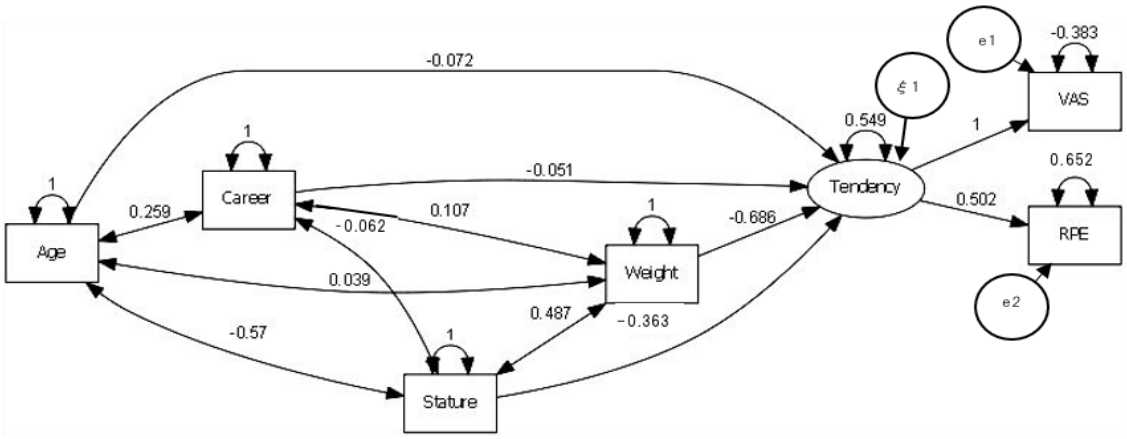

Figure 5. MIMIC model from the data concerning fundamental factors of inexperienced subjects, and their VAS and RPE scores.

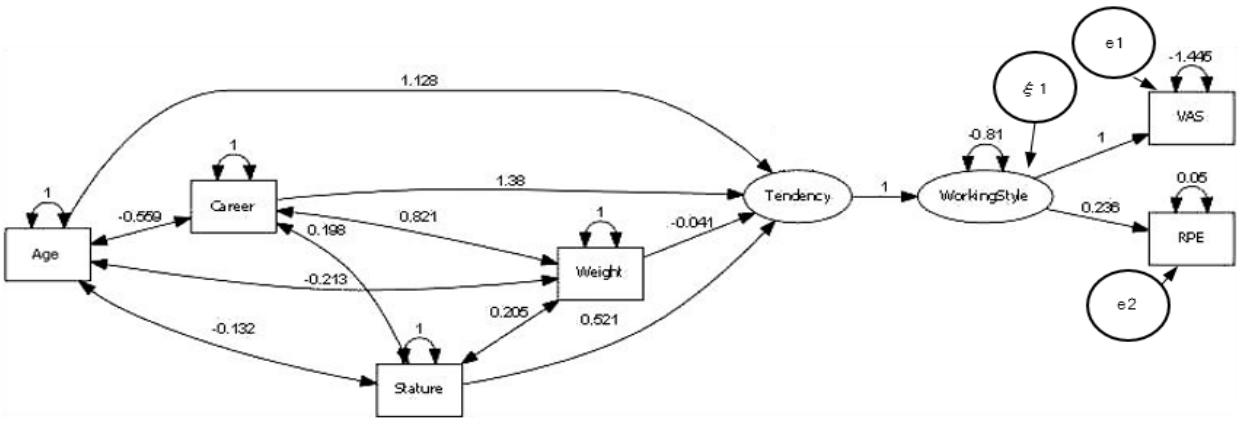

Figure 6. PLS model from the data concerning fundamental factors of inexperienced subjects, and their VAS and RPE scores.

2) VAS and RPE scores of experienced and inexperienced subjects

According to experienced subjects' model (see Fig. 7), the average scores of VAS and RPE, VAS score differed little (about 1.4 points of 100 degrees), and the average score of RPE showed an apparent difference (about two points of 15 degrees). The absolute values of four numbers of path coefficients directing to "Tendency" in the PLS model were rather small, from 0.159 to 0.394. The path coefficient from "Working Style" to "RPE" was -2.431, a large negative value. According to inexperienced subjects' model (see Fig. 8), the values of three numbers of path coefficients from "Age", "Career" and "Stature" to "Tendency" in PLS model were rather large. Among inexperienced subjects, the values of these three factors were scattered and significant, especially "Age" (1.128) and "Career" (1.380). The path coefficient from "Working Style" to "RPE" was 2.431, a large negative value.

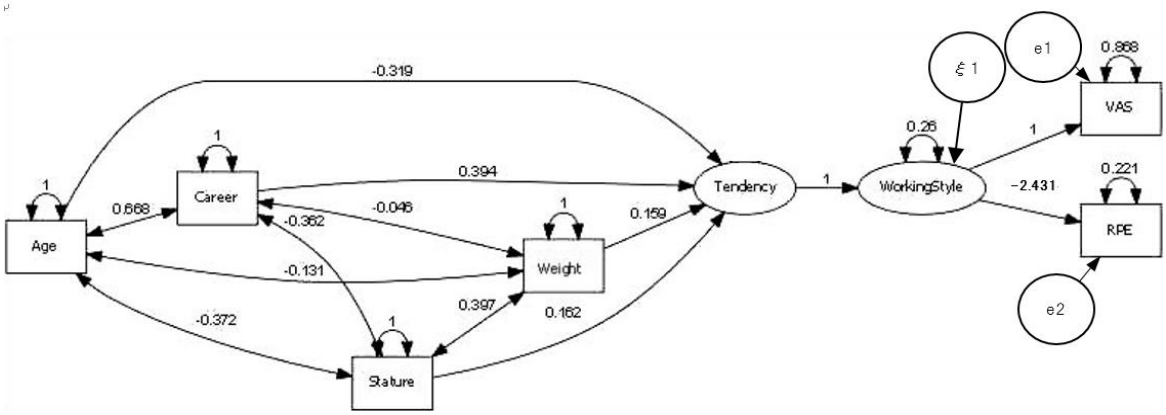

Figure 7. PLS model from the data concerning fundamental factors of experienced subjects, and their VAS and RPE scores.

