

# **The Use of Statistical, Computational and Modelling Tools in Higher Learning Institutions: A Case Study of the University of Dodoma**

Gilbert M. Gilbert  
College of Informatics and Virtua Education,  
The University of Dodoma

## **ABSTRACT**

In this paper an assessment was made to the use and awareness of statistical, modelling and computational tools and methods in higher learning institutions, the University of Dodoma taken as a case study in which 112 instructors were randomly sampled. Data were analyzed using Chi-square tests with p-values to determine the statistical significances in the use of tools. Results show that there is no significant association between instructors of the colleges of the university in the use and awareness of popular statistical tools, however the study reveals very high statistical significant relationships in the awareness of computational and modelling tools and concepts between the instructors. Moreover, a considerable gap was observed in terms of training of the tools and concepts in which a good number of instructors, 75%, never had any formal training and 94.7% have shown the need for training of the tools and concepts. It is an opportunity for relevant bodies to provide effective training of the tools to ensure quality research and better education for the benefit of all stakeholders.

## **General Terms**

Statistical, Computational, Modelling, Instructors, Higher learning, university.

## **Keywords**

Statistical tools, Computational and modelling tools, higher learning institution.

## **1. INTRODUCTION**

In higher learning institutions like universities and colleges, one of the core attributes of effective and competent instructors is to continually involve themselves in quality research of various nature where they search, create or use of existing knowledge creatively for the purpose of furthering new ideas, concepts, methodology and understanding (Hansen 2009).

At the heart of many research activities lies statistical methods which play significant and vital role in analyzing data to help in assessing current situation, objectives setting and reach conclusion more reliably (Hsu 2005). These methods are crucial in ascertaining proposed theories to minimize knowledge gaps in our societies where suitable policies could be made out of these informed-decision. It is therefore uttermost important for researchers in higher learning institutions to possess widening skills in statistics and the interpretation of their analyses.

Over the years, computers have had profound effects on the practical application of statistics in teaching, research and consultancy. The use of computers has enabled development of complex statistical software tools which integrate different

statistical techniques which in turn have allowed researchers to be able to apply different research methods and analysis with little difficult (Duckworth and Stephenson 2002). However, computer have revolutionized statistics in many aspects with the introduction of new, progressive, extendible statistical or mathematical packages (e.g. Splus, Gauss, Matlab, and Mathematica) and special-purpose software (Diaconis and Efron, Duckworth and Stephenson 2002). Also, with increasingly powerful computers, robust networking platforms and reliable software, varieties of computational and modelling tools are being developed to support many aspects of research works in all disciplines. With the use of these tools, researchers are widely exposed to various research techniques, greater flexibility, and time saving when tackling complex situations in order to accelerate discovery (NIBIB, 2013). Modelling concepts such as Markov Chains (Meyn and Tweedie 1993), Queuing Theory (Sundarapandian, 2009) and Fuzzy logic (Zadeh et al, 1996), form the theoretical base of computational and modelling tools.

Consequently, this paper examines the use of statistical and computational tools and explores their extent of application among instructors in higher learning institutions with the University of Dodoma taken as a case study. The university is located in Dodoma region in Tanzania and is the largest in the country, with six colleges, namely, College of Informatics and Virtual Education (CIVE), College of Earth Sciences (CES), College of Humanities and Social Sciences (CHSS), College of Education (CE), College of Natural and Mathematics (CNMS) and College of Health Sciences (CHS).

This paper is organized as follows: section 2 describes the proposed methodology used in the study and section 3 presents findings and discussion of the results. Section 4 includes conclusion remarks and future works.

## **1.1 Research Questions**

The objective of this study was to assess the use of statistical, modelling and computational tools in higher learning institutions, a case study of the University of Dodoma, in which the focus was on differences in the use between instructors from different five colleges. The following were the guiding research questions: -

1. To what extent instructors in higher learning institutions differ in the use of statistical, computational and modeling tools with respect to the college they teach?
  - a. How are computational and modeling tools used in education, engineering, computing, social and law fields?
  - b. How do instructors involve the use of statistical, computational and modeling tools in teaching, consultancy or studies?

