CHARACTERIZATION OF HUMAN AND INSTRUMENT RELATED MEASUREMENT ERRORS

Thesis submitted under the Faculty of Engineering

for the award of the degree of

Doctor of Philosophy

by

VINODKUMAR JACOB

Under the Supervision of

M. BHASI
Professor & Director, School of Management Studies,
R. GOPIKAKUMARI
Professor & Head, Division of Electronics Engineering

SCHOOL OF ENGINEERING,
COCHIN UNIVERSITY OF SCIENCE AND TECHNOLOGY,
KOCHI – 682 002

February - 2013

COCHIN UNIVERSITY OF SCIENCE AND TECHNOLOGY
Kochi – 682 022

Ph. : +91 - 484 - 2577290 / 2577550
Fax. : +91 - 484 - 2577595
Website : http://www.cusat.ac.in

13 Feb., 2013

Certificate

This is to certify that the thesis entitled ***"Characterization of Human and Instrument Related Measurement Errors"*** *which is being submitted by* ***Vinodkumar Jacob*** *in partial fulfillment of the requirements for the award of the Degree of Doctor of Philosophy, to the Cochin University of Science and Technology, Kochi-22 is a record of the bonafide research work carried out by him under our supervision and guidance, in School of Engineering, Cochin University of Science and Technology, Cochin – 682 022 and no part of the work reported in this thesis has been presented for the award of any degree from any institution.*

Dr. M. BHASI (*Supervising Guide*)
Professor and Director, School of Management Studies
Mobile: 09447419863, *Email*: mbhasi@cusat.ac.in

Dr. R. GOPIKAKUMARI (*Co-Guide*)
Professor and Head, Division of Electronics
Mobile: 09446129193, *Email*: gopikakumari@gmail.com

DECLARATION

I hereby declare that, the work presented in this thesis entitled ***"Characterization of Human and Instrument Related Measurement Errors"*** *is based on the original research work carried out by me under the guidance and supervision of Dr. M. Bhasi (Supervising Guide), Professor and Director, School of Management Studies and Dr. R. Gopikakumari (Co-Guide), Professor and Head, Division of Electronics, Cochin University of Science and Technology, Cochin – 682 022 and no part of the work reported in this thesis has been presented for the award of any degree from any institution.*

Kochi,
13 Feb., 2013

VINODKUMAR JACOB

ACKNOWLEDGEMENT

I have received a large amount of guidance and assistance from many people with whom I had to interact during my Ph.D. programme at Cochin University of Science and Technology (CUSAT), Kochi. To all of them I express my gratitude. This endeavor would not have been possible without the guidance and support of my research gurus, Dr. M. Bhasi and Dr. R. Gopikakumari. Words are inadequate to express my profound sense of gratitude for both of them.

Dr. M. Bhasi is an unmatched person. The interactions with him have been a great learning experience to me. As a research guide, he has always shown total commitment, enthusiasm and involvement. He always keeps himself abreast with latest developments in the subjects and research area, and he is always a source of ready reference. I salute his kindness and empathy. I extremely appreciate the values, which he always upholds. He has been very generous with both his time and patience in providing insights and new windows during my investigation.

Dr. R. Gopikakumari has been a great source of strength and inspiration to me in each and every phase of the research. Her commitment to research is outstanding. The kindness and attention she showered on me are unforgettable. I admire her style of working and thankful for her critical comments and valuable suggestions. I shall always cherish her support and encouragement.

I am deeply indebted to the Director and Staff of Sophisticated Test and Instrument Centre (STIC), Kochi, Fluid Control and Research Institute (FCRI), Palakkad and the State Legal Metrology Lab, Kakkanad, Ernakulam for providing the calibration data that was used extensively in my research. I place on record my sincere gratitude to the authorities of Electronic balance manufactures ELCON, Valakam, Muvattupuzha, INTERTECH, Kayamkulam, and MARVEL, Perumbavoor, Kerala for providing type approval and verification data. I remain grateful to the technicians and engineers of various companies such as Cochin Refineries, Sevana Electricals, Keltron, OEN and faculty members and instructors of various Engineering Colleges for their co-operation and support in the experimental and survey study. I devote special thanks to Mr. Ismail of legal metrology, Mr. Pramod, STIC, Mr. Alex, FCRI, Mr. Babu, CUSAT, Prof. Unni A. M., CUSAT and Prof. Mathew K., MACE for their constant support during my research programme.

I would also take this opportunity to thank Dr. Sam Thomas, SMS, CUSAT, Dr. P. S. Sreejith, Principal, School of Engineering, Faculty members of Division of Electronics, CUSAT, Mr. Jays Chandy and my fellow research scholars.

I express my sincere gratitude to the authorities of M. A. College of Engineering, Kothamangalam for giving me permission to do the doctoral work. The help given by the Electronics and Communication Engineering Department and the colleagues of M.A College of Engineering is also acknowledged.

Finally, I express my heartfelt thanks to my beloved ones for the love and support bestowed on me over the years and the forbearance and cheerful disposition, which were vital for sustaining the effort, required for completing my doctoral programme.

VINODKUMAR JACOB

ABSTRACT

Measurement is the act or the result of a quantitative comparison between a given quantity and a quantity of the same kind chosen as a unit. It is generally agreed that all measurements contain errors. In a measuring system where both a measuring instrument and a human being taking the measurement using a preset process, the measurement error could be due to the instrument, the process or the human being involved. The first part of the study is devoted to understanding the human errors in measurement. For that, selected person related and selected work related factors that could affect measurement errors have been identified. Though these are well known, the exact extent of the error and the extent of effect of different factors on human errors in measurement are less reported. Characterization of human errors in measurement is done by conducting an experimental study using different subjects, where the factors were changed one at a time and the measurements made by them recorded.

From the pre-experiment survey research studies, it is observed that the respondents could not give the correct answers to questions related to the correct values [extent] of human related measurement errors. This confirmed the fears expressed regarding lack of knowledge about the extent of human related measurement errors among professionals associated with quality. But in post-experiment phase of survey study, it is observed that the answers regarding the extent of human related measurement errors has improved significantly since the answer choices were provided based on the experimental study. It is hoped that this work will help users of measurement in practice to better understand and manage the phenomena of human related errors in measurement.

The second part of this research work concentrated on the characterization of errors observed during calibration done periodically (effect of time) of selected sophisticated instruments and selected standards used in legal metrology. The extent of errors due to passage of time and use, were found for some sophisticated instruments and some standards used in legal metrology. These studies have enabled the researcher to characterize errors in these instruments and thus add to the understanding of measurement errors. In order to make the data collected more useful, Regression and Artificial Neural Network [ANN] based models have been developed to predict error [extent] for instrument type and standard types studied. A theoretical method to determine the uncertainty budget through mathematical model from the calibration data of Digital multimeter and Digital thermometer has been also derived.

This research also investigates how to introduce the concept of uncertainty in the criteria of conformity in type approval and verification by analyzing the type approval and test data of non-automatic weighing instrument [NAWI]. It is observed that, the uncertainty in type approval is different from the uncertainty in the performance of measuring instruments. In type approval, if the uncertainty is equal to or less than one-third of the maximum permissible errors, the uncertainty should be included in the criteria of conformity. If the measurement values inclusive of the uncertainty are within the maximum permissible errors, it will be categorized as conforming, and if it is outside the maximum permissible errors, categorized as non-conforming. Only what can be measured, can be controlled. Following this logic, this thesis presents some studies to understand, measure and control measurement errors.

Key Words: *Human Errors, Characterization of Errors, Sophisticated Instruments, Working Standards, Prediction of Errors, Artificial Neural Network.*

Bio - Data

Personal:

Name	:	***Vinodkumar Jacob***
Date of Birth	:	*04 May, 1969*
Marital Status	:	*Married*
Permanent Address	:	*Pazhookudy* *Ramalloor,* *Kothamangalam P. O.,* *Pin 686 691.*
Official Address	:	*Associate Professor* *Electronics and Communication Engineering* *M. A. College of Engineering,* *Kothamangalam – 686 666*
Telephone	:	*0485 – 2823222 (Res.) Mob. 98461 21223* *0485 – 2822363 (Off.)*
Email	:	*vinodkumar_jacob@rediffmail.com*

Academic Qualifications:

- *B. Tech. from Madurai Kamaraj University with First class in 1990*
- *M. Tech. in Biomedical Engineering from IIT Bombay in the year 1998*
- *Doing Ph. D. in Characterization of Human and Instrument Related Measurement Errors in Cochin University of Science and Technology.*

Experience:

22 years of teaching experience in M. A. College of Engineering, Kothamangalam.

Areas of Interest:

Measurement Errors, Human Factor Engineering, Control Systems and Signal Processing

Vinodkumar Jacob
13 February, 2013

Table of Contents

Chapter V

Measurement Uncertainties in Sophisticated Instruments

Chapter VI

Errors and Uncertainties in Type Approval and Verification in Legal Metrology

List of Tables

List of Figures

Abbreviations

AAG	:	Annual Average Growth
AC	:	Analog Current
AMM	:	Analog Multimeter
AN	:	Afternoon
ANN	:	Artificial Neural Networks
ANOVA	:	Analysis of Variance
AUC	:	Area Under Curve
B. Tech.	:	Bachelor of Technology
CRO	:	Cathode Ray Oscilloscope
CS	:	Commercial Standards
DC	:	Direct Current
Df	:	Degrees of Freedom
DMM	:	Digital Multimeter
DSO	:	Digital Storage Oscilloscope
DUT	:	Device Under Test
ER1	:	Student subjects with Kerala State Engineering Entrance Rank between 1 and 5000
ER2	:	Student subjects with Kerala State Engineering Entrance Rank between 5001 and 15000
ER3	:	Student subjects with Kerala State Engineering entrance rank above 15000
ET	:	Experienced Technicians
ET1	:	Experienced Technicians in the age group of 31-40 years
ET2	:	Experienced Technicians in the age group of 41-50 years
FCRI	:	Fluid Control and Research Institute
FET	:	Field Effect Transistor
FN	:	Forenoon
FPR	:	False Positive Rate
IE1 / IQ1	:	Inexperienced Engineers / Diploma holders with above average IQ
IE2 / IQ2	:	Inexperienced Engineers / Diploma holders with average IQ
IE3 / IQ3	:	Inexperienced Engineers / Diploma holders with below IQ
IQ	:	Intelligent Quotient

MPE	:	Maximum Permissible Error
MPES	:	Maximum Permissible Error in Service
MPEV	:	Maximum Permissible Error on Verification
NAWI	:	Non Automatic Weighing Instrument
PG	:	Pressure Gauge
RAP	:	Resistance Analog Parameter
RDP	:	Resistance Digital Parameter
ROC	:	Receiver Operating Characteristic
SG	:	Signal Generator
STIC	:	Sophisticated Test and Instrument Calibration
TPR	:	True Positive Rate
VAP	:	Voltage Analog Parameter
VDP	:	Voltage Digital Parameter
VSWR	:	Voltage Standing Wave Ratio
WS	:	Working Standards

Chapter I

Introduction

1.1 MEASUREMENT SYSTEMS

Measurement systems are used every day in manufacturing, research, and commerce. They are a critical component in the quality of product and service the company provides to its customers and is one of the areas of significant investments [1]. Measurements are the window through which one can look at products and processes, and it is necessary to know whether the image that is seen is accurate or, perhaps, somewhat distorted [2]. Quality or accuracy of measurement systems is essential to the quality of a manufacturing process, because the measurement process itself is subject to variation, and excessive variation in the measurement systems can make critical variation in the manufacturing process [3]. Often measurements are made with little regard for the quality of such measurements [4]. Yet all too often, the measurements are not perfectly representative of the true value of the characteristic being measured and this could be because the measurement system is not accurate enough or not precise enough [5]. The moral is that before one embarks on using a new measurement system for a characteristic which has not been previously measured on it, one should perform a measurement system analysis because this is critical to the success of every measurement and will ensure that future measurements are representative of the characteristic being measured [6].

Measurement is a process of gathering information from a physical world and comparing this information with agreed standards [7]. Measurement is carried out with instruments that are designed and manufactured to fulfill given specifications [8]. After an instrument is designed and prototyped, various evaluation tests are conducted. These tests are typically made under reference conditions or under simulated work environment conditions [9]. Some examples of reference condition tests are accuracy, response time, drift, and warm up time

[10]. Simulated environmental test are based on statutory requirements regulated by government and other authorities [11]. Some simulated environment tests include climatic test, drop test, dust test, insulation resistance test, vibration test, electromagnetic compatibility (EMC) tests, and safety and health hazard tests [12]. Many of these tests are prescribed by company, industry, national and international standards [13]. Adequate testing and proper use of instruments is important to achieve the accurate results when in use [14]. After instruments are installed, regular calibration is necessary to ensure the accuracy over the period of life of operation [15].