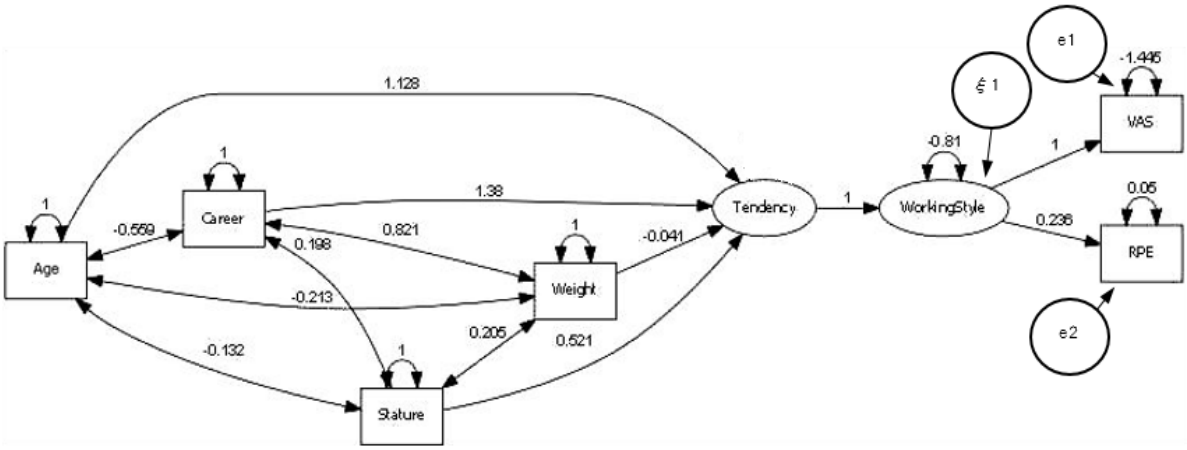

Figure 8. PLS model from the data concerning fundamental factors of inexperienced subjects, and their VAS, RPE scores.

3) Acceleration data of experienced and inexperienced subjects

According to experienced subjects' model (see Fig. 9-Fig. 10), concerning 1st, the absolute values of four numbers of path coefficients directing to "Tendency" in PLS model were rather large. On the other hand, concerning the 3rd trial, these path coefficients were small (from -0.662 to -0.248). Concerning 3rd trial, subjects' character of their movement were tend to effect to "DCAcceHoe3rd" and "SdAcceHoe3rd". In the 3rd trial, the path coefficient from "Working Style" to "DcAcceHoe3rd" was 0.998. In short, it was almost equal to the path coefficient from "Working Style" to "SdAcceHoe3rd"(1.000). According to inexperienced subjects' model (see Fig. 11 and Fig. 12), concerning the 1st and 3rd trials, the values of path coefficients from "Weight" to "Tendency" were fairly large, the path coefficients from "Working Style" both to "DcAcceHoe1st" and to "DcAcceHoe3rd" were moderate (-0.330 and 0.297).

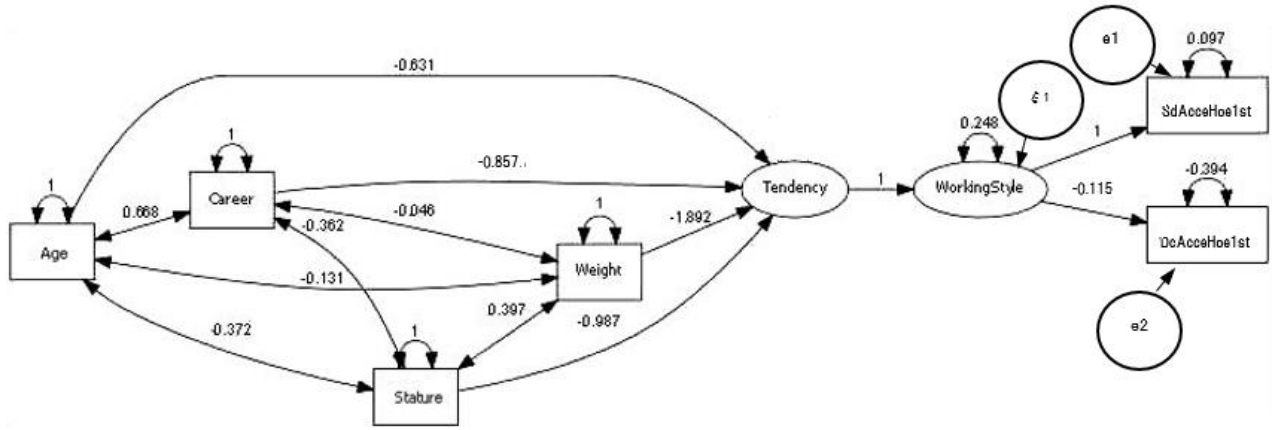

Figure 9. PLS model from the data concerning fundamental factors of experienced subjects, and their vertical acceleration data (S.D. and DC component) of their first trial.