2. What are the status instructors in awareness and attendance of training about statistical/computational and modelling tools with respect to the college belong?

## 1.2 Research objectives

In order to answer the above questions, the researcher designed the following objectives:

1. To find out the most commonly used statistical, modelling and computational tools and their extent of use instructors in higher learning institutions
2. To establish the level of awareness and training attendance about statistical, modelling and computational tools between instructors in higher learning institutions

## 2. METHODOLOGY

In addressing the research question, a case study approach was used in order to achieve an in-depth description (Dooley 2002; Marczyk et al. 2005) in the use of computerized research tools by male and female instructors. The case study design is also useful in integrating a variety of sources, research methods, and data in the investigation (Robson, 2002; Yin, 2003).

### 2.1 Study Design

A cross-sectional descriptive design was used in the study using both qualitative and quantitative approaches for data collection, analyses and reporting. This design was chosen because it is relatively quick and easy to conduct (no long periods of follow-up), data on all variables is only collected once, multiple outcomes and exposures can be studied. This design is also good for descriptive analyses and for generating hypotheses.

### 2.2 Sampling and Data Collection

Population in this study included academic staff from the University of Dodoma which have six colleges. All the colleges' staff were picked randomly during the study.

From each of the colleges, data was collected from 112 instructors and postgraduate (PG) students of the university by filling in survey questionnaires. The questions in the questionnaires focused on identifying teachers' education level, age, role (staff, student or both) and sex. Furthermore, questions also captured the use of statistical, computational and modelling tools (such as SPSS, STATA, R, Matlab and Maya), awareness of the use of the tools and awareness of concepts in modelling (machine learning, fuzzy logic, etc.). Respondents were also asked their attendance status and willingness to attend short courses or training in the use of the tools.

Table 1 provides the summary of the characteristics of the respondents involved in the research. The study involves majority (above 60%) of instructors in age group of 26-35. Furthermore, the profile of respondents show the majority of the participants have second university degree (Master degree).

**Table 1 Demographic Characteristics of the respondents**

Variables	Sub-division of variables	Frequency				
		CIVE	CHSS	CE	CES	CNMS
Gender	Male	21	1	16	11	17
	Female	2	0	5	8	3
Age	18-25	2	0	0	0	1

(Years)	26-35	19	0	7	16	15
	36-45	1	0	13	3	4
	46 and above	1	1	1	0	0
Education Level	Bachelor Degree	9	0	4	8	9
	Master Degree	12	1	13	11	8
	PhD	2	0	4	0	3
Role at the university	PG student	3	0	0	0	7
	Staff	17	1	18	18	14
	Both	3	0	3	1	0

### 2.3 Study Variables

During the analysis, the outcome variable was taken to be the college and the independent variables (covariates) were the use of statistical, computational and modeling tools and awareness of the tools and attendance in training.

### 2.4 Methods of Data Analysis

The data collected during the study were analyzed with a Pearson Chi-square tests using a statistical package IBM SPSS Version 22. This method compares the binary outcome and other independent variables.

## 3. RESULTS AND DISCUSSION

This section explains the results obtained during the analysis of the data. Chi-Square test results together with statistical p-values were used in order to answer research questions.

### 3.1 The Extent Of Use Of Statistical, Modelling And Computation Tools Among Instructors In Respective Colleges

From Table 2, majority of the respondents (71.4%) reported the use of statistical, computational and modeling tools in their teaching, consultancy or other academic endeavors with no any statistical significance between instructors of the colleges ( $p\text{-value}=.961, >.05$ ) and only 15.2% involving those tools in more than 5 research works. Out of all the fields, education sector attracted a good number of instructors (43.1%) who work with the tools.

A very high statistical significance could be observed in research fields, which could be explained by the fact that most instructors participate only in research work pertaining to their field of expertise. However, more than 74% of the participants agreed the usefulness of the tools in adding values to their research data. With no any significant relationship by being an instructor of a particular college ( $p\text{-values} > 0.05$ ).