1.2 MEASUREMENT ERRORS AND ERROR CONTROL SYSTEMS

Measurement error is defined as the difference between the distorted information and the undistorted information about a measured product, expressed in the measurement units [16]. In short, an error is defined as real (untrue, wrong, false, no go) value at the output of a measurement system minus ideal (true, good, right, go) value at the input of a measurement system.

$$\text{i.e.; } \Delta x = x_r - x_i \qquad (1.1)$$

where Δx is the error of measurement, x_r is the observed measurement value, and x_i is the true value being measured. The aim is to make x_r as close as possible to x_i. That is, reducing error to zero [17]. Besides zero error, an ideal or perfect instrument would have perfect sensitivity, reliability, and repeatability over the entire range of values measured [18]. However in practice, most measurements will be imprecise and will give inaccurate results due to many internal and external factors [19]. This departure from the expected perfection is called the error. Often, sensitivity analyses are conducted to study the effect of individual components that cause these errors [20]. Sensitivity analysis is conducted by varying one parameter at a time while keeping the others constant. This can be done physically by using the developed instruments or estimated mathematically by means of appropriate models [21]. In instrumentation systems, errors have been broadly classified as systematic and random error [22].

1.2.1 Systematic Errors

Systematic errors remain constant with repeated measurements [6]. They can be divided into two basic groups: the instrument related errors and the environment related errors [23]. Instrumental related errors are inherent within the instrument, arising because of the mechanical structure, electronic design, improper adjustments, wrong applications, and so on. These can also be sub classified as loading error, scale error, zero error, and response time error [24]. Environmental errors are caused by environmental factors such as temperature, humidity, etc. [25].

1.2.1.1 Types of Systematic Errors

The three most common types of systematic errors are: i. Instrumental error ii. Operator error and iii. Method error.

i. Instrumental error

A common systematic error is an incorrectly calibrated instrument that systematically gives results that are either high or low [26]. Calibration determines the error associated with a measurement, and, if possible, reduces that error [8]. This means that, calibration is more than just adjusting the measurement capability of a device. Calibration improves measurement accuracy and ensures that the product meets the required specifications [27]. Calibration also provides a number of other benefits, such as: increased production yields, overall measurement consistency, standardization [28]. To maintain measurement quality, the measuring device must be calibrated at set intervals.

The calibration process includes three parts:

- verifying that the measurement capability of the measurement device is within specifications;
- adjusting the device to reduce its measurement error;
- verifying the new measurement capability of the device to ensure that is operating within specifications [29].

ii. Operator error

It is the error, which is observed during measurement process even after the possibility of instrument error and method error is detached by proper calibration of the instrument and with repeated measurements. Operator errors may occur in the process of observations or during the recording and interpretation of experimental results [30]. A large number of errors can be attributed to carelessness, improper setting up of instruments, the lack of knowledge about the instrument and the process, and so on [31]. These systematic errors, sometimes cannot be traced, and often can create quite large errors [32]. Though experimentation and observation, scientists are learning how to minimize the human factor related measurement error. This is one of the problems studied in this thesis.

iii. Method error

Systematic errors caused by measuring the same quantity using two different methods and not getting the same result is known as method error [33]. As an example, if one measures a given voltage using two different methods -

analog and digital voltmeters and gets different readings, the error involved is said to be method error.

1.2.2 Random Errors

The errors that remain after the instrument operator and method errors have been removed may be called random errors. They are small and within the limits for the measuring instruments. Random errors appear as a result of rounding of reading, noise and interference, backlash and ambient influences, and so on [15]. In measurements, the random errors vary by small amounts around a mean value. Therefore, the future value of any individual measurement cannot be predicted in a deterministic manner [34]. Random errors may not easily offset electronically; therefore, in the analysis and compensation, stochastic approaches are adapted by using the laws of probability [35]. Depending on the system, the random error analysis may be made by applying different probability distribution models [36]. This thesis presents a method of using random error effect in uncertainty budget.

1.2.2.1 Types of Random Errors

- Transient qualities of the individual (mood, motivation, degree of alertness, boredom or fatigue) [37]
- Situational factors involve the physical setting such as (noise level, lighting, ventilation etc...), anonymity, presence of peers [38].
- Administrative factors involve the actual administration of the instrument or the amount of subjectivity influencing the measurement process [39].

Studies show that, in applied science and engineering, all observations contain errors. It is reported that errors can be due to measurement device and/or due to the human being involved in the measurement process [40]. It is also reported that humans are not very good at thinking logically (algorithmically), but for measurement, following procedures is often necessary [41]. Deviations from set procedure are a common source of error in measurement [42].

Literature review on measurement errors and discussions with professionals also indicated that if error reduction techniques can be effectively applied in the manufacturing and service industries, they would gain better control over the processes and expenses [43][13]. It would also improve customer satisfaction and thus strengthen their level of competitiveness [44]. Therefore, it is evident that a clear understanding of the critical factors causing measurement errors in method, measurement and instrument would help the implementation of management practices to improve quality and enable better control over processes and cost [45]. Motivated by these observations, the research problem, objectives and methodology of the research work have been formulated.

1.3 RESEARCH PROBLEM

Measurement is essential for scientific investigations. It is so fundamental and important to Science and Engineering that the whole science can be said to be dependent on it. Instruments are developed for measuring and displaying physical variables. Every act of measurement has to deal with errors [46]. Errors can result in negative consequences [47] [example: loss of time, faulty products] as well as positive ones [example: learning, innovation]. The large negative consequences - for example, accidents such as the Chernobyl or Challenger disasters – tend to be widely observed, and have been of high interest to scholars and laypeople alike [48]. The scientific understanding of the negative effects of errors is much better developed than that of the potential positive effects of errors [49]. One way to contain the negative and to promote the positive consequences of errors is to use error management. This approach assumes that human errors per se can never be completely prevented, and, therefore, it is necessary to ask the question of what can be done after an error has occurred [50]. Errors are not easily defined. Errors may be unintended deviation from goals, standards, a code of behavior, the truth, or from some true value [51]. Literature review and discussions with professionals and experts was used to identify the area of research as ***"Characterization of Human and Instrument related Measurement Errors"***. Measurement system in this research study can be visualized as consisting of three parts as shown in *Figure. 1.1*.

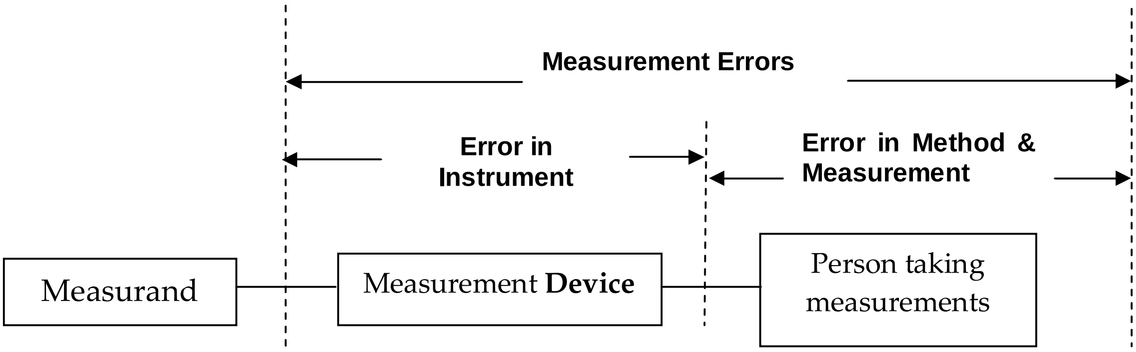


Figure 1.1 Measurement system

Measurand is the physical parameter being measured. The measuring device can be of different types such as Electrical, Electronic and Mechanical etc. Measurement errors may be due to one or more of factors such as measuring device and/or the method and the person involved in measurements. Measurement error is defined as the difference between the output of the measurement system and the reference [known, actual, true, master, and standard] value. The measurement system could be defined as only the measuring instrument [narrow view] or as comprising of the measuring instrument and the person taking or doing the measurement and reporting the measured output [broader view used by us for this research]. Discussions with test engineers and professionals in the area of measurement reveals that they know that human related measurement errors exist and there is a possibility for calibration error growth with ageing, they are unsure of the extent of such errors. Therefore it is decided to conduct a study on these two factors. The next section gives each of the objectives (in italics) and the methodology followed to achieve it.

1.4 OBJECTIVES AND METHODOLOGY

- *To study the effect of selected human related factors on measurement errors.*

Experimental analysis on human related measurement errors was conducted for understanding and quantifying the critical factors. Many experts supervising production and quality control in different production environments and engineering academicians were approached to find out the factors influencing human error in measurement. Simple experimental setups and experiments were designed and developed to analyze the impact of the identified human related factors on measurement errors. Experiments were conducted using subjects of different categories.

- *To study the perception of professionals on the extent [values] of human related measurement errors.*

A survey based research study was conducted to understand the perception on extent of human related measurement errors among professionals associated with measurement and testing. The survey instrument was developed after discussions with experts to measure the understanding of human related measurement errors. Survey research was done in two phases. Pre-experiment survey was conducted in the first phase with two sets of questions. Section A questions dealt with the human related factors that impact measurement error and section B dealt with quantification of these impacts. Answer choices were given for section A while, open ended questions were used for getting the perception of professionals about the range of values of possible human errors in section B in this phase. The second phase was conducted as a post experiment survey where answer choices was given for both section A and B. Pre experiment survey results and Post experiment survey results were compared.

- *To study measurement errors and characterize them for selected types of measuring instruments.*

Characterization of errors with respect to time was done by analyzing calibration data of various sophisticated instruments such as digital multimeter, digital thermometer, cathode ray oscilloscope, signal generator and pressure gauge. Similarly effect of ageing was analyzed using the calibration data of standards such as, non automatic weighing instrument, weight measures and volumetric measures used in legal metrology.

- *To develop a method for finding out uncertainty budget from the calibration data of digital multimeter and digital thermometer.*

  Calibration results of DMM and digital thermometer were used to determine uncertainty budget and then develop a theoretical method determining uncertainty budget using a mathematical model.

- *To study the uncertainty in type approval and verification using the calibration data of non-automatic weighing instrument and to develop a system for comparison of calibration and verification.*

- *To introduce the concept of uncertainty in the criteria of conformity in type approval and verification using the type approval and verification data of non-automatic weighing instrument.*

The above two objectives were achieved by taking the type approval and verification data of Class I and II Non Automatic Weighing Instrument [NAWI]. Fifteen instruments, each in the category of class I and class II were taken from three different companies which manufacture NAWI. Out of the four different classes of NAWI, only class I and II have been taken for this study since the influence of uncertainty is very large for class I and II where as for class III and IV, it is negligible.

- *To develop systems for predicting human related measurement errors using artificial neural network [ANN].*

The error data obtained from the experimental study on human related measurement errors were used for developing, training and testing an ANN model. The observations from the experiments were used as the input for the error prediction model of human related measurement errors, for training and testing ANN, and were built to predict the probable errors due to the influence of various human factors under the circumstance of a typical measurement.

- *To develop systems for predicting calibration errors of sophisticated instruments and working standards in legal metrology using ANN prediction model and regression analysis techniques.*

A methodology similar to the one used above was used with the calibration data of various sophisticated instruments and working standards in legal metrology to develop, train and test the neural network prediction model. The same input data was used in the regression model.

1.5 ORGANIZATION OF THE THESIS

The thesis is presented in eight chapters. The chapters are organized as follows:

The *second chapter* discusses the survey of literature relevant to the study. In the *third chapter*, the experimental methods and results to quantify the extent of human related factors on measurement errors are given. The *fourth chapter* deals with the survey based study to understand the perception of professionals on human related measurement errors. The *fifth chapter* is dedicated to characterization of errors observed during calibration in terms of time for selected sophisticated instruments and to suggest a theoretical method for calculating uncertainty budget. The *sixth chapter* explains the effect of ageing in selected standards used in legal metrology and describes the proposed methodology of the concept of uncertainty in the criteria of conformity in type approval and verification. The *seventh chapter* presents regression analysis and artificial neural network based models for measurement error prediction. The *eighth chapter* reports the research findings as a summary, lists the limitations of the work and ends with scope for future research. The list of references comes after this. The questionnaire, photograph of experimental set up, features of instruments used in the experiment and the specification of the components used in the pattern generator are given as an appendix at the end of the thesis.

Chapter II

Literature Review

2.1 INTRODUCTION

In this chapter, a review of literature is presented. Theoretical concepts in the area of measurements, measurement errors and types of measurement errors are reviewed. Since this work focuses on error in method, measurement and error in instrument, for clarity of presentation, the literature review is divided into three sections. Error in method and measurement is nothing but human error. Being a practice dominated area; the impact of human errors, reported in the literature is reviewed. Error in instrument is studied in this work by characterizing errors observed during calibration and measurement uncertainties of various sophisticated instruments and standards in legal metrology; a detailed review has been done of this area. Prediction error model is developed in this work using Artificial Neural Network. Estimation of errors is also done using regression analysis technique. Therefore, concepts of ANN and Regression analysis and the research work in this area are also reviewed.

It is explained in the literature that Measurement is a process of gathering information from a physical world and comparing this information with agreed standards [31]. Measurement is carried out with instruments that are designed and manufactured to give correct measurement [7], [44].

Measuring instruments are supposed to maintain prescribed relationships between the parameters being measured and the physical factors under investigation [52]. Literature says that measurement is essential for observing and testing scientific and technological investigations [47]. Instruments are developed for monitoring the conditions of physical factors and converting them into symbolic output forms. Measurement error might be because, the measurement system is not accurate enough, or precise enough, or can be because of the human being involved in the measurement [53].