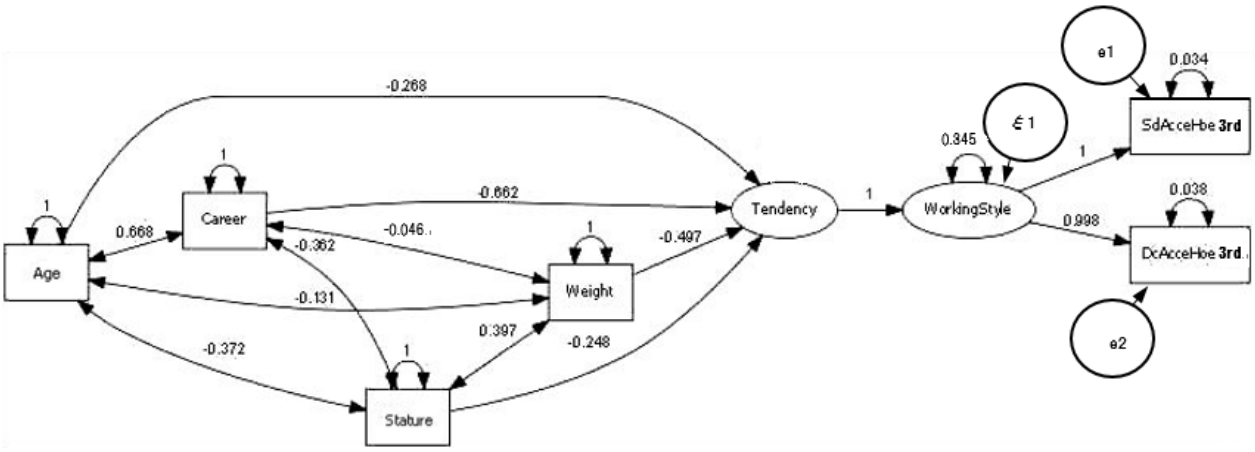

Figure 10. PLS model from the data concerning fundamental factors of experienced subjects, and their vertical acceleration data (S.D. and DC component) of their third trial.

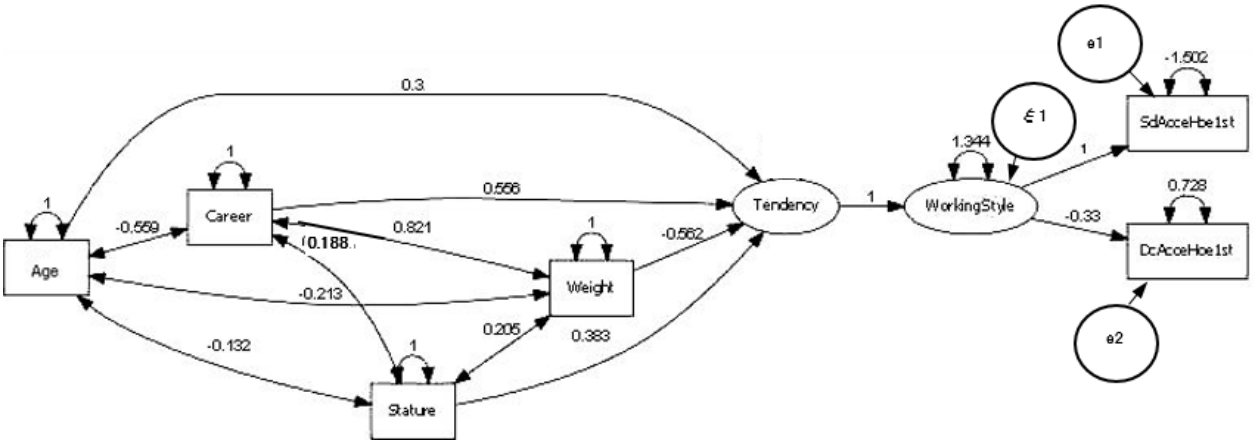

Figure 11. PLS model from the data concerning fundamental factors of inexperienced subjects, and their vertical acceleration data (S.D. and DC component) of their first trial.

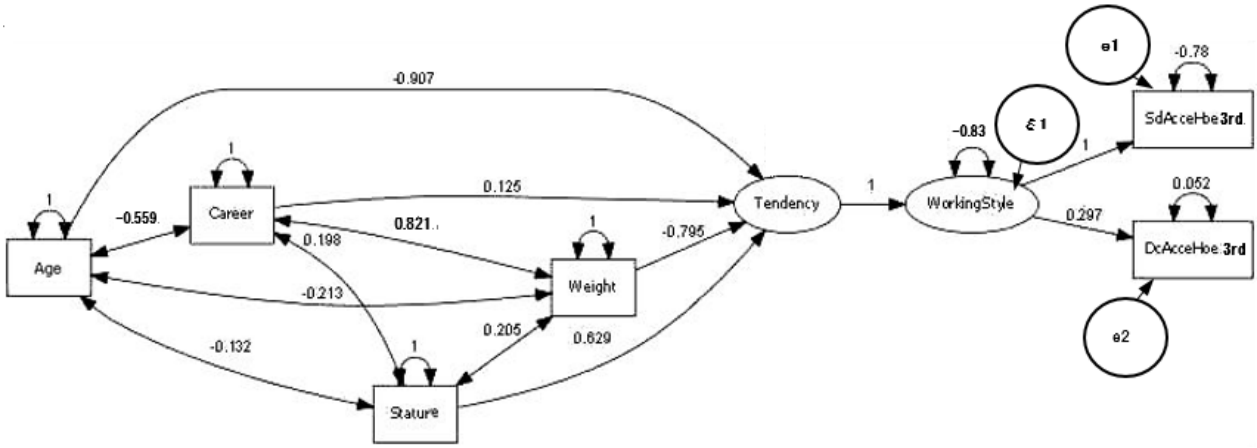

Figure 12. PLS model from the data concerning fundamental factors of inexperienced subjects, and their vertical acceleration data (S.D. and DC component) of their third trial.