#### 3.1.1 Awareness of the use of statistical tools

Based on the analyses of the data as shown in Table 3 and Figure 4, it could be revealed that the level of awareness in the use of statistical tools, particularly SPSS was at basic level to majority of instructors (71.4%) with no significance in the use of the package between instructors of difference college ( $p\text{-value}>0.05$ ). Almost 50% of the instructors who participated in the research were reported to never used or heard about STATA and R statistical tools. A significant relationship ( $p\text{-value} = 0.035$ ) was observed in awareness of R statistical tool.

### 3.1.2 Awareness of the use of modelling and computational tools

Figure 3 and Table 4 show the statistics of the level of the awareness of computation and modeling tools. It can be seen that majority of the respondents have never used or heard about the tools (Simulink (81%), 3Ds Max (90%), Blender

(94%), Maya (96%), Cinema 4D (94%), Unity 3D (94%) and Python (90%)). It could be that many instructors have neither been exposed to or attend training regarding the use of computational and modeling tools. Further results on the distribution of the use of computational and modeling tools are shown in Figure 1.

**Table 2 Descriptive statistics of analysis of the covariates with respect to colleges**

Covariate	Response	Colleges (n, (%))					P-value
		CIVE	CHSS	CE	CES	CNMS	
Have you ever involved the use of statistical/computing/modelling tools in your teachings/consultancy/studies?	Yes	15 (25)	1 (1.7)	14 (23.3)	15 (25)	15 (25)	<b>.961</b>
	No	7 (21.1)	0 (0)	3 (15.8)	5 (26.3)	4 (15.8)	
	Not sure	1 (20)	0 (0)	2 (40)	1(20)	1 (20)	
In how many research/study/consultancy works have you involved the use of the tools?	1-2	10 (25.6)	1 (2.6)	11 (28.2)	8 (27.3)	9 (23.1)	<b>.961</b>
	3-5	3 (27.3)	0 (0)	2 (18.2)	3 (33.3)	3 (27.3)	
	More than 5	1 (11.1)	0 (0)	2 (22.2)	3 (33.3)	3 (33.3)	
In which fields have you used these tools?	Education	2 (8)	0 (0)	13 (52)	2 (8)	8 (32.2)	<b>0.00**</b>
	Engineering	5 (55.6)	0 (0)	0 (0)	2 (22.2)	2 (22.2)	
	Computing	7 (43.8)	0 (0)	0 (0)	8 (50)	1 (6.3)	
	Law	0 (0)	0 (0)	0 (0)	0 (0)	0 (0)	
	Social	0 (0)	1 (0)	2 (66.7)	0 (0)	0 (0)	
	Medicine	0 (0)	0 (0)	0 (0)	2 (100)	0 (0)	
Science	0 (0)	0 (0)	0 (0)	0 (0)	3 (100)		
Do you think these tools have/ had any value addition in the analysis of your research data?	Yes, very much	11 (22.9)	1 (2.1)	11 (22.9)	12 (25)	13 (27.1)	<b>.831</b>
	Yes, to some extent	4 (30.8)	0 (0)	4 (30.8)	3 (23.1)	2 (15.4)	
	No, they do not add anything	0(0)	0 (0)	0 (0)	0 (0)	1 (100)	

\*\* means very high significance.

**Table 3 Descriptive statistics of the awareness and use of statistical tools by academicians with respect to their colleges of engagement**