2.2 IMPACT OF HUMAN RELATED MEASUREMENT ERRORS

Measurement error can occur due to the human factor, because the measuring system has a human component [54]. Literature says that, a human factor is a physical or cognitive property of an individual or social behavior, which is specific to humans and influences functioning of technological system as well as human environment equilibriums [30]. Human factors involves the study of all aspects of the way humans relate to the world around them, with the aim of improving operational performance, safety, through life costs and/or adoption through improvement in the experience of the end user [55].

Researches show that human factors study is a multidisciplinary field, incorporating contributions from psychology, engineering, industrial design statistics, operations research and anthropometry [56]. It is a term that covers:

- the science of understanding the properties of human capability (Human factor science)
- the application of the understanding to the design, development and deployment of systems and services (Human factors engineering)[31]
- the art of ensuring successful application of Human Factors Engineering to a program (sometimes referred to as human factors integration). This is a part of ergonomics [57].

2.2.1 Human Factors

Human factors involve the study of factors and development of tools that facilitate the achievement of these goals. After defining the problem there are five different approaches that can be used in order to implement the solution [58]. These are as follows:

- ***Equipment design***: changes the nature of the physical equipment with which humans work.
- ***Task design***: focuses more on changing what operators do than on changing the devices they use. This may involve assigning part or all of tasks to other workers or to automated components [59].
- ***Environmental design***: implements changes, such as improved lighting, temperature control and reduced noise in the physical environment where the task is carried out.
- ***Training the individuals***: better preparing the worker for the conditions that he or she will encounter in the job environment by teaching and practicing the necessary physical or mental skills [60].
- ***Selection of individuals***: is a technique that recognizes the individual differences across humans in every physical or mental dimension that is relevant for good system performance. Such a performance can be optimized by selecting operators who possess the best profile or characteristic for the job [20].

Human error literature [61] has revealed several perspectives on the nature and causes of human error. Some of the perspectives that have been reported are cognitive, ergonomics and system design, psycho-social and organizational [62]. In a cognitive model, the assumption is that information progresses through a series of stages of mental operations (eg. Attention allocation, pattern recognition and decision making) that mediate between stimulus input and response execution [63]. Within this approach, errors occur when one or more of these mediating operations fail to process information appropriately. Literature gives a detailed development of taxonomic algorithm for classifying various types of information processing failures associated with the erroneous actions of operators [64]. According to the system perspective, the human is rarely, if ever, the sole cause of an error or accident. Rather, human performance (both good and bad) involves a complex interaction of several factors. In fact, "System models recognize the inseparable link between individuals, their tools and machines, and their general work environment" [65]. Some of the psycho-social models seen in the literature highlight the social and interpersonal aspects of human performance within the context of measurement, a perspective that has historically been overlooked by those in the industry [66]. Organizational approaches to human error have been utilized in a variety of industrial settings for management of human errors. In fact, it is this emphasis that organizational models place on the fallible decision of managers, supervisors, and others in the organization that sets them apart from the other perspectives previously discussed [21].

Human error is a topic that researchers and academicians in the fields of human factors and psychology have been grappling with for decades. Indeed, there are a number of perspectives on human error, each of which is characterized by a common set of assumptions about the nature and underlying causes of errors. Unfortunately, from the practitioner's point of view, there often appears to be as many human error models and frame works as there are people interested in the topic [67]. Even worse, most error models, and frame works tend to be theoretical and academic, making them of little benefit to the applied needs of practitioners. Therefore, having been left without adequate guidance and objective criteria for choosing a particular approach, many practitioners have resorted to developing error – management programs based on intuition or "Pop Psychology", rather than on theory and empirical data. The end results are safety programs that, on the surface, produce a great deal of activity (eg. Safety seminars and "Error awareness" training), but in reality only peck around the edges of the true underlying causes of human error. This is the reason, why, it was decided to conduct an experimental analysis, varying selected factors that possibly influence human related measurement errors as part of this research work.

2.3 MEASUREMENT ERRORS IN SOPHISTICATED INSTRUMENTS

Literature shows that the standard uncertainty associated with estimate has the same dimension as the estimate itself [68]. In some cases, the relative standard uncertainty of measurement may be appropriate, which is the standard uncertainty associated with an estimate divided by the modulus of that estimate, and is therefore dimensionless. This concept cannot be used if the estimate equals zero [69]. The standard uncertainty of the result of a measurement, when that result is obtained from the values of a number of other quantities is termed combined standard uncertainty. An expanded uncertainty is obtained by multiplying the combined standard uncertainty by a coverage factor [70]. This, in essence, yields an interval that is likely to cover the true value of the measurand with a stated high level of confidence.

Studies related to uncertainty and error analysis of sophisticated instruments are very few and some of the research studies are as follows: Literature gives a rigorous uncertainty analysis to compare the nose-to-nose swept-sign and electro-optic–sampling system based calibrations [29]. They have also reported that characterization of systematic error in the nose-to-nose calibration is statistically significant. Study of memory soft error measurement on production systems was reported [71]. An understanding on the memory soft error rate demystifies an important part of whole-system reliability in today's production computer systems [1]. It was reported in the literature, a technique based on multiple linear regressions for predicting memory soft error rate [72]. A method for diagnosing and correcting systematic errors in the spectral imaging system was reported [23]. In that, several sets of imaging system calibration and spectro-photometric targets can be employed to both diagnose and correct systematic errors. In their study, the systematic errors were characterized as a function of wave length. It is reported that, the R^2 value as an index of regression fitting performed well in imaging system when imaging system calibration and spectro-photometric correction targets were identical [73].

Literature gives a specific strategy to address the issue of taking into account the error bars in the measurements used as inputs to fuzzy logic systems [74]. This approach of considering the uncertainty of the measurements as an independent axis of complexity, is conceptually sound for many real systems, and provides a clear improvement in the performance of the fuzzy logic classifiers, when the statistic of the noise is properly taken into account. They clarify that the proposed analysis can provide an independent confirmation of the error bars estimated for various measurements. They also say that with regard to future developments, it would be better to confirm the potential of the strategy for a concrete regression problem. Literature also explains different types of errors in various sophisticated instruments like Digital Multimeter [75], Digital Thermometer [76], Cathode Ray Oscilloscope [77], Signal Generator [78] and Pressure Gauge [68].

2.4 MEASUREMENT ERRORS AND UNCERTAINTIES IN LEGAL METROLOGY

Correctness of measurements and measuring instruments is one of the most important prerequisites for the assurance of the quality and quantity of products and services, and the accuracy of the instruments must be consistent with their intended use.

In compliance with the ISO 9000 standard series and the ISO/IEC 17025 standard, traceability of measuring and test equipment to the realization of SI units must be guaranteed by an unbroken chain of comparison measurements to allow the necessary statements about their metrological quality [79]. It is reported that, the most important actions to ensure the correct indication of measuring instruments are [80]:

- **in industrial metrology** : regular calibration of the measuring instruments according to the implemented quality systems : and
- **in legal metrology:** periodic verification or conformity testing of the measuring instruments according to legal regulations.

Both actions are closely related and are mostly based on the same measuring procedures.

Historically, however, these actions have been established with separate rules and metrological infrastructures and activities. Verification has become a principal part of legal metrology systems and calibration is widely used in quality assurance and industrial metrology. Accreditation bodies prefer calibration as a primary action to provide proof of the correctness of the indication of measuring instruments [81].

The relationship between the state and metrology is symbiotic. The State needed measurements to provide the information necessary to organize, plan, defend and tax system with efficiency. Such accounting depended on uniform measurements across wide geographical areas and across a broad spectrum of farming and manufacturing practices and work organization [14].Metrology on the other hand required the mandate of the State to ensure conformity to measurement requirements [82]. The fundamental requirement, to ensure consistency, was that all measurement be derived from (royal) standards and this can be defined as traceability [83].

According to literature [84], one institutional aspect of the development of metrology in the 20th century was the separation of metrology in many countries into scientific metrology, led by the National Measurement Institute [NMI's,] and practical or legal metrology, administered by weights and measures authorities. The need to ensure international consistency of trade and regulatory measurements, and to "resolve internationally the technical and administrative problems raised by the use of measuring instruments", led to the establishment in 1955 of a second metrology Treaty organization, the International Organization of Legal Metrology (OIML) [85]. In the past, various interests as well as regional and historical differences led to differing units and systems [26] reported that as cross-border trade increased in significance, pressure grew for harmonization; this resulted in the introduction of the SI system which not only became the legal basis for official dealings and commercial transactions, but also gained in importance in the non-regulatory field of industrial metrology. An efficient metrological infrastructure is the basis of all modern industrial societies and from this point of view; legal metrology was the pioneer of uniform measurement [25]. In type approval tests, a design of a new type of specified measuring instrument is examined in order to decide the conformity to technical requirements [86]. Type approval tests consist of tests on the characteristics that solely depend on the design. Therefore, one or a few samples of measuring instrument are subjected to type approval tests. When type approval is obtained, the design of this type of instrument shall meet the requirements of legal regulation and perform adequately under the conditions of practical use [69].

According to literature [87], verification is carried out in order to decide the conformity of individual items of type-approved specified measuring instrument to the requirements for performance. Verification tests consist of the tests on characteristics that depend on each item [27]. If each item makes a major instrumental error, some of the tests performed in type approval could be applicable to verification as well. Actually, the test results of type approval can be substituted in verification [88].

It is reported that, calibration is carried out in order to provide a quantitative statement about the correctness of the measurement results of a measuring instrument [89]. For economic reasons, laboratories strive for broad recognition of their calibration and measurement results. Confidence in results, therefore, is achieved through both establishing the traceability and providing the uncertainty of the measurement results.

The International vocabulary of basic and general terms in metrology (VIM) [90] defines a calibration as "a set of operations that establish, under specified conditions, the relationship between values of quantities indicated by a measuring instrument or measuring system, or values represented by a material measure or a reference material, and the corresponding values realized by standards". This means that the calibration shows how the measured value or the nominal value indicated by an instrument relates to the true or conventional true values of the measurand. It is assumed that the conventional true value is realized by a reference standard traceable to national or international standards.

Not only is the measurement uncertainty but also the environmental conditions during the calibration has significant impact. Calibration is often carried out in a place with standard environmental conditions, which leads to low measurement uncertainties [69]. When the calibrated instrument is used in a different environment the measurement uncertainty determined by the calibration laboratory will often be exceeded if the instrument is susceptible to its environment. There can also be a problem if instrument performance deteriorates after prolonged use. The user of the calibrated instrument must therefore consider any environmental or secular stability problems [91]

It was observed from literature review that, research on the sophisticated instruments are based and focused on calibration errors and other types of errors in the instrument. Literature also provides a theoretical understanding of the possible errors in various instruments. Measurement uncertainty study in sophisticated instruments was not found. In legal metrology, measuring instruments for commercial transactions and certifications are regulated by the Measurement Law of the respective country and called "specified measuring instruments". Calibration and uncertainty study of weigh-bridge and weighing instruments were reported [79], [80]. Studies considering accuracy for deciding instrumental errors in inspection of verification standard were reported [27]. Reliability studies which demonstrate the difference of scope in the measurement of change [25] were also seen. No work-related to the characterization of calibration errors of various instruments were found. Discussions with the professionals and practitioners revealed that characterization of calibration errors would help them to get an idea about the extent of error growth with ageing and would enable them to decide about the period of calibration. The legal regulatory system demands that measuring instruments pass type approval tests and be subject to verification before they enter the market. In the current type of approval tests and verification, evaluation is only based on whether or not the test results fall within the maximum permissible error on verification (MPEV) specified by the Measurement Law [36]. Here the uncertainty of measurement is not considered at all. The current criteria for deciding conformity has less reliability because of the uncertainty of measurement. It is also reported that, to improve the reliability, measurement uncertainty also needs to be taken into account [69]. Therefore the research work was extended to capture the concept of the measurement uncertainty into the criteria for deciding conformity.

2.5 PREDICTION ERROR MODEL USING ARTIFICIAL NEURAL NETWORKS

An Artificial Neural Network (ANN) is inspired by the biological brain, which consists of billions of interconnected neurons working in parallel. An ANN is a network of highly interconnecting processing elements (neurons) operating in parallel [92]. As in nature, the connections between elements largely determine the network function. A subgroup of processing element is called a layer in the network. The first layer is the input layer and the last layer is the output layer. Between the input and the output layer, there may be additional layer(s) of units, called hidden layer(s) [93]. Learning in ANNs takes place through an iterative training process during which node interconnection weight values are adjusted. Initial weights, usually small random values, are assigned to the interconnections between the ANN nodes. One can train a neural network to perform a particular function by adjusting the values of the connections (weights) between elements [94].

A novel approach to credit risk evaluation using a neural network is seen in the literature [95]. In this approach they trained a three-layer supervised neural network, which is based on the back propagation learning algorithm, following seven learning schemes. A comparative research review of use of three famous artificial intelligence techniques, i.e., artificial neural networks, expert systems and hybrid intelligence systems, in the financial market were reported [96]. Literature also gives the design of a hybrid intelligent system for credit ranking using reasoning-transformational models [93]. Expert system as symbolic module and artificial neural network as non-symbolic module are components of this hybrid system. An ANN training algorithm inspired by the neurons' biological property of meta-plasticity was reported [94]. It is reported that this algorithm is especially efficient when few patterns of a class are available, or when information inherent to low probability events is crucial for a successful application, as weight updating is overemphasized in the less frequent activations than in the more frequent ones [97].