B. SEM Model

The authors described the scores of GM analysis based path coefficients below.

1) VAS and RPE scores of experienced and inexperienced subjects

According to experienced subjects' model (see Fig. 13), the length of career was inversely proportional to the scores of RPE. Tall or heavy experienced subjects were likely not to register high RPE scores related to this work, and they did not feel that the work is intense.

Further, tall subjects seemed not to feel tired after work, based on VAS. According to inexperienced subjects' model (see Fig. 14), tall or heavy inexperienced subjects were also likely not to register high VAS scores related to

this work. On the other hand, heavy subjects registered high RPE scores.

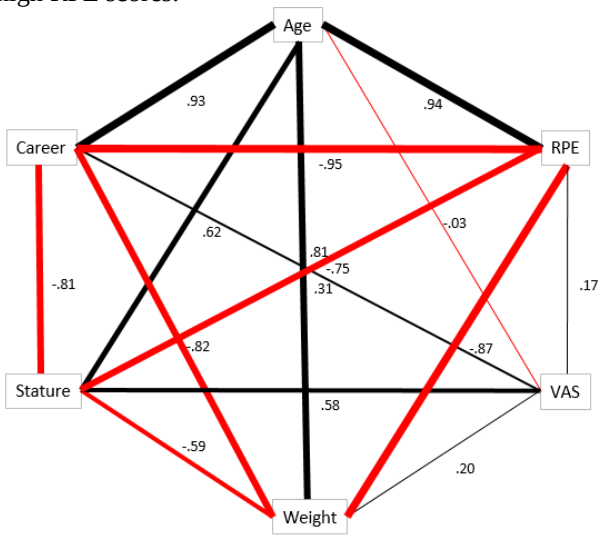

Figure 13. GM model from the data concerning fundamental factors of experienced subjects, and their VAS and RPE scores.

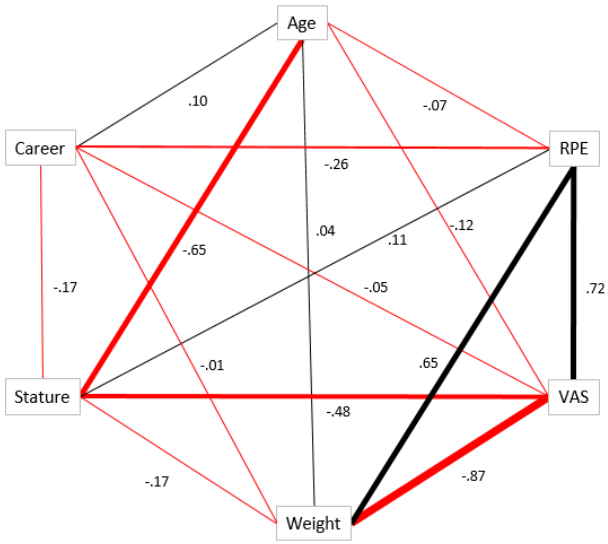

Figure 14. GM model from the data concerning fundamental factors of inexperienced subjects, and their VAS and RPE scores.

2) Acceleration data of experienced and inexperienced subjects

For experienced subjects (see Fig. 15 and Fig. 16), age, career, stature and weight were not proportional to S.D. values. In short, these four factors were likely to affect to their swing's characters slightly and to the value of DC component proportionally.

However, about their 3rd trial, these four factors did not affected to the values of S.D. and DC component strongly. For inexperienced subjects (see Fig. 17 and Fig. 18), concerning the 1st trial, the values of S.D. and DC components concerning the 1st and 3rd trials were not proportional. Heavy subjects were likely to keep a hoe pointing in a downward direction and swing with a small range. However, this trend reversed in the 3rd trial.

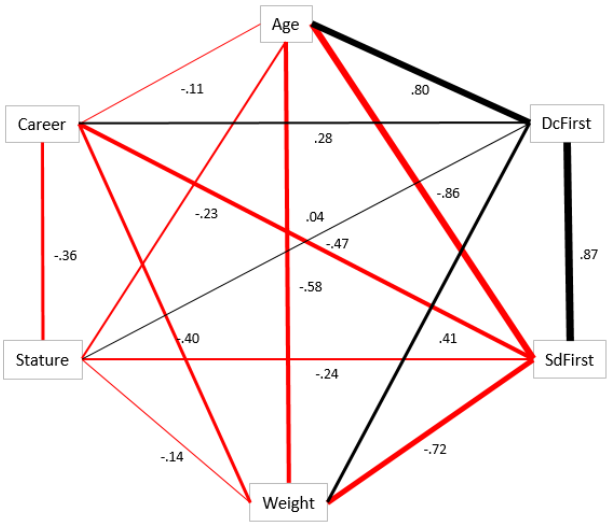

Figure 15. GM model from the data concerning fundamental factors of experienced subjects, and their vertical acceleration data (S.D. and DC component) of their first trial.