Statistical Tool/Covariate	Response (Level of use or awareness)	Colleges (n, (%))					P-values
		CE	CHSS	CIVE	CNMS	CES	
SPSS	expert	1 (14.3)	0 (0)	1 (14.3)	2 (28.6)	3 (42.9)	<b>0.587</b>
	know the basics	17 (34)	1 (2)	10 (20)	10 (20)	12 (24)	
	never used	2 (18.2)	0 (0)	6 (54.5)	2 (18.2)	1 (9.1)	
	never heard	1 (50)	0 (0)	1 (50)	0 (0)	0 (0)	
STATA	expert	0 (0)	0 (0)	0 (0)	1 (50)	1 (50)	<b>0.052</b>
	know the basics	0 (0)	0 (0)	0 (0)	2 (40)	3 (60)	
	never used	5 (20.8)	1 (4.2)	8 (33.3)	2 (8.3)	8 (33.3)	
	never heard	11 (61.1)	0 (0)	3 (16.7)	2 (11.1)	2 (11.1)	
R	expert	1 (33.3)	0 (0)	0 (0)	2 (66.7)	0 (0)	<b>0.035</b>
	know the basics	0 (0)	0 (0)	0 (0)	2 (40)	3 (60)	
	never used	3 (15)	1 (5)	8 (40)	3 (15)	5 (25)	
	never heard	13 (52)	0 (0)	4 (16)	3 (12)	5 (20)	

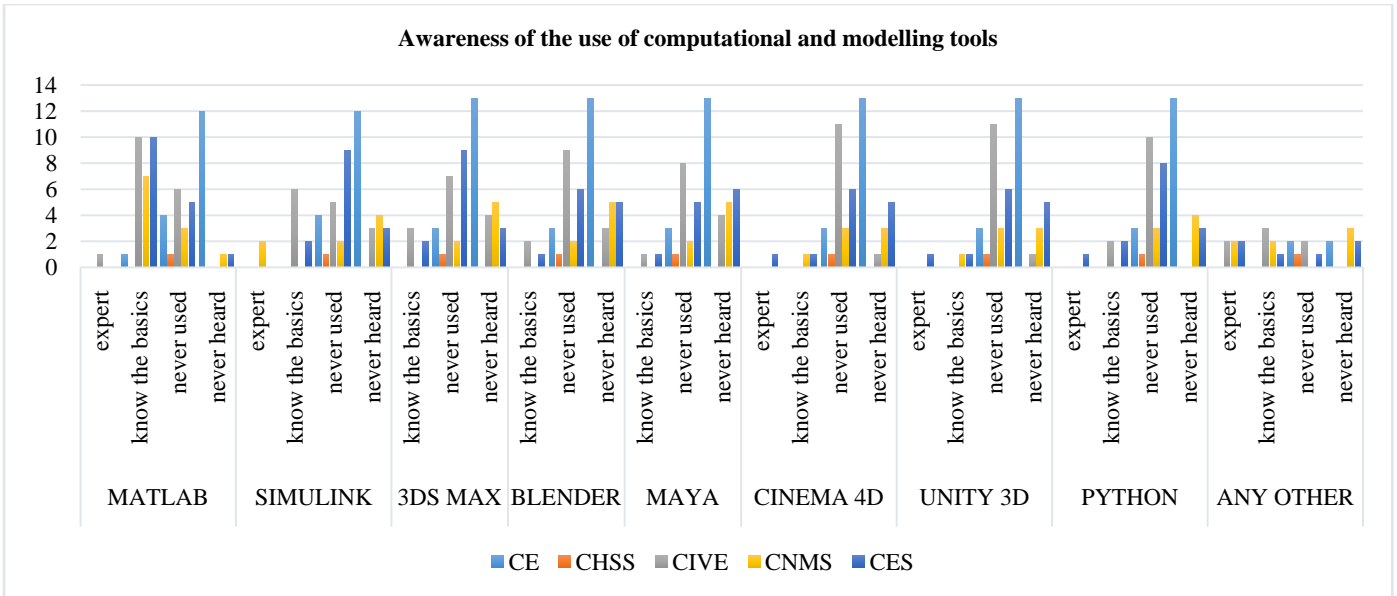


Fig 1: Awareness of computational and modelling tools between instructors with respect to the college they belong