Although the long-term goal of the neural-network community remains the design of autonomous machine intelligence, an important modern application of artificial neural networks is in the field of pattern recognition [19]. In the sub-field of data classification, neural-network methods have been found to be useful alternatives to statistical techniques such as those which involve regression analysis or probability density estimation [8]. It is reported that the potential utility of neural networks in the classification of multisource satellite-imagery databases has been recognized for well over a decade, and today neural networks are an established tool in the field of remote sensing [99]. The most widely applied neural network algorithm in image classification remains the feed-forward back-propagation algorithm [100].

It has also been shown that ANN can generalize traditional constitutive laws well (e.g. a hyperbolic model) by considering their descriptive parameters [101]. Despite their good performance on the available data, ANN models give no clue on the way inputs affect the output and are therefore considered as a black box class of model. The lack of interpretability of ANN models has inhibited them from achieving their full potential in real world problems [102] as the credibility of the artificial intelligence paradigm frequently depends on its ability to explain its conclusion [103]. Therefore for verification of such models, as well as the accuracy measuring of ANN based models with available data; a methodology should be adopted to extract the meaningful rule from the trained networks, which are comparable with trends inferred from experiments.

2.6 PREDICTION ERROR MODEL USING REGRESSION ANALYSIS TECHNIQUE

Regression analysis is a statistical tool for the investigation of relationships between variables. Usually, the investigator seeks to ascertain the causal effect of one variable upon another—for example the effect of a price increase upon demand, or the effect of changes in the money supply upon the inflation rate[104]. To explore such issues, the investigator assembles data on the underlying variables of interest and employs regression to estimate the quantitative effect of the causal variables upon the variable that they influence. The investigator also typically assesses the "statistical significance" of the estimated relationships, that is, the degree of confidence that the true relationship is close to the estimated relationship. Regression analysis with a single explanatory variable is termed "simple regression." "Multiple regression" is a technique that allows additional factors to enter the analysis separately so that the effect of each can be estimated. It is valuable for quantifying the impact of various simultaneous influences upon a single dependent variable. Further, because of omitted variables bias with simple regression, multiple regression is often essential even when the investigator is only interested in the effects of one of the independent variables. At the outset of any regression study, one formulates some hypothesis about the relationship between the variables of interest [105].

Another common statistic associated with regression analysis is the R^2. This has a simple definition—it is equal to one minus the ratio of the sum of squared *estimated* errors (the deviation of the actual value of the dependent variable from the regression line) to the sum of squared deviations about the mean of the dependent variable. R^2 statistic is a measure of the extent to which the total variation of the dependent variable *is* explained by the regression [106]. The R^2 statistic takes on a value between zero and one. A high value of R^2, suggesting that the regression model explains the variation in the dependent variable well, is obviously important if one wishes to use the model for predictive or forecasting purposes.

It is required to understand the capacities, assumptions, and applicability of various approaches, and maximally exploit the complementary advantages of these approaches in order to develop better intelligent systems. Such an effort may lead to a synergistic approach which combines the strengths of ANNs, regression analysis techniques and other approaches in order to achieve a significantly better performance for challenging problems [107]. In such a synergistic approach, not only are individual modules important, but a good methodology for integration is inevitable [108].

Chapter III

Experimental Analysis of Human Related Measurement Errors

3.1 INTRODUCTION

Human fallibility is natural. The internal and external factors that aggravate human fallibility introduce the concept of vulnerability [109]. Humans are vulnerable to their internal make-up as well as external work place factors. These factors are referred to in literature as performance shaping factors, error precursors and error forcing contexts. Ability to perform, in the presence of these negative conditions, decreases, and error probabilities escalate [110]. Most of the time, these factors are neither good nor bad, they are part of work reality, that must be managed. Management of task and environment, primarily external concerns, are the simpler performance shapers to identify and deal with [111].

A three dimensional framework is proposed at this stage. The models must contain at least three critical elements:

i. The human operator

ii. His or her task

iii. The environment or context in which the task is performed [40]

There are also several task taxonomies that entail all or parts of these elements. Human's taxonomy has three major branches: environment, subject and task [112]. The proposed three dimensional matrix for mapping human errors and technological innovations is given below in *Figure 3.1*

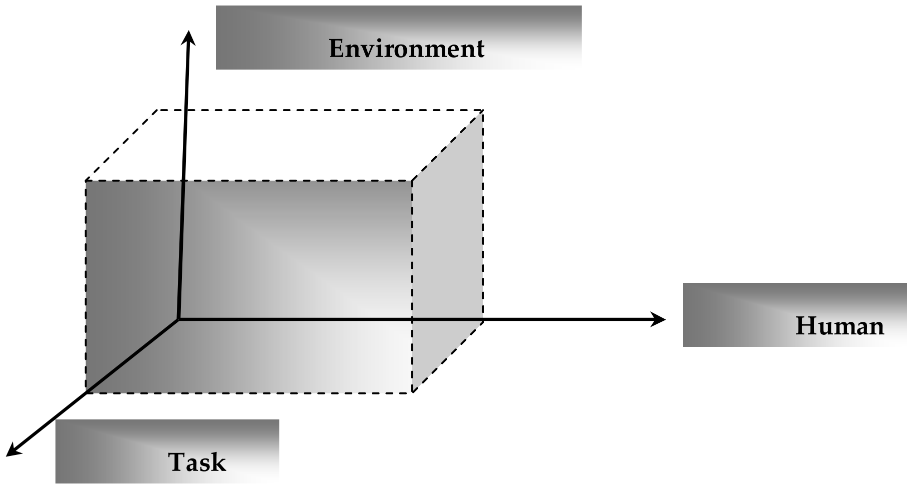


Figure 3.1. Three Dimensional Frame Work of Human Taxonomy

However, while some address human error directly, others do so indirectly. Some attempt to eliminate the occurrence of errors altogether, whereas others look to reduce the negative consequences of these errors [113]. In many works, personal error in a measurement procedure is calculated by taking the ratio of estimated true score or universe variance to observed score variance [114]. In some other practical approaches, the error was determined as the estimated standard deviation of the score distribution that would be obtained if an experiment is done by many examinees [115].

Human error is synonymous to human performance [i.e. poor performance or failure to perform] [116]. Bailey proposed a general qualitative model of human performance that can be generalized to all performance situations as given in *Figure* 3.2 [40].

It is clear that the model must contain at least three critical elements which are, the human operator, his/her task, and the environment in which the task is performed. Although the majority of these elements remain unknown at worst, and poorly understood at best, and although the number of factors and their potential interactions can be bewildering, this research work offers some startling benefits in linking such disparate realms as human performance and factors that affect human performance.

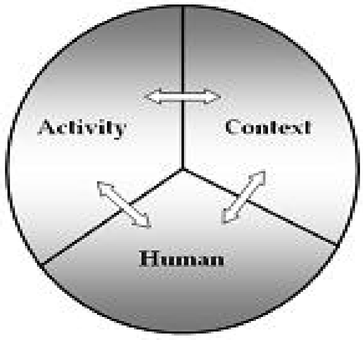


Figure 3.2. Bailey's Qualitative Model of Human Performance [40]

Production and quality control engineers who deploy human resource to take readings from instruments need to understand the effect of various factors on the errors induced by the human resource. Errors induced by human during measurement can be further split into those:

(a) From human related factors
(b) From work related factors.

In order to identify the factors to be considered in this study, a review of literature was done to identify the factors influencing human induced errors in measurement. A survey was also carried out among experts supervising production and quality control in different production environments and engineering academicians to generate a list of possible factors usually found in industry, the effect of which would be useful to explore. These two sources were used to make the list of factors for the study. These factors were then classified into stable and transient. Stable variables were those work related variables that would remain the same over time for the experimental set-up [Voltage, Resistance, Analog, Digital, A/c or Non A/c environments, Forenoon afternoon etc.]. Transient variables would change over time for a given experimental set-up like instrument temperature, input values, ageing etc. A preliminary survey among professionals also revealed that, they are aware that, human errors exist, but they do not have any idea about the range of values of possible errors in different person or work related environment. [Pre-survey and Post-survey analysis results were given in Chapter IV.] Therefore an experimental study was designed and conducted to identify and quantify the impact of a few selected factors on the magnitude of measurement error. The subjects used for doing the experiments were technicians conversant with measurement, inexperienced B. Tech. and Diploma holders and B. Tech. students.

The factors that could possibly impact measurement errors were identified and classified as follows:

Person related factor (Human)	*Work related factor* [Activity (A) and Context (C)]
1.Intelligent Quotient	1. Instrument differences (A)
2. Age	2. Time of work (C)
3. Experience	3. Task differences (A)
4. Gender	4. Time pressure (A)
5. Effect of training	5. Environment (C)

3.2 DESIGN OF EXPERIMENTS FOR TESTING HUMAN ERRORS IN MEASUREMENT

In order to measure the effect of the Person, Activity and Context related factors given above on measurement errors, experiments were designed. The parameters chosen to be measured were Voltage, Resistance, Length and Breadth. Experiments were conducted with all possible combinations of factors, by varying one factor at a time. Twenty replications for each set of factor combination were taken using twenty subjects. The details of the experimental setup and procedure are given below.

i. Voltage

To study the effect of different factors, there was a need to isolate the effect of each factor, therefore in the experiments one factor was changed at a time. In this case, all subjects were asked to do voltage measurement one by one. A setup was made to generate predetermined set of 50 voltages, one after the other to be provided as an input for the subject doing the experiment to measure voltage using two types of voltmeters one analog and the other digital.

The test voltage generator was to ensure that all subjects were given the same set of values. For doing this a Microcontroller based test pattern generator which gives 50 different voltage outputs one after another on press of a button was designed and used for the study. This setup kept errors due to system being measured out of the experiment and the focus could be maintained on error due to observation and noting. Different experiments were carried out using the Microcontroller based test pattern generator where analog and digital meters were used for measurement and the factors such as intelligent quotient, age, experience, gender, effect of training, instrument differences, time of work, task differences, time pressure, and environment of human resource were varied one at a time.

A block schematic of digitally controlled Analog test pattern generator is shown in *Figure 3.3*. This generator gave different prefixed voltages for each hit of a switch.

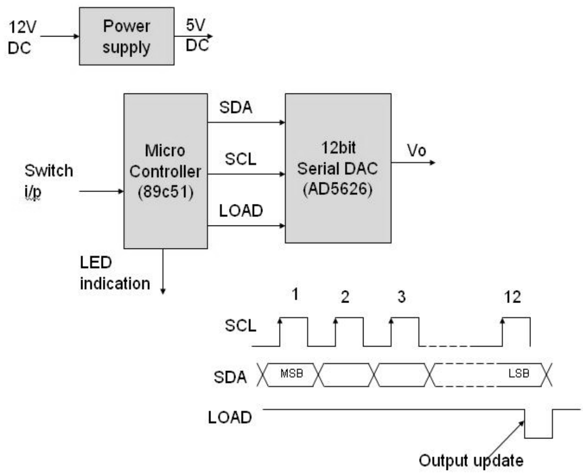


Figure 3.3 Test Pattern Generator

Display devices such as Analog Multi-meter (Make: SANWA, Model: YX-360TRF) and Digital Multi-meter (Make: EZ Digital DM 333) were used for testing Voltage Analog Parameter [VAP] and Voltage Digital Parameter [VDP]. In an experiment a subject had to push the switch for the next reading and make note of fifty such readings consecutively.

ii. Resistance

A set of fifty different valued [code covered] resistances were used for studying the effect of person related and work related factors on human errors in measurement. Resistances numbered from 1 to 50 were given to subjects in the experiments, where they used digital and analog multimeters to make the measurements and note the same. In different experiments, the factors such as intelligent quotient, age, experience, gender, effect of training, instrument differences, time of work, task differences, time pressure, and environment were varied one at a time. Though the method involved in this case is very simple, all subjects were also trained in it to avoid the effect of difference in method of measurement on the error generated.

iii. Length and Breadth

Influence of human factors on measurement error was also studied by providing fifty different sized serially numbered square metal pieces to the subjects. The subjects were asked to measure the length and breadth using analog caliper [Make : Swastik] and digital caliper [Make : Swastik] and to note down the length and breadth of each metal pieces separately on the sheet provided to the subjects.

3.3 SUBJECTS FOR THE EXPERIMENTS

Experiments were conducted using different sets of subjects. The key differentiating factors of the subject groups are given below:

i. Experienced Technicians [ET*] in the age group of 31 to 40 years and 41 to 50 years. Their IQ test showed that all were in the below average [IQ3] category.

ii. Inexperienced [IE] B. Tech. [Bachelor of Technology, four year engineering degree program in India] and diploma holders in the age group of 21 to 30 years and with different IQ levels, above average [IQ1], average [IQ2] and below average [IQ3] category.

iii. B. Tech. students [17 to 22 years] with Kerala State [Country: India] Engineering Entrance Rank between 1 and 5000 [ER1], 5001 and 15000 [ER2] and above 15000 [ER3].

iv. First year [17 to 18 years] and final year [21 to 22] male and female B. Tech students in ER1, ER2 and ER3 categories.

v. 20 subjects from each of the different categories [e.g. ER1, ER2 and ER3] and hence a total of 220 subjects were participated in the experiment. This was found to be statistically sufficient for the mean error measurement which has been taken for analysis in this study [there is no significant change in the measure of mean and standard deviation of the error when sample was increased from 15 to 20].