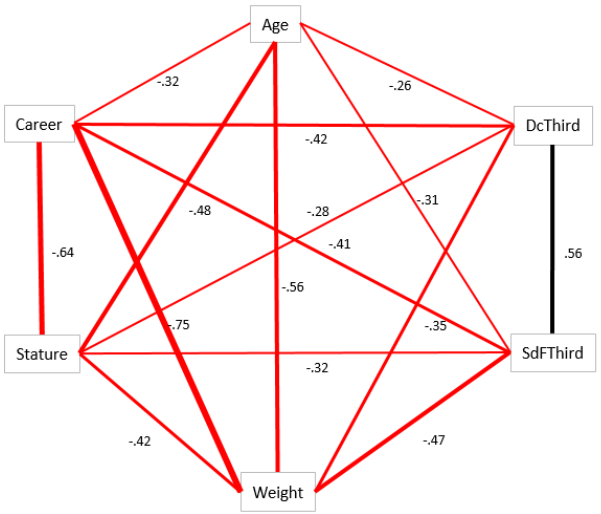

Figure 16. GM model from the data concerning fundamental factors of experienced subjects, and their vertical acceleration data (S.D. and DC component) of their third trial.

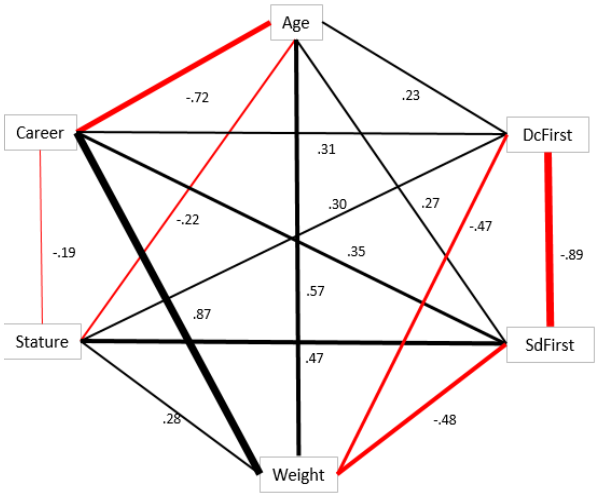

Figure 17. GM model from the data concerning fundamental factors of inexperienced subjects, and their vertical acceleration data (S.D. and DC component) of their first trial.

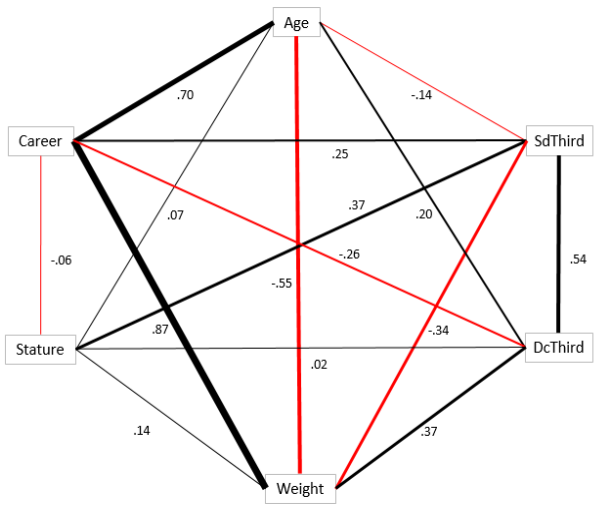

Figure 18. GM model from the data concerning fundamental factors of inexperienced subjects, and their vertical acceleration data (S.D. and DC component) of their third trial.

3) VAS, RPE scores and acceleration data of experienced and inexperienced subjects

In total, all absolute values of path coefficients were higher for experienced subjects than for inexperienced subjects (see Fig. 19 and Fig. 20). Concerning experienced subjects (see Fig. 19), there were high correlations between 1st trial and 3rd trial S.D. and DC components, these features were also confirmed in inexperienced them (see Fig. 20). Their all RPE scores did not be both proportional to 1st S.D. and proportional to 3rd S.D. (see Fig. 19 and Fig. 20). Their 1st and 3rd trial DC component was difficult to analyze. Concerning inexperienced subjects (see Figure 20), inexperienced subjects' path coefficients were rather small values, and were more dispersed than those of experienced subjects. These path coefficient's data concerning experienced subjects were more clear and characteristic, in other words, higher correlations between 1st and 3rd S.D. and DC components.

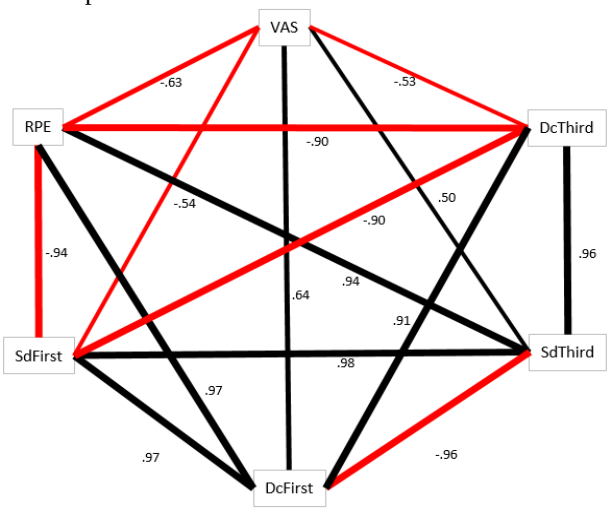

Figure 19. GM model from VAS, RPE scores of experienced subjects, and their vertical acceleration data (S.D. and DC component) of their first and third trial.