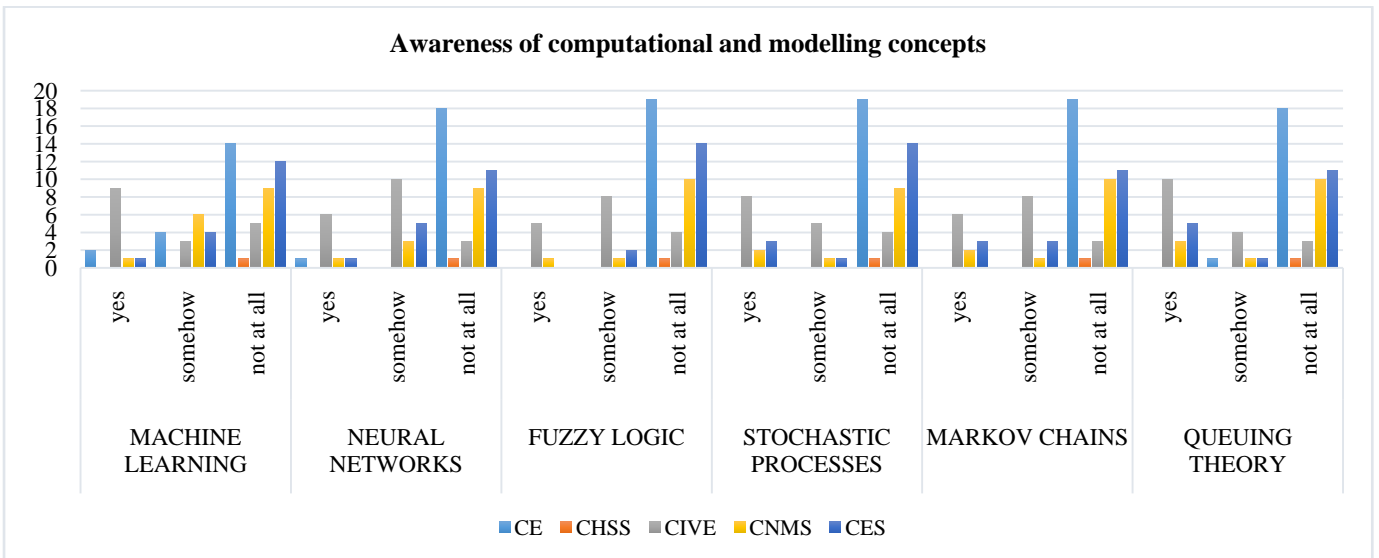


Fig 2: Awareness of computational and modelling concepts between instructors with respect to the college they belong

Table 4 P-values of the awareness of computational and modelling tools among instructors in respective colleges

Computational and modelling Tools	P-value
Matlab	0.00**
Simulink	0.001**
3Ds Max	0.023
Blender	0.068
Maya	0.223
Cinema 4D	0.027
Unity 3D	0.027
Python	0.009
Other	0.375