3.3.1 Intelligent Quotient Test

The intelligent quotient was determined by conducting psychometric test using questionnaire taken from Psychometric test work book and by conducting an online test. The average score obtained by a subject for these two tests was used to categorize the subjects into above average [IQ1 - score above 110], average [IQ2 – score 100 to 110] and below average [IQ3 – score below 100] [117,118].

3.4 EXPERIMENTS

The impact of person related and work related factors on human related measurement error were studied by asking subjects to measure resistances, voltages, length and breadth using analog and digital measuring devices. To check the effect of intelligent quotient, subjects with different IQ levels such as IQ1, IQ2, IQ3 and engineering students in ER1, ER2, and ER3 categories were asked to measure resistances, voltages, length and breadth using analog and digital measuring devices. The effect of experience and age were studied by comparing the measurement error observed in experienced technicians in the age groups of 31 to 40 years, 41 to 50 years and with inexperienced subjects in the age group of 21 to 30 years. The measurement error noticed in the first year and final year B. Tech. students was also compared. The subjects were made to do the experiments during forenoon, afternoon and night hours. The impact of instrument differences was studied by comparing the errors that occurred when using analog and digital readouts. To check the influence of task differences subjects were asked to measure resistances, voltages, length and breadth using analog and digital measuring devices. The effect of the gender factor was studied using first year and final year B.Tech. male and female students with different engineering entrance rank ranges ER1, ER2 and ER3. First, the experiments were done by the subjects without giving any training. But it was observed that training reduces measurement error. The following procedure was adopted for the experiments.

i. Twenty subjects of each category were selected.
ii. On a given day, one category subjects were made to take one set of measurement (say only resistance measurement using analog device) both in the forenoon and afternoon. This was repeated on different days till all subjects had done all types of measurement experiments. Student subjects were also asked to do the measurements during night hours.
iii. Subjects were asked to enter the readings of resistance measurement using analog and digital devices on a laptop along with the preparation of hand written record.
iv. The measurements were conducted within a time frame as given below :
    a. 30 minutes to make fifty measurements when using analog device for measuring both length and breadth.
    b. 20 minutes to make fifty measurements when using digital device for measuring both length and breadth.
    c. 30 minutes to make fifty measurements when using analog device for measuring both resistance and voltage.
    d. 20 minutes to make fifty measurements when using digital device for measuring both resistance and voltage

e. 40 minutes to make fifty measurements when using analog device for measuring resistances and voltages, and for both entering on to a laptop and preparing hand written copy.
f. 30 minutes to make fifty measurements when using digital device for measuring resistances and voltages, and for both entering on to a laptop and preparing hand written copy.
g. Inexperienced subjects were also allowed to do the measurement in a relaxed environment without time limit to complete the fifty measurements
h. Male and female engineering students of first year and fourth year also did the measurements.
i. The experiments were done in a normal laboratory environment and air conditioned environment
j. No feedback on their performance was given to the subjects.

(Features of the components used in pattern generator, instruments and photograph of experimental set up are given in Appendix I)

3.5 RESULTS AND DISCUSSIONS

The results from the experiments were analyzed and descriptive statistics such as mean, standard deviation and coefficient of variance of error were calculated. The mean percentage error for each category was used for further analysis to remove the effect of random errors. Parametric analysis such as one sample t-test, Paired sample t-test and ANOVA were carried out to understand the effect of different factors on measurement error. The results of the analysis are discussed factor-wise in the section that follows.

3.5.1 Person Related Factors

i. Experience

Human interaction with the task at hand is primarily driven by human's knowledge. Knowledge level is both fundamental knowledge and experience related Training programs, lessons learned from reported events, and the use of industry, operating experience applied in a continuous manner, establish a minimum acceptable knowledge level and enable the development of better mental models of operation over time. It can be proved by comparing the results of experimental measurements carried out by the experienced technicians [IQ3], inexperienced subjects [IQ1, IQ2, IQ3] and first year and final year B. Tech. students [ER1, ER2, ER3]. From this given in *Table* 3.1, the mean error is seen to range from 2.0113% to 77.0453%.

Table 3.1 shows that experienced technicians in the age group of 31 to 40 are making less error than the IQ2 category and are comparable to IQ1. Experienced technicians in the age group of 41 to 50 years are better than IQ2 category.

Table 3.1 Comparison of Mean Percentage Error of ET, IQ, ER categories

Subjects	RAP	RDP	VAP	VDP
ET Age :31-40 yrs [IQ3]	6.77	5.09	16.68	3.61
ET Age :41-50 yrs [IQ3]	7.78	7.12	18.22	4.54
IQ1 Age :21-30 yrs	6.2	3.51	10.76	2.37
IQ2 Age :21-30 yrs	10.74	6.24	17.28	5.28
IQ3 Age :21-30 yrs	12.93	7.84	32.44	8.17
Final year [ER1]	5.4216	2.0113	18.8743	2.9574
Final year [ER2]	22.3763	3.4151	35.6916	5.8650
Final year [ER3]	22.1904	2.8818	57.7818	8.8006
First year [ER1]	20.1475	2.6696	27.1661	4.3999
First year [ER2]	44.4396	3.3295	47.5379	8.2129
First year [ER3]	19.8829	2.6882	77.0453	12.0510

The error reduction in the case of experienced technicians compared to inexperienced subjects with intelligent quotient [IQ3] is about 10.96% for analog devices and it is 3.66% for digital. Experienced technicians are making 15% and 27% less error than final year and first year students respectively in analog measurements. The final year B. Tech. subjects are making 12.31% and 1.24% less error than first years for analog and digital measurements respectively. It is quite evident from the studies that experience reduces error irrespective of the subjects, measurement type, technological differences and intelligent quotient of the human being.

ii. Training

Training is defined as the instruction or education, on the job or self development provided to all personnel, and units enable them to acquire the essential job skills and knowledge. Training is required to bridge the gap between the target audiences existing level of knowledge and that required to effectively operate, deploy/employ, maintain and support the system. Training is particularly crucial in the acquisition and deployment of a new system. The inexperienced subjects were initially asked to make measurements without giving any training. Then they were given training and were asked to repeat the measurements.

The *Table* 3.2 shows that training reduces error in both analog and digital measurements. The error reduction is 9.5%, 14% and 9.6% respectively for IQ1, IQ2 and IQ3 in analog measurements. In digital it is 1.6%, 2.6% and 7.7% respectively. The result is significant at 0.05 confidence level.

Table 3.2 Training

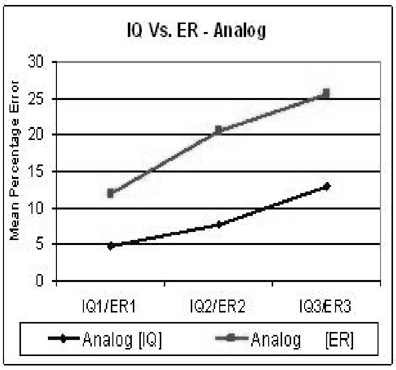
Category	Parameters	Training	Mean	t	Sig. (2-tailed)
IQ1	RAP	With training	11.03	3.272	.004
		Without training	20.49		
	RDP	With training	2.34	-2.149	.045
		Without training	3.90		
IQ2	RAP	With training	15.33	2.148	.046
		Without training	29.31		
	RDP	With training	3.89	-3.185	.005
		Without training	6.50		
IQ3	RAP	With training	16.41	3.468	.003
		Without training	25.96		
	RDP	With training	3.05	-2.174	.043
		Without training	10.75		

iii. Intelligent Quotient

The intelligent quotient of a person is indicative of his capacity to perform any task especially technical and logical. The measurement error created by subjects with intelligent quotient IQ1 (highest) is minimum in all the types of measurements followed by IQ2 and IQ3 in IQ category. In the ER (entrance test rank) category, ER1 is making least error, then ER2 and the maximum by ER3.

IQ1 category is making 3% less error than IQ2 in the case of analog measurement and 1% in digital measurement, IQ2 reduces analog measurement error by 5% compared to IQ3 and 2% less error in digital. When IQ1 is compared with IQ3, IQ3 is making 8% more error in analog measurement and 3% more in digital measurement.

ER1 is making 9% less error compared to ER2 in analog measurements and 1.23% in digital measurements; ER3 creates 5% more error compared to ER2 in analog and 1% in digital measurements where as the ER1-ER3 error difference is 14% and 2% respectively in analog and digital measurements.



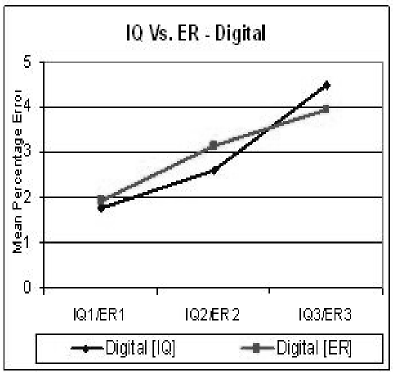


Figure 3.4 Intelligent Quotient

The comparison of inexperienced subjects [IQ] with B. Tech. students [ER] is given in *Figure* 3.4. It shows that ER1 is making 7% more error than IQ1 in analog and 0.1% more in digital measurements. ER2-IQ2 error difference is 13%

and 0.6% more in analog and digital measurements respectively. ER3 is producing 13% more error than IQ3 in analog measurements where as IQ3 is making 0.6% more error than ER3 in digital measurements. Hence it can be observed that the engineering students are making 11% more error in analog measurements and 0.05% more error in digital measurements than inexperienced subjects.

ANOVA test was done to check the significance of error difference between groups, within groups and it is found that in all the cases the error difference is significant at 0.05 confidence level.

The multiple comparison test [Post Hoc Test] results also validated the error difference among the inexperienced and student subjects. Hence it can be concluded that intelligent quotient of an individual is one of the factors which play a vital role in measurement errors. It can be safely said from the observations that people with higher IQ tend to make less errors in measurement.

iv. ***Age***

The other person related factor, age, may also have impact on measurements. To characterize error in terms of age, the measurement errors of experienced technicians with intelligent quotient [IQ3] and in the age groups of 31 to 40 years and 41 to 50 years were evaluated. Then they have also been compared with inexperienced subjects in the age group of 21 to 30 years and having different IQ levels IQ1, IQ2 and IQ3.

The mean percentage error created by the technicians in the age group of 31 to 40 years is less than that of the age group 41 to 50 years with lesser within variation than the later age group. Independent sample t-test was conducted to check whether the relation holds in the population or not. The results of the test suggest that only in the case of RDP, there is a significant difference. Hence it is concluded that the difference in the measurement error is only a sample characteristic for the case of RAP, VDP and VAP. But In the case of RDP the difference holds with $p<0.05$.

The first year and final year B. Tech. students in the ER1, ER2 and ER3 categories were asked to do the measurements. The first year students are making more error than the final years in all the cases. The error increase with first years compared to final years are 13.14%, 6.84%, 2.35% and 0.13% respectively for VAP, RAP, VDP and RDP. It can also be observed that the first years are making 10% and 1.25% more error than final year students in analog and digital measurements respectively. B. Tech students are making 11% more error than inexperienced subjects in analog measurements where as in digital it is 0.05%.

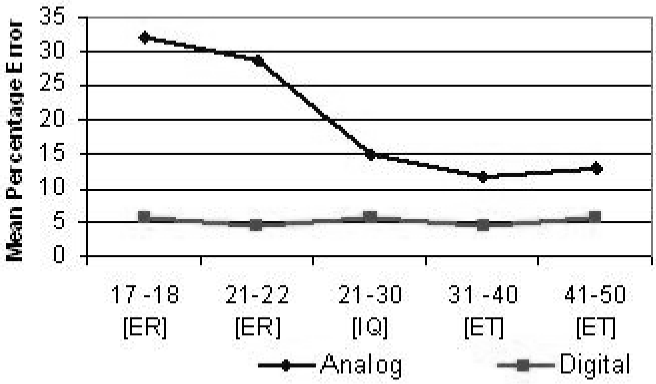


Figure 3.5 Graph Showing Age Factor – IQ, ER, ET

It can also be noticed from *Figure* 3.5 that the mean percentage error produced by Experienced Technicians in the age group of 31 to 40 years is 3.32% less compared to Inexperienced B. Tech. and Diploma holders in the age group of 21 to 30 years in analog and 1.22% in digital. Experienced Technicians in the age group of 41 to 50 years are making 1.27% and 1.48% more error compared to

Experienced Technicians in the age group of 31 to 40 years for analog and digital measurements respectively. Experienced Technicians in the age group of 41 to 50 years are making about 0.26% more error for digital measurement than inexperienced B. Tech. and Diploma holders in the age group of 21 to 30 years. This may be because of over confidence, poor vision and inability to maintain concentration for a longer time.

v. Gender

Experiments have been conducted to study whether the gender factor has any effect on measurement error. The analog and digital measurement errors of male and female B. Tech. students in ER1, ER2 and ER3 categories were studied. The female subjects in the ER1 and ER2 categories are making more errors than male subjects in RAP, ER2 and ER3 female are making more errors than male, in VAP and VDP. In all other cases male subjects are making more errors than female. But when the average was taken female is making 6.2%, 3.2% and 0.2% more errors than male, in RAP, VAP and VDP respectively. In RDP male is making 0.4% more error than female. The t-test has been conducted and it shows that the error difference between male and female is insignificant.