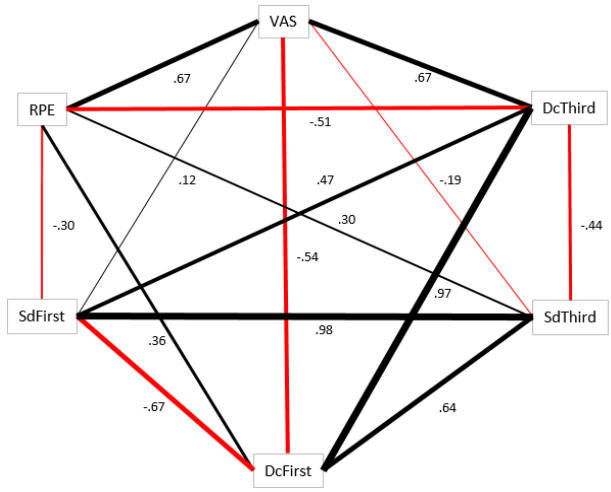

Figure 20. GM model from VAS, RPE scores of inexperienced subjects, and their vertical acceleration data (S.D. and DC component) of their first and third trial.

C. Correlation Ratio

In light of the results in Table II concerning correlation ratio (η^2: 0 - 1) between experienced and inexperienced subjects, the range of η^2 were rather scattered (0.0009 - 0.77), concerning data from survey sheets, η^2 of subjects' age, career, stature and weight were relatively high (0.56 – 0.75). On the other hand, concerning weight and VAS, the values were quite low (0.0009 - 0.03).

TABLE II. CORRELATION RATIO (η^2: 0 - 1) BETWEEN EXPERIENCED AND INEXPERIENCED SUBJECTS

Index			η^2
From survey sheets		Age	0.75
		Career	0.62
		Stature	0.56
		Weight	0.03
		VAS	0.0009
		RPE	0.56
From the time series of hoe acceleration data	1st trial	Max	0.03
		Min	0.39
		S.D.	0.01
		DC	0.61
	3rd trial	Max	0.01
		Min	0.56
		S.D.	0.01
		DC	0.77

IV. DISCUSSION

A. SEM Model

1) Contrast between MIMIC model and PLS model

Between MIMIC models and PLS models, these coefficients' values were quite different. The authors thought that this tendency was because of MIMIC models' simplicity and forcefulness. Inherently, the PLS model is more precise model and more minute than MIMIC one, in addition, for the season of the fundamental accuracy about matching between their real data and these two path models' style. The values of MIMIC models' path coefficients have lower believability than those of PLS models, which is why

obtaining these tendencies is difficult. Because of this, the authors mainly analyzed and described the aforementioned PLS models, and the values of the associated path coefficients.

2) VAS and RPE scores of experienced and inexperienced subjects

According to the score of VAS and RPE of experienced and inexperienced subjects, these results were believed to have been caused by differences in their experience and physical condition. The absolute values of four numbers of path coefficients directing to "Tendency" in the PLS model were rather small. The path coefficient from "Working Style" to "RPE" was a large negative value. In short, these experienced subjects' features effected to their fatigue negatively, inversely. The values of three numbers of path coefficients from "Age", "Career" and "Stature" to "Tendency" in PLS model were rather large. Among inexperienced subjects, the values of these three factors were scattered and significant. That was why the authors supposed that those effected to their subsequent factors “Tendency” and “Working Style” more directly than experienced them. The path coefficient from "Working Style" to "RPE" was a large negative value, in short, these experienced subjects' features effected to their fatigue negatively, inversely.

3) Acceleration data of experienced and inexperienced subjects

According to acceleration data of experienced and inexperienced subjects, concerning 1st, the absolute values of four numbers of path coefficients directing to "Tendency" in PLS model were rather large, however, concerning the 3rd trial, these path coefficients were small. The authors guessed that in the 1st trial, these four factors were likely to affect hoe acceleration directly, however concerning 3rd trial, subjects' character of their movement were tend to effect to "DCAcceHoe3rd" and "SdAcceHoe3rd". Relating to the 1st and 3rd trials, the values of path coefficients from "Weight" to "Tendency" were evenly large, so the authors thought that was because inexperienced subjects’ weight was related to their muscles and power, their swing style and character were affected by their weight.

On the other hand, concerning their "Stature", it was difficult to comment. Concerning the 1st and 3rd trials, the path coefficients from "Working Style" both to "DcAcceHoe1st" and to "DcAcceHoe3rd" were moderate. "Working Style" looked like to impact largely on Standard Deviations of vertical accelerations, not to their average swing angle.

B. GM Model

1) VAS and RPE scores of experienced and inexperienced subjectsVAS and RPE scores of experienced and inexperienced subjects

Concerning VAS and RPE scores of experienced and inexperienced subjects, tall subjects seemed not to feel tired after work, based on VAS. These data were quite reasonable in light of this movement, the weight of this hoe and our commonsense. Tall or heavy inexperienced subjects were also likely not to register high VAS scores, probably because of their physical strength. Additionally, the authors thought that the scores of VAS and RPE were also are reasonable. On the other hand, heavy subjects registered high RPE scores, seemingly because of the shortage of subject data (N = 6).