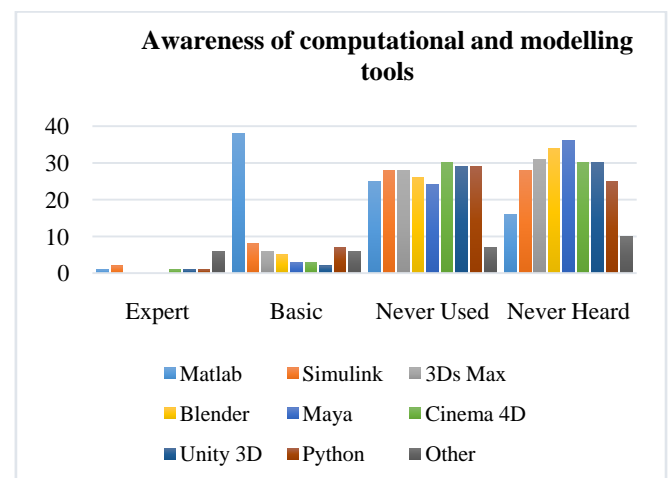


Fig 3. Awareness in the use of computational and modeling tools

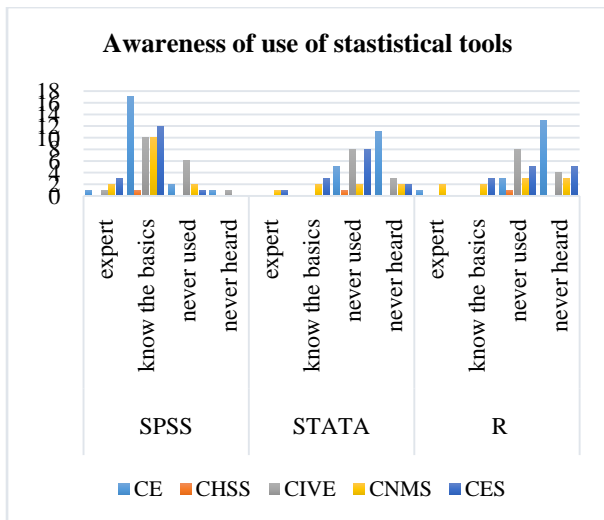


Fig 4. Awareness in the use of tools between instructors

Using Chi-square tests, the researcher could establish the significant relationship between the use of the tools and belonging to a particular college as shown in Table 4. Those who reported to use the tools have background in computer studies and they belong to CIVE college which teaches computer courses. Figure 1 elaborates the same.

### 3.1.3 Awareness of the computational and modeling concepts

In this section, analyses of the awareness of computational and modeling concepts are presented.

Using Figure 2, one could observe the pattern of the responses with in which the gap in the awareness of computational and modeling concepts between those who responded to know and not at all with the majority on the far right side. Table 5 again shows that there are very high statistical significant relationships between instructors of the colleges when it comes to the awareness of computational and modeling tools with those responding yes or somehow engaged to science colleges (CIVE, CES, CNMS).

Table 5 P-values of the awareness of the computation and modeling concepts between instructors with respect to their colleges

Concept	P-value
Machine learning	0.009
Neural networks	0.001
Fuzzy logic	0.00*
Stochastic processes	0.001
Markov chain	0.00*
Queuing theory	0.002

### 3.1.4 Training attendance

#### 3.1.4.1 Status of training attendance

It can be seen clearly in Figure 5 that more than three quarters of the responded to not have attended any sort of training regarding the use of statistical, computational and modeling tools. This agrees with what which have been observed in the above paragraphs. A p-value of 0.015(<0.05) explains the differences in the attendance among instructors.

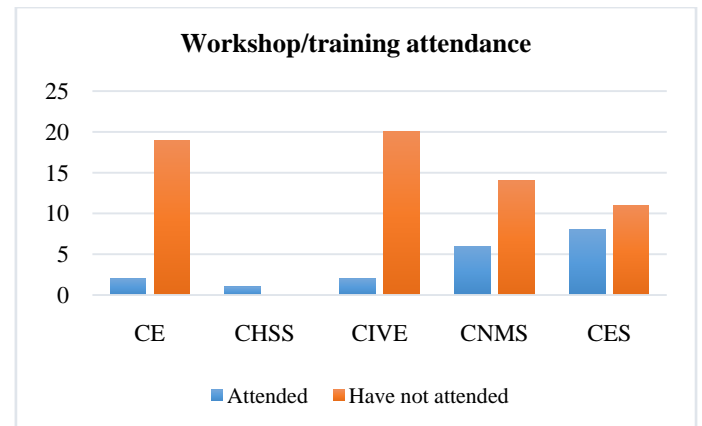


Fig 5: Training attendance status between and female instructors

#### 3.1.4.2 Willingness to attend training about statistical, computation and modelling tools

According to statistics as shown in Figure 6, majority of instructors seem to be interested in using the tools and are willing to attend training about the use of the tools. About 94.7% of the responded are willing to attend if these training are made available to them. The willingness is consistent throughout colleges as confirmed by the results of the analyses using Chi-square test in which a p-value of 0.53 (> 0.05).

Table 6 P-values of willingness to attend training among instructors with respect to the colleges they belong

Covariate	P-values
Attendance	0.015
Willingness	0.53

Table 6 shows the result of the Chi-square between instructors, which indicates that there is no significant association in attending trainings. In all cases p-value > 0.005.