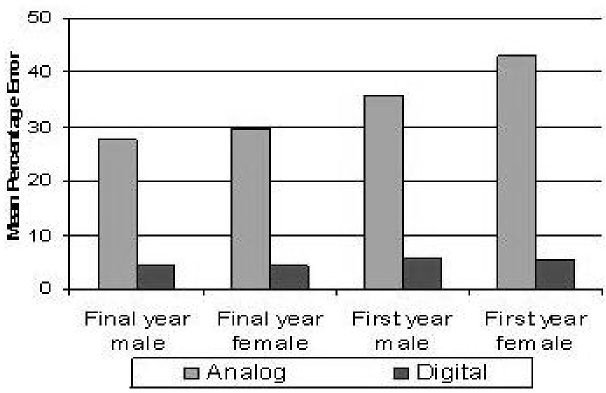


Figure 3.6. Gender Factor - Final and First Year B. Tech Students

First and final year male and female also has been compared to check the gender influence. The *Figure* 3.6 shows that final year female are making 2.1% and 0.1% more error than male in analog and digital measurements respectively, whereas the first year male subjects are generating 0.3% more error than female in digital measurements and the first year female subjects are making 7.3% more error than first year male in analog measurements. But the error difference between male and female is insignificant as per t-test. Hence it can be concluded that the factor-gender does not significantly contribute to measurement errors.

3.5.2 Work Related Factors

i. Instrument Differences

Human errors can be related to various technologies. A common element to both humans and technologies is task. Humans use technology as tools to accomplish certain task, or technologies may require human to perform tasks on them. To link human errors with technologies, the subjects with experience, without experience and having different IQ levels and B. tech. students in ER1, ER2 and ER3 categories were asked to do measurements using analog and digital devices and to note these readings using pen and paper.

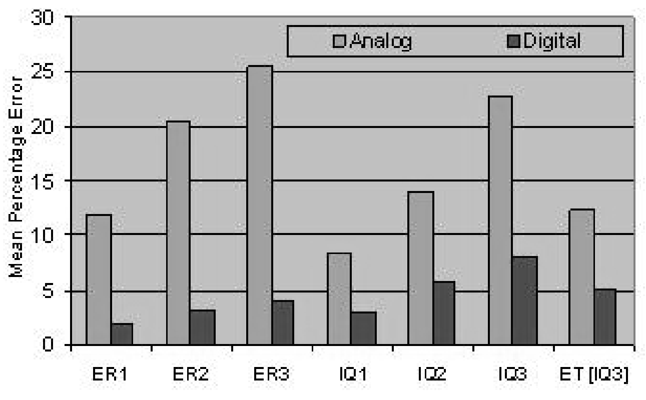


Figure 3.7. Instrument Differences - Multimeter

Error occurring when using digital multimeter for measurement is very less than analog multimeter as it is evident from *Figure* 3.7. Experienced technicians [ET] reduce measurement error by 7.3% with digital multimeter. The inexperienced subjects [average error of IQ1, IQ2 and IQ3 were taken] causes an error reduction of 9.5%, whereas the student subjects on an average reduces measurement error by 16.3% with digital multimeter. Hence the average of experienced, inexperienced and student subjects shows that measurement using digital multimeter makes an error decrease of 11% compared to analog multimeter.

Inexperienced and the student subjects of various categories were asked to measure length and breadth using analog and digital calipers. The average of digital caliper measurement error that occurred in the case of inexperienced subjects with different IQ levels IQ1, IQ2 and IQ3 is 1% less compared to analog calipers and it is 1.4% in the case of student subjects [ER1, ER2 and ER3] as is given in *Figure 3.8*. Therefore the digital caliper measurement error on an average is 1.2% less compared to analog caliper measurement error. It is observed that the average errors with digital devices are always less than analog devices irrespective of the experience, intelligence or age of the subjects.

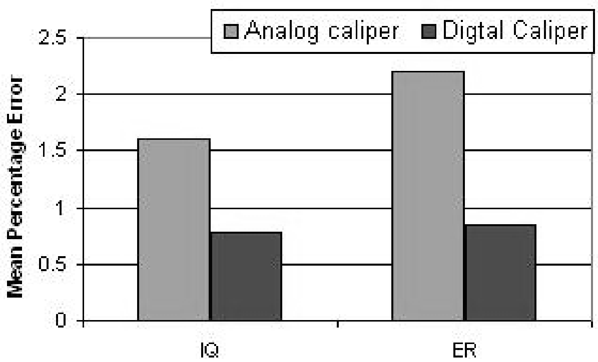


Figure 3.8. Instrument Difference – Calipers

It is also noticed that the measurement error with digital devices in the afternoon session for IQ1 category is even less than the measurement error created by them in the forenoon with analog devices. Thus it can be concluded that digital measurement devices are better than analog devices.

ANOVA test has been carried out to check the effect of technology (instrument difference). It was observed that in all the cases it is significant at 0.05 levels. Multiple comparison tests also show that Digital measurement is superior to Analog Measurement.

Measurement error observed due to technological differences can also be called as method error. Method error is a systematic error and can be found by measuring the same quantities using two methods and not getting the same results.

ii. Environment

The IQ1, IQ2 and IQ3 categories did the measurements of resistance both in normal and in better conditions.

Figure 3.9 gives that, better environment provides error reduction of 12%, 28.2% and 17.3% respectively for IQ1, IQ2 and IQ3 categories in measurements using analog devices. In the case of digital measurements it is 1.1%, 1.6% and 1.4% respectively.

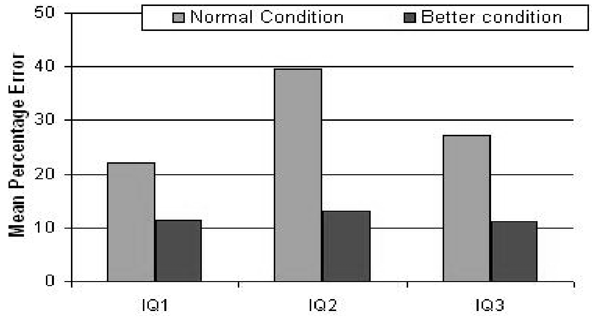


Figure 3. 9. Environment

iii. Time pressure

Inexperienced subjects in the IQ1, IQ2 and IQ3 categories were also asked to do the resistance measurement using analog and digital multimeter without any upper time limit for completing the work. *Table* 3.3 shows that with relaxed time the analog measurement causes an error reduction of 13.4%, 18.6% and 21.3% respectively where as for digital it is 1.4%, 0.6% and 0.8%.

The results of the t-test leads us to conclude that there is significant reduction in error when the time limits for doing the work was not constrained and the work environment was relaxed, in the case of analog measurement of resistance. But in digital measurement, there is a significant error difference only in the IQ1 category. Similar results are observed in the case of forenoon and afternoon sessions. Thus it can be concluded that relaxed time gives better result especially when a keen observation is involved in measurements [as in analog meter] rather than just read out [as in digital meter].

Table 3.3 Time Pressure – t-test

		Time	Mean	t	Sig. (2-tailed)
IQ1	RAP	Normal	20.4850	5.802	<.001
		Relaxed time	7.1237		
	RDP	Normal	3.7471	-2.775	.012
		Relaxed time	2.3405		
IQ2	RAP	Normal	25.8860	11.012	<.001
		Relaxed time	7.3075		
	RDP	Normal	3.6338	-.922	.369
		Relaxed time	3.0521		
IQ3	RAP	Normal	29.3066	4.113	.001
		Relaxed time	8.0431		
	RDP	Normal	4.6823	-1.191	.249
		Relaxed time	3.8865		

iv. Time of work

It is observed that subjects are making more error in the afternoon irrespective of their experience, intelligent quotient, instrument differences and the different tasks they are carrying out. It can be noticed from *Figure 3.10* that, in the case of Experienced Technicians, the error increase for analog devices in the afternoon compared to the forenoon is about 2.6% and for digital devices it is 1.3%. But for the inexperienced subjects with IQ levels IQ1, IQ2 and IQ3, the error increase is in the order of 9.74%, 15.7% and 26.6% respectively with analog devices. For the digital devices it is 1.4%, 2.3% and 5.7% respectively. It can also be noticed that even though the inexperienced B. Tech. and Diploma holders with IQ above average [IQ1] is making less error compared to the Experienced subjects, the error increase in the afternoon is very high with inexperienced subjects. This may be because of, the experienced technicians are more tuned to long and tiring working hours and the increased room temperature in tropical climate afternoons and therefore make fewer errors in adverse working conditions when compared to inexperienced subjects.

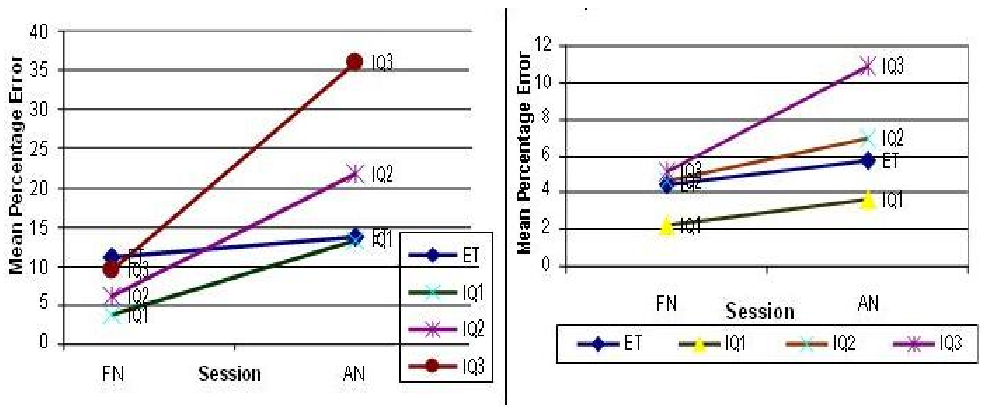


Figure 3.10. Time of Work – Experienced Technicians Vs. Inexperienced Subjects

The t-test also has been conducted to check the significance of the error difference between the forenoon and afternoon sessions for different types of measurements and found that it is significant at 0.05 levels in all the cases.

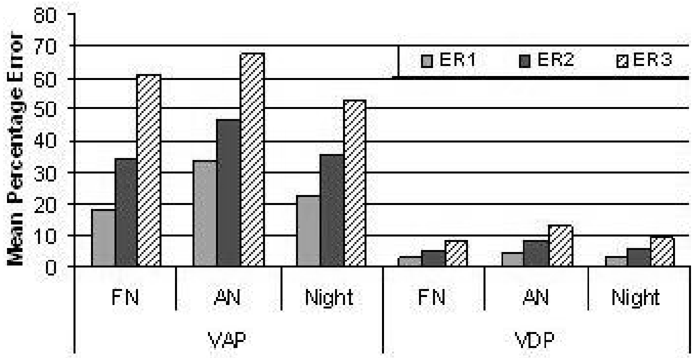


Figure. 3.11. Time of Work - B. Tech. Students [ER]

The student subjects were also asked to do the voltage measurements using Analog and Digital Multimeter during forenoon afternoon and night hours [*Figure* 3.11]. It was noticed that the error is maximum during afternoon in all the cases. ER1 is making 8.6% and 2.2% more measurement errors in the afternoon and night hours compared to forenoon. ER2 creates 7.5% and 1.12% more measurement error, each during afternoon and night compared to forenoon. But for ER3, the measurement error is minimum during night hours, 3.42% less compared to forenoon and is 6% more in afternoon compared to forenoon. The t-test also shows that the results are significant and thus it can be concluded that time of work has a significant impact on measurement errors.

Error observed in the experiments is maximum during the afternoon irrespective of the subjects, task differences and instrument differences. It is an indicative of the transient qualities of individual such as mood, motivation, degree of alertness, boredom or fatigue which affects measurements [15]. Measurement error caused by the experienced technicians with IQ, IQ3, in the afternoon is less than inexperienced subjects even though their IQ is above average. This may be because the former are conscious about their work and have a mind set about the work they have to do. But the Inexperienced and B. Tech. student subjects may have an impression that they are not supposed to do this type of work in their subconscious mind. This may be the reason why the boredom or fatigue affects more in the measurements of inexperienced and student subjects with higher intelligent quotient. Errors due to analog measuring devices are more in the afternoon than the digital measurements. This is also pointing to the transient qualities of human being. Attention on a task can only be sustained for a fairly short period of time, depending on the specifications of the task, after which fatigue sets in and errors are more likely to occur.

v. Task differences

The other work related factor is the difference in parameters to be measured. In this study the parameters measured were resistance, voltage, length and breadth using both analog and digital devices, thus there were eight tasks. It can be noticed that irrespective of the subjects experience or IQ, the errors are maximum for comparatively simple voltage measurement task as is given in *Figure* 3.12. But the error difference is not significant in IQ1 category and experienced technicians with IQ level IQ3 [*Figure* 3.12] .