2) Acceleration data of experienced and inexperienced subjects

Concerning experienced subjects, age, career, stature and weight were not proportional to S.D. values, so the authors thought that these four factors were likely to affect their swing's characters slightly. Concerning their 3rd trial, these four factors did not affected to the values of S.D. and DC component strongly, it would not be reasonable. Possibly, those were merely because of their fatigue. For the values of S.D. and DC components concerning the 1st and 3rd trials were not proportional. These results meant that holding a hoe horizontally is connected to the smallness of their swing. Heavy subjects were likely to keep a hoe pointing in a downward direction and swing with a small range. The authors could observe these characteristics from the video data.

3) VAS, RPE scores and acceleration data of experienced and inexperienced subjects

The tendency that all absolute values of path coefficients were higher for experienced subjects than for inexperienced subjects was because inexperienced subjects’ physical and motion features were more variable than experienced subjects. There were high correlations between 1st trial and 3rd trial S.D. and DC components, the authors also confirmed these features in inexperienced them. Their RPE scores did not be both proportional to 1st S.D. and proportional to 3rd S.D., so the authors guessed that was why subjects concentrated extremely hard on the 1st trial and were stiff, so in 3rd trial, they were exhausted; their motions became more discursive later. The authors cannot comment on their 1st and 3rd trial DC component, these were troublesome to be commented. The authors also could see that experienced subjects’ motions were more typical, rigid, and not so flexible from visual data.

4) Total discussion concerning path analyses of MIMIC and PLS models, and hexagonal shaped GM analyses

Through path analyses of both MIMIC and PLS models, and hexagonal shaped GM analyses, we tried to indicate superficial, sequential cause-and-effect links, but also to reveal hidden relationships relating to the typical agricultural tillage. However, the authors had difficulties to describe about those differences explicitly, because those index values had their own complexity and effected multiply. That is why agricultural leaders, managers and newcomers should consult and refer these achievements concerning the tendencies of tilling movements on a case-by-case basis, like a common coaching and training in a sports-related field.

C. Correlation Ratio

Considering the raw numerical data of correlation ratio, the weight and VAS data were too distributed among experienced and inexperienced groups, and those results

of η^2 seemed reasonable. Especially their age was strongly concerned with other factors, on the other hand, their weight and VAS were not so. Related to the hoe vertical acceleration, DC components concerning first and third trials were thought as key factors, secondly, these minimum values were significant. In short, with what angle should workers have a hoe was the critical points.

That is why inexperienced workers should carefully concentrate on and pay attention to aforementioned DC components and minimum acceleration values when they are training about the working skill and involved in such tasks. Additionally, strictly speaking, the authors should include other factors like weather, humidity and soil condition etc., however, in the trials, these features were decided as enough typical, not specific in Japanese farmlands. In addition, the authors wanted to focus on the analysis based on human dynamics and physiological sciences, they did not include them.

V. Conclusion

Our studies have developed and tested various basic steps of the authors' original systems and methods by examining actual agricultural working sites. In this research, the authors have obtained promising prospects concerning the validity of SEM and GM analysis these are relating to other path analyses in diverse scientific fields, furthermore, have also obtained correlation ratio data. Considering them, the authors have demonstrated significant, practical suggestions for agricultural leaders, managers and newcomers. In short, the authors believed that our results were achievements in the fusion of agricultural informatics, statistics, and human dynamics to some extent.

However, authors' systems' long-term endurance, accuracy and coverable area must be confirmed more accurately later. For that, the authors should add a greater variety of worker data into these statistical data, and execute both more varieties of and a greater amount of path analysis.

The authors also thought that the measure of precision for diagnosing critical situations and bad conditions (e.g. serious disease) would be improved in future. But also, the stability of the whole system against sudden accidents would be improved, too. The reason was that irregular movements and posture data could certainly be diagnosed. Furthermore, utilizing these results and many sets of experimental evidence, the authors could plan to launch practical supporting projects for workers. From the viewpoint of global agricultural dynamics, the authors could use these methods with arranging variables to match the requirements of foreign situations and agricultural methods.

Acknowledgement

My heartfelt appreciation goes to the members of CSIS and IIS at The University of Tokyo in Japan, who have provided careful supports, feedbacks and valuable comments.

References

[1] Y. Fujii, T. Nanseki, H. Kobayashi, and T. Kojima, "The characteristics of expert know-how in agricultural planning on large-scale paddy field farms: A case study of a corporate farm in Shiga prefecture," *Agricultural Information Research*, vol. 22, no. 3, pp. 142-158, April 2013.

[2] A. Shinjo and M. Kudo, "The practical use of IT in agriculture: The movement into high-value-added crops and integrated solutions," *The Journal of the Institute of Electronics, Information, and Communication Engineers*, vol. 96, no. 4, pp. 280-285, April 2013.

[3] Y. Fujii, T. Nanseki, H. Kobayashi, and K. Nishitani, "Information management and cultivation of employee capabilities at large-scale paddy field farms: A case study of raising rice seedlings," *Agricultural Information Research*, vol. 21, no. 5, pp. 51-64, June 2012.

[4] T. Nanseki, S. Takeuchi, and Y. Shinozaki, "Business development, ICT use, and personal training in agricultural corporations: An analysis of nationwide questionnaire survey," *Agricultural Information Research*, vol. 22, no. 3, pp. 159-173, May 2013.