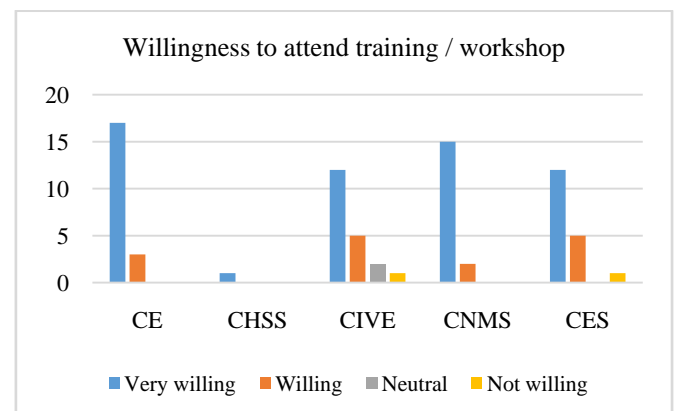


Fig 6: Willingness to attend training

## 4. CONCLUSION

This study shows the use of statistical, computational tools in assisting researchers fulfill their mission. Some tools have been popular among researchers which implicate sameness of research areas and results, particularly in education and social fields. Research works in science field have been very rare, which could be attributed to the fact that most of the tools

used in this field are not known to the researchers. In any higher learning institution, research forms the backbone of the quality of their output in terms of useful results that can be absorbed in the society. As a matter of fact, most instructors have shown extremely willingness to acquaint themselves with the tools and concepts.

As a primary goal of enhancing research culture, higher learning institutions should take a significant part to commit to the mission of building competent researchers from basic to advanced statistical techniques as well as develop their creativity and improve them as critical thinkers. Since instructors are the main researchers, these institutions should continue to equip them with necessary statistical and methodological skills in order to develop them as competitive professionals in the society.

Despite the low number of female instructors in higher learning institutions and falling behind in most aspect, still there are strong indications that their use of these tools is very much comparable to their male counterparts. It is an opportunity for relevant authority to address the issue of training equally to all parts.

## **5. ACKNOWLEDGMENTS**

Special thanks to the Carnegie-SIG Regional Initiative in Science and Education (RISE) through its competitive fund award for supporting this research. Furthermore, the author acknowledges the support on technical and academic issues regarding this publication from Computation and Modelling Research Group at University of Dodoma. Moreover, this study wouldn't have been possible without a generous support from academic and non-academic staff at the University of Dodoma.

## **6. REFERENCES**

- [1] Robson, C.(eds.) 2002. Real World Research (2nd ed.). Oxford: Blackwell Publishing.
- [2] Yin, R.K. (eds.) 2003. Case study research: Design and methods (3rd Ed.). Thousand Oaks:Sage Publications.
- [3] Dooley, L. M. 2002, 'Case study research and theory building', *Advances in Developing Human Resources* vol. 4, no. 3, pp. 335-354.
- [4] Marczyk, G., et al. 2005, *Essentials of research design and methodology*, Wiley & Sons, New Jersey.
- [5] Hansen, H .F. (2009). *Research evaluation: methods, practice, and experience*. Danish Agency for Science, Technology and Innovation. Bredgade.
- [6] Hsu, T. (2005). Research methods and data analysis procedures used by educational researchers. *International Journal of Research & Method in Education*, 28(2), 109-133.
- [7] Duckworth, W. M., and Stephenson, W. R. (2002), "Beyond traditional statistical methods," *The American Statistician*, 56(3): 230-233.
- [8] Diaconis, P., & Efron, B. (1983). Computer-intensive methods in statistics. *Scientific American*, 248(5), 116–130. doi:10.1038/scientific American 0583-116
- [9] NIBIB. (2013). Computational Modeling. Accessed 20/4/2016 at <https://www.nibib.nih.gov/science-education/science-topics/computational-modeling>
- [10] Meyn, S and Tweedie,R. (1993). *Markov Chains and Stochastic Stability*. London: Springer-Verlag,
- [11] Sundarapandian, V. (2009). "7. Queueing Theory". *Probability, Statistics and Queueing Theory*. PHI Learning.
- [12] Zadeh, L. et al. (1996). *Fuzzy Sets, Fuzzy Logic, Fuzzy Systems*, World Scientific Press,