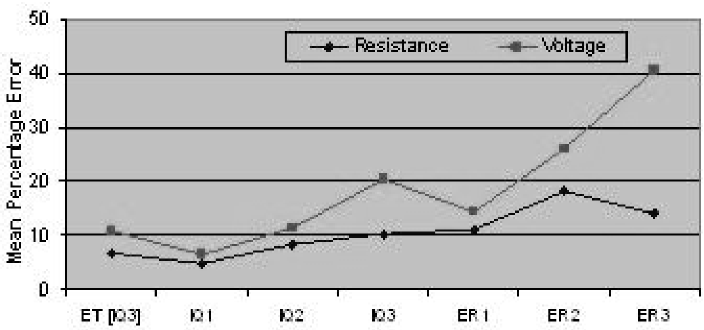


Figure 3.12 Task Differences [Resistances Vs. Voltage]

The results of paired sample test shown in *Table* 3.4 shows that there is significant change in error when doing both resistance and voltage measurements using digital and analog devices. Error difference between the measurement of resistance and voltage using digital devices is not significant at .05 levels.

An interesting observation in this experimental work is that though the voltage measurement work is simpler compared to resistance measurement, more errors are seen in voltage measurements. A possible explanation could be that when a simple task is given subjects may pay less attention to the work and thus human errors could become higher.

Table 3.4 Task Differences

Paired Samples Test				
		t	df	Sig. (2-tailed)
Pair 1	RAP - RDP	5.087	59	<.001
Pair 2	RAP - VAP	-4.224	59	<.001
Pair 3	RAP - VDP	6.863	59	<.001
Pair 8	RDP - VAP	-6.268	59	<.001
Pair 9	RDP - VDP	1.048	59	.299
Pair 14	VAP - VDP	6.155	59	<.001

The comparison of the length – breadth measurement error between the inexperienced subjects with different IQ levels and student subjects with different state entrance ranks[ER] given in *Figure* 3.13 shows that, the error rate is more with the measurement of breadth parameter using analog and digital devices. Length measurement error is 0.7% less compared to breadth. The reason for this may be because the smaller quantity measurement creates more error than a larger quantity.

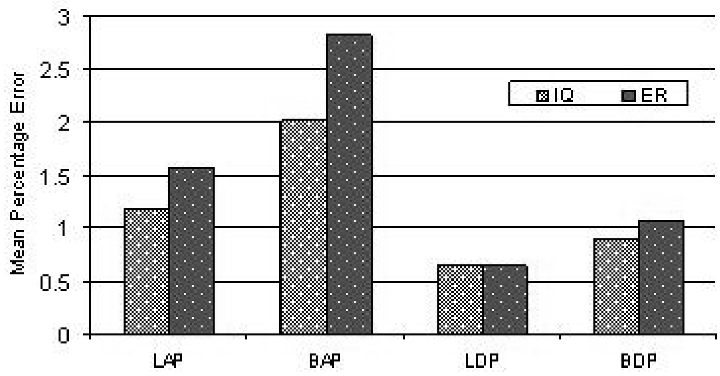


Figure 3.13. Task Differences – Length-Breadth

Since computers are now used widely to process quality related data and their use requires data entry, randomly selected subjects from ER1, ER2 and ER3 categories were asked to enter the measured resistance values into a laptop computer. The error in typing is found to be around 0.03%.

3.6 CONCLUSION

Error is an inevitable part of any kind of measurements. Measurement errors can be due to different factors. One of the important factors is the human being involved in measurement. Human error depends on many factors. In this experimental research work, an attempt has been made to identify some of these factors and characterize them as person related and work related factors. Simple experimental setups and experiments have been designed and developed to study the impact of the above factors on measurement errors. Descriptive statistics and parametric analysis such as t- test, Paired sample test and ANOVA were carried out to study the significance of each factor.

The outcome of a manual measuring system includes the results produced by the measuring device and what is observed and noted by the human subject involved in the manual measurement exercise. This study focused on the effect of a few person related factors such as experience, training, intelligent quotient, age and gender, and work related factors such as type of instruments used, time of work, task, time pressure and environmental differences when experiment was done. Using a standardized setup for the experiments, and only changing one parameter at a time, the measurement error in this study has been reduced to human errors in observation and noting. It is seen from the experiments conducted that, humans produce significant errors in measurement. The human errors due to observation and noting ranged from a minimum of 0.5 % to a maximum of 31.87%. It was necessary to study and bring out these values, since knowing the values would help in dealing with the errors when they are significant for the system. This study brings out values of human errors under different person and work related factor combinations. These may be taken as expected error values under those conditions for work system design.

In the decreasing order of influence on human errors, the person related factors are: Experience, Training, Intelligent quotient, age and gender. The ascending order of influence of work related factors on humans errors are instrument differences [digital, analog], working environment, time pressure, time of day and type of work. The practitioners who rely on measurement for research, control and production, and quality control would be benefited by the identification of the above parameters and especially the assessment of the extent of errors they produce. This work has also demonstrated a simple methodology that can be used for such work. It is hoped that this work will help users of measurement in practice to better understand and manage the phenomena of Human errors in measurement.

Chapter IV

Survey to Analyze Perception on Human Errors Among Professionals

4.1 PURPOSE OF THE SURVEY

The quality revolution, which is spreading all over the world and into all types of organization, makes error management a potential area of research. The primary objective of the study is to identify the critical factors affecting human errors in measurement, from the professional's perspective, based on an empirical analysis. Such an empirical study demands a rigorous research methodology using reliable and valid instruments. This can be achieved by measuring the perceptions of the practitioners in the industry. A questionnaire survey is widely accepted as an appropriate tool for assessing the perceptions of individuals on a particular subject. The results of this survey provided factors affecting human errors in measurement that was used in designing the experimental study presented in chapter III.

4.2 DESIGN OF THE QUESTIONNAIRE SCHEDULE AND SAMPLE SURVEY

Factors that could have possible impact on human related measurement errors were identified from the literature survey and through discussions with professionals in industry, calibration and testing laboratories and academic institutions. Details have been discussed in Chapter III.

Since there is no comprehensive instrument available to measure the critical factors of human related measurement errors from the professional's point of view, an instrument has been developed in this study. The instrument development is based on the critical dimensions of human errors identified based on literature review and discussions with the professionals in the calibration laboratories and quality control engineers in the industry. The questionnaire which has been designed has two sections:

Section A.

This section contains questions related to the impact of factors of human related measurement errors. These factors are:

i. Different types of equipments

ii. Different hours of the day

iii. Different tasks

iv. Subjects with different IQ levels and experience

v. Subjects with different age groups

vi. Subjects with different gender

Section B.

Questions in Section B were formulated to get the perception of professionals on the range of probable mean percentage human error in measurement caused by a given factor (given in Section A.)

A pilot survey was done using 25 respondents (five from each class). The instrument has been refined based on the findings of the pilot study, and based on the comments and suggestions of the experts [119]. The survey instrument that was finally developed and used consists of 37 items [30 questions and 7 subdivisions]. The complete instrument consisting all these 37 items, that epitomize the ten critical factors of human related measurement errors on measurement is presented in Appendix II. The items were jumbled and presented in a random order when given to the respondents [120]. The respondents have been asked to indicate their answers based on their perception.

4.2.1 Validity Analysis

Validity is defined as the extent to which any measuring instrument measures what is intended to measure [121]. Different validity terms are used to illustrate various aspects of construct validity [122]. The validity types that are relevant for this study are, face and content validity. The two aspects of validity, namely, face validity and content validity has been tested as explained below.

4.2.1.1 Face Validity

Generally, a measure is considered to have 'face validity', if the items are reasonably related to the perceived purpose of the measure [123]. Face validity is the subjective assessment of the correspondence between the individual items and the concept through rating by expert judges [124]. In face validity, one looks at the measure and judges whether it seems a good translation of the construct under study.

The face validity can also be established through review of the instrument by experts in the field [124]. The questionnaire used in this study has been given to five groups of experts in the area, namely, Design Engineers, Production and Control Engineers, Professionals in calibration laboratories, Quality Control Engineers and Engineering Academicians. Five each from each group were briefed about the purpose of the study and its scope. The experts were then requested to scrutinize the questionnaire and to give their impressions regarding the relevance and contents of the questionnaire. They have also been asked to critically examine the questionnaire, and to give objective feedback and suggestions with regards to comprehensiveness/coverage, redundancy level, consistency and number of items in each variable. In the initial questionnaire, there were 42 items. Based on the feedback from experts, five items were dropped, leaving 37 items in the questionnaire for the study.

4.2.1.2 Content Validity

Content validity of an instrument refers to the degree to which it provides and adequate depiction of the conceptual domain that it is designed to cover [124]. In the case of content validity, the evidence is subjective and logical, rather than statistical. Establishment of content validity warrants sound logic, good intuitive skills and high perseverance on the part of the instrument designer [123]. Content validity can be ensured if the items representing the various constructs of an instrument are substantiated by a comprehensive review of the relevant literature [125]. The instrument used in this study has been developed on the basis of a detailed literature review and consultations with experts, so as to ensure the content validity.

4.3 DATA COLLECTION AND SAMPLE

The survey population consists of professionals such as, Design Engineers, Production Engineers, Technicians, Quality Control Engineers and Academicians (the five classes of respondents) with more than two years experience. The population could not be determined hence random sampling method was not adopted. A stratified random sampling technique has been used for sample selection [126]. Twenty five respondents each were selected at random from each of the five classes (given above). The respondents were briefed about the purpose of the study and were given instructions on how to mark responses in the questionnaire before they were asked to give their responses. Survey research has been done in two phases with respondent sample selection as given below.

Phase-1:Pre-experiment Survey: Pre-experiment survey was conducted by framing 17 different questions in section A and 19 questions in section B. Answer choices were given for section 'A', but for section 'B', open ended questions were used. The questionnaire was given to ten subjects each from the five different categories of experts as is given in *Table* 4.1. Section B questions were used for getting the range of values of human related measurement errors produced by the selected factor.

Phase-2:Post-Experiment Survey: After phase-1 of the survey, the experiment described in chapter III to quantify the human related measurement errors was conducted. From the results of the experiments, multiple (three) answer choices were made, for marking answers to questions in section B. Then the second phase of the survey was carried out, using the modified set of questions, which gave answer choices to both section A and B. The numbers of respondents were also increased in the second phase by adding 15 more subjects to each of the five different categories as shown in *Table* 4.1.

Table 4.1. Respondents of the Survey

Respondents	Pre Experiment* [Without Answer Choice]	Post Experiment* [With Answer Choice]
Design Engineers	10	25
Production Control Engineers	10	25
Technicians	10	25
Quality Control Engineers	10	25
Engineering Academicians	10	25
**Experimental analysis on human errors*		

4.4 METHODS OF ANALYSIS

Survey questionnaire responses were statistically analyzed and the validity of the results were confirmed with t test and using receiver operating characteristic (ROC).

Receiver Operating Characteristic (ROC)

In signal detection theory, a **receiver operating characteristic** (**ROC**), or simply **ROC curve**, is a graphical plot which illustrates the performance of a binary classifier system as its discrimination threshold is varied [127]. It is created by plotting the fraction of true positives out of the positives (TPR = true positive rate) vs. the fraction of false positives out of the negatives (FPR = false positive rate), at various threshold settings. TPR is also known as sensitivity, and FPR is one minus the specificity or true negative rate.

The ROC is also known as a relative operating characteristic curve, because it is a comparison of two operating characteristics (TPR and FPR) [128]

A ROC space is defined by FPR and TPR as x and y axes respectively, which depicts relative trade-offs between true positive (benefits) and false positive (costs). Since TPR is equivalent with sensitivity and FPR is equal to 1 − specificity, the ROC graph is sometimes called the sensitivity vs (1 − specificity) plot.

The best possible prediction method would yield a point in the upper left corner or coordinate (0, 1) of the ROC space, representing 100% sensitivity (no false negatives) and 100% specificity (no false positives). The (0, 1) point is also called a *perfect classification*. A completely random guess would give a point along a diagonal line (the so-called *line of no-discrimination*) from the left bottom to the top right corners (regardless of the positive and negative base rates) [129].

The diagonal divides the ROC space. Points above the diagonal represent good classification results (better than random), points below the line poor results (worse than random). Note that the output of a consistently poor predictor could simply be inverted to obtain a good predictor.

The closer a result from a contingency t*able* is to the upper left corner, the better it predicts, but the distance from the random guess line in either direction is the best indicator of how much predictive power a method has. If the result is below the line (i.e. the method is worse than a random guess), all of the method's predictions must be reversed in order to utilize its power, thereby moving the result above the random guess line [130].