[5] S. Kawakura and R. Shibasaki, "Supporting systems for agricultural worker's skill and security," in *Proc. Asian Conference Remote Sensing*, South Kuta, Bali, Indonesia, 2013, pp. 71-77.

[6] S. Kawakura and R. Shibasaki, "Timeline effects of vocal instructions from computer programs on agricultural technical teaching," *Journal of Advanced Agricultural Technologies*, vol. 1, no. 2, pp. 104-112, December 2014.

[7] D. Karantonis, M. Narayanan, M. Mathie, N. Lovell, and B. Celler, "Implementation of a real-time human movement classifier using a triaxial accelerometer for ambulatory monitoring," *IEEE Transactions on Information Technology in Biomedicine*, vol. 10, no. 1, pp. 156–167, January 2006.

[8] P. Mistry, P. Maes, and L. Chang, "WUW-wear Ur world: A wearable gestural interface," in *Proc. CHI '09 Extended Abstracts on Human Factors in Computing Systems*, vol. 2009, pp. 4111-4116, April 2009.

[9] S. C. Mukhopadhyay, "Wearable sensors for human activity monitoring: A Review," *IEEE Sensors Journal*, vol. 15, no. 3, pp. 1321-1330, March 2015.

[10] P. P. Pradhan and S. P. Dhanure, "Child activity monitoring using tri-axial accelerometer," *International Journal of Advanced Research*, vol. 3, no. 6, pp. 222-224, June 2015.

[11] D. E. Pritchard, Y. J. Lee, and L. Bao, "Mathematical learning models that depend on prior knowledge and instructional strategies," *Physical Review Special Topics Physics Education Research*, vol. 4, no. 1, pp. 010109-1 - 010109-8, May 2008.

[12] S. Qiu, Z. Wang, H. Zhao, and H. Hu, "Using distributed wearable sensors to measure and evaluate human lower limb motions," *IEEE Transactions on Instrumentation and Measurement*, vol. 65, no. 4, pp. 939-950, April 2016.

[13] N. Raveendranathan, *et al.*, "From modeling to implementation of virtual sensors in body sensor networks," *IEEE Sensors Journal*, vol. 12, no. 3, pp. 583–593, March 2012.

[14] P. Augustyniak, M. Smoleń, Z. Mikrut, and E. Kańtoch, "Seamless tracing of human behavior using complementary," *Sensors*, vol. 14, no. 5, pp. 7831-7856, May 2014.

[15] L. Bao and S. I. Stephen, "Activity recognition from user annotated acceleration data," in *Proc. 2nd Int. Conference on Pervasive Computing*, 2004, pp. 1–17.

[16] T. Terada, Y. Kobayashi, and M. Tsukamoto, "A context-aware system based on scent," *Computer Software*, vol. 29, no. 4, pp. 324-334, February 2012.

[17] N. Funao, *The R Tips—Data Kaiseki kankyou R no Kihonwaza・Graphics Katsuyou Syuu*, Tokyo, Japan: Ohmsha, 2009.

[18] T. Kan, *Excel de Manabu Tahenryou Kaiseki Nyuumon*, Tokyo, Japan: Ohmsha, 2013.

[19] T. Kan, *Excel de Manabu Toukei Kaiseki Nyuumon*, Tokyo, Japan: Ohmsha, 2013.

[20] T. Kan, *Excel de Manabu Toukeiteki Yosoku*, Tokyo, Japan: Ohmsha, 2014.

[21] T. Kojima and M. Yamamoto, *Excel de Manabu Kyoubunsan Kouzou to Graphical Modeling*, Tokyo, Japan: Ohmsha, 2013.

[22] N. Matloff, *Art of R Programming*, Tokyo, Japan: O'Reilly Japan, 2012.
[23] T. Oomori, M. Handa, and H. Yadohisa, *R Commander Niyoru Data Kaiseki*, Tokyo, Japan: Kyouritsu Syuppan, 2014.
[24] P. Teetor, *R Cook Book*, Tokyo, Japan: O'Reilly Japan, 2011.
[25] N. Toyama and M. Tsujitani, *Jissen R Toukei Bunseki*, Tokyo, Japan: Ohmsha, 2015.
[26] T. Yamada and T. Sugisawa, *R Niyoru Yasashii Toukeigaku*, Tokyo, Japan: Ohmsha, 2008.
[27] G. W. Juette and L. E. Zeffanella, "Radio noise currents in short sections on bundle conductors," presented at the IEEE Summer Power Meeting, Dallas, TX, June 1990, pp. 22-27.
[28] L. Bao, "Physical activity recognition from acceleration data under semi-naturalistic conditions," M. Eng. thesis, Massachusetts Institute of Technology, 2003.

Shinji Kawakura received Ph.D. in Environmentology from The University of Tokyo, 2015. (Meguro-ku, Tokyo, Japan). He was born on July 14, 1978. He received his B.A in control system engineering, Tokyo Institute of Technology, 2003, and M.A. in human-factors engineering, Tokyo Institute of Technology, 2005. (Meguro-ku, Tokyo, Japan). His past career: system engineering, researching for private companies, and Development and Verification of Wearable Sensing Systems with Real-time Spoken Commands for Outdoor Agricultural Workers.

Ryosuke Shibasaki is with Department of Socio-Cultural and Socio-Physical Environmental Studies, The University of Tokyo/Kashiwa-shi, Chiba, Japan. Dr. in Engineering.
He is Professor at the Center for Spatial Information Science, University of Tokyo.