Area under Curve:

The area under curve (AUC), when using normalized units, is equal to the probability that a classifier will rank a randomly chosen positive instance higher than a randomly chosen negative one (assuming 'positive' ranks higher than 'negative'). It can be shown that the area under the ROC curve is closely related to the Mann–Whitney U, which tests whether positives are ranked higher than negatives. It is also equivalent to the Wilcoxon test of ranks [130]. The AUC is related to the Gini coefficient (G1) by the formula, G1=2AUC-1, where:

$$G_1 = 1 - \sum_{k=1}^{n} (X_k - X_{k-1})(Y_k + Y_{k-1}) \qquad (4.1)$$

In this way, it is possible to calculate the AUC by using an average of a number of trapezoidal approximations. The machine learning community most often uses the ROC AUC statistic for model comparison.

Other Measures:

In engineering, the area between the ROC curve and the no-discrimination line is often preferred, due to its useful mathematical properties as a non-parametric statistic. This area is often simply known as the discrimination.

4.5 RESULTS AND DISCUSSIONS

4.5.1 Pre-Experiment Survey

In the pre-experiment survey phase, ten respondents each from five different categories of subjects as given in *Table* 4.1 participated. Section A questions were used to collect the general know how about the possibilities of human related measurement errors, when they do the measurements using different devices (analog or digital), different hours of the day (FN, AN and night hours) and when they are assigned different tasks (simple, difficult, large sized and small sized objects). This set of questions also collects responses of the possibility of human related measurement errors in subjects with different IQ levels, age, experience and gender. Respondents being technical professionals, majority of them were able to give the correct answers about the human related measurement errors in the above conditions. The pre-experiment survey answers of section A question is similar to the responses of the post-experiment phase of the survey. Therefore the detailed analysis of each response to each question in section A is included in post-experiment survey analysis.

In section B, nineteen different questions were there to get the range of possible values of human related measurement error causes by selected factors. Analysis of the responses to section B questions clearly indicates that the respondents do not have a fair idea about the extent of human related measurement error causes by selected factors. This justifies the experimental study explained in chapter III.

4.5.2 Post Experiment Survey results

The analysis of responses related to Section A of the questionnaire is presented next. In the question of comparison of digital and analog devices error, 52.8% of the respondents answered that digital multimeter [DMM] makes minimum measurement error, 45.6% replied digital storage oscilloscope [DSO] makes minimum error and 1.6% thinks that analog multimeter creates minimum error, 97.6% answered that digital calipers makes minimum error compared to analog calipers. The experiments on different subjects reveal that error is minimum with digital multimeter and maximum with Analog multimeter. Similarly Digital calipers shows very less error compared to Analog calipers.

In the question of possible errors in different hours of the day, 91.2% subjects responded that, the measurement error will be minimum during forenoon, 7.2% are of the impression that the error will be minimum during night hours and 1.6% think that the error is minimum during afternoon hours. 76.8% think that the measurement errors will be maximum during afternoon hours.

The experiments show that the measurement error is minimum during forenoon and maximum during night hours. The questions regarding the effect of task differences, 96.8% says that task difference will have effect on measurement error. The 67.2 percentage of the participants thinks that measurement of smaller quantities may create more error than larger quantities. Experiment proves that, smaller quantities create more error than larger quantities.

The Questions were asked to compare the measurement error dependence on subjects' Intelligent Quotient. It is observed from the questionnaire analysis that, 84% of the respondents feels , subjects with average IQ [IQ2] make more error than subjects with above average IQ [IQ1] and 95% believes that subjects with below average IQ[IQ3] make more error than IQ1 both in analog and digital measurements. The experimental findings also confirm with this.

Questions for comparing the measurement errors of experienced technicians whose IQ is below average [ET-IQ3] with inexperienced engineers and diploma holders with above average IQ [IQ1], average IQ [IQ2] and below average IQ [IQ3] were also analyzed. 72.8% respondents feel that the experienced technicians with below average IQ [ET-IQ3] are comparable to inexperienced engineers and diploma holders with average IQ [IQ2]. 14.4% and 12.8% respectively thinks that ET-IQ3 is comparable to IQ1, and IQ3 categories. The experimental results show that ET-IQ3 is comparable to IQ2.

In the comparison of Engineering Students with State Engineering Entrance Rank between 1 and 5000 [ER1] with inexperienced engineers having different IQ level, it is observed that 56.8% feels that ER1 is comparable to IQ2, 42.4% feels that ER1 is comparable to IQ1. 68.8% equates ER2 to IQ2 and 28% equates them to IQ3. 71.2% of the respondents feels that the performance of student subjects with state entrance rank above 15000 [ER3] is comparable to IQ3 and 11.2% feels that the performance of ER3 is not comparable to either IQ1 or IQ2 or IQ3. The experimental results shows that ER1 category is comparable to IQ2 , ER2 is comparable to IQ2 and ER3 is not comparable to any of the three categories IQ1, IQ2 and IQ3 being lower.

91.2% of the respondents feel that first year students make more error than final year students and it matches experimental results. The experimental study shows that the measurement errors with experienced technicians in the age group of 41 to 50 [ET2-IQ3] is more than experienced technicians in the age group of 31 – 40 [ET1 to IQ3]. 78.4% of the respondents also agree with the above findings.

72% of the respondents said that there won't be any measurement error difference according to the gender of the subjects, 14.4% says that male creates more error and 13.6% feels that female makes more error. But in the experimental study, it is observed that female makes slightly more error than male. However, the t-test showed that the error difference between male and female is insignificant. 58% of the participants in the pre-experiment questionnaire schedule believe that, difficult task will result in more error, where as 53% of the respondents are of the impression that simple task is more error prone. Experimental study shows that there are more errors in simple tasks. The outcome of experiments on human errors in measurement is that, the possible mean error that may be created by subjects with different IQ levels, experience, subjects in different age group, gender and when the subjects were provided with different equipments and tasks were found out. Using these values three choices were included as part of the question in section B. The post experiment questionnaire analysis of the section B related questions shows that 43.76% of the respondents could give correct answers against 7.34% of pre experiment questionnaire response. The details are given in the following section.

4.5.3 Comparison of Pre and Post Questionnaire Analysis

i. Section A category questions

The pre and post experiment questionnaire analysis of questions related to section A and B categories were compared separately. The comparison of pre and post experiment questionnaire analysis of six type of questions in section A shows that the questionnaire answers are similar in both the phases of survey with a small aggregate variation of 1.76%, which is insignificant.

The ROC curve given in *Figure* 4.1 shows that there is not much difference between the pre experiment and post experiment questionnaire analysis for the questions related to section A category. The t-test result given in *Table* 4.3 confirms that the variation between the two sets is insignificant.

Table 4.2 Comparison of Pre and Post Questionnaire Analysis

Category	*Group*	*Mean*	*N*	*Std. Deviation*
Design Engineers	First set	43.0000	10	2.16025
	Second set	43.0000	10	2.10819
	Total	43.0000	20	2.07745
Production and Control Engineers	First set	42.0000	10	2.26078
	Second set	42.0000	10	2.30940
	Total	42.0000	20	2.22427
Technicians	First set	41.7000	10	1.49443
	Second set	41.4000	10	1.71270
	Total	41.5500	20	1.57196
Quality Control Engineers	First set	43.4000	10	1.71270
	Second set	43.0000	10	1.56347
	Total	43.2000	20	1.60918
Engineering Academicians	First set	41.9000	10	1.96921
	Second set	41.7000	10	1.88856
	Total	41.8000	20	1.88065
Total	First set	42.4000	50	1.97949
	Second set	42.2200	50	1.97215
	Total	42.3100	100	1.96790

Table 4.3 t-Test Showing the Comparison of Pre And Post Experiment Survey Analysis

Group		Mean	Std. Deviation	T	Df	Sig. (2-tailed)
total score	First set	42.4000	1.97949	.456	98	.650
	Second set	42.2200	1.97215			

The ROC curve is given in *Figure* 4.1. The area under the curve is found out as .474. Since the area under the curve is only .474, which indicate that there is no much difference between the two groups.

ROC Curve

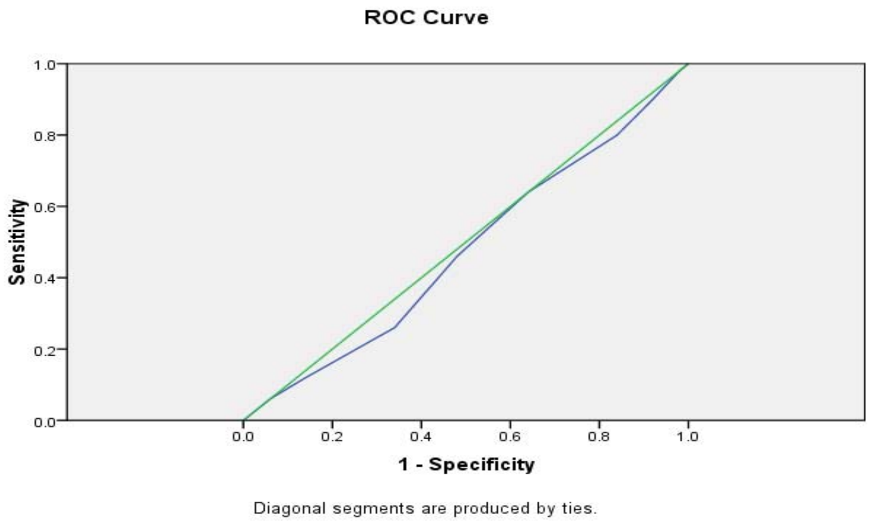


Figure 4.1 ROC Curve -Pre and Post Survey Analysis – Questions in Section A

ii. Section B category questions

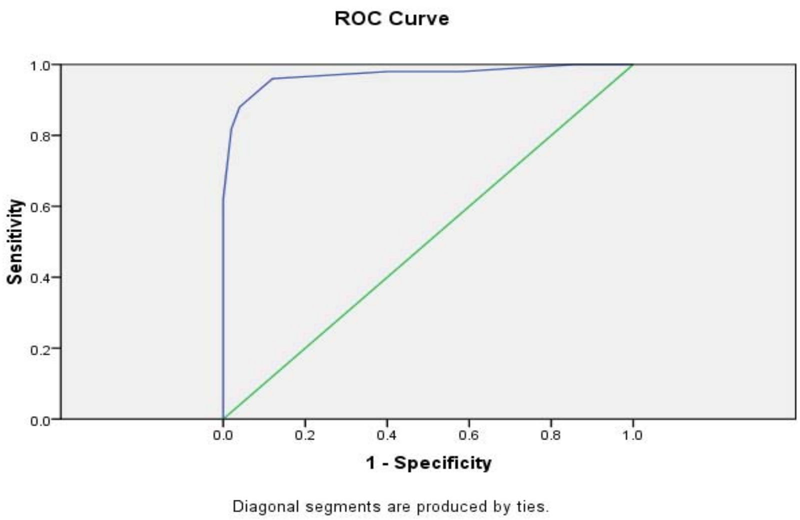
The pre and post experiment questionnaire answers related to second category of questions were compared as given in *Table* 4.4.

Table 4.4 Comparison of Pre and Post Experiment Survey Analysis

	Without choice [pre experiment]	With Choice [post experiment]
Design engineers	10.5	42
Production Control Engineers	8.5	40.8
Technicians	8.6	40
Quality Control Engineers	4.7	52.6
Engineering Academicians	4.4	43.4
Average	7.34	43.7

It can be observed from *Table* 4.4 that on an average only 7.34% could give correct answer when no choices about the possible range of measurement errors that can be caused by subjects having different experience, IQ, age, gender, in different type of measurements, in different time of day, in different environment for measurement and using different technology for measurement, were given [pre experiment survey].

The experimental study conducted after the pre experiment survey provided the possible mean percentage error values for answer choices. Then for the nineteen questions in second category, choices of range of mean percentage error were given and 25 subjects from each of the above said categories were asked to answer to the questions as schedule. It can be observed from *Table* 4.4 that, on an average 43.7% could answer the question correctly against 7.34% when choices of answers were given post-experiment survey. Quality control engineers scored maximum and the mean percentage of 19 questions is 52.6% as can be noticed from *Table* 4.4. Technicians scored minimum and the mean percentage is 40%.



Area under the Curve = 0.970

Figure 4.2 ROC Curve Showing the Significance of the Difference Between Pre and Post Survey Analysis – Questions in Section B

The ROC curve for comparing the analysis results of questionnaire related to section B category questions of pre experimental and post experimental survey is given in *Figure* 4.2. It can be observed that the area under the curve in this case is 0.97 which is close to 1, therefore, it can be concluded that the difference between the pre experiment and post experiment survey analysis of section B questions is significant. Section B questions, being to provide the perception of professionals on the extent of human related measurement errors under various circumstances, it can be said that the results derived out of the human error in measurement analysis experiment is quite useful.

4.6 CONCLUSION

In this survey research study the questionnaire has Sections A and B. Section A category, which being, questions to study the relative importance of various factors which influences the human related measurement error. It is observed that there is no significant difference between the pre-experiment and post-experiment survey responses, and that the largest fraction of the respondents, is able to give the correct answers, though they are not a large majority. This leads us to the conclusion that the factors that result in human related measurement error are understood by a large fraction of the respondents but not by a large majority. This justifies the effort to identify and study the factors and then widely share the findings.

In section B, the questions were to study the extent of human related measurement errors. There is a remarkable difference between pre and post experiment survey responses. Pre-experiment survey analysis shows that only 7.34% of the respondents could give correct answers regarding the extent of errors, whereas in post-experiment phase, on an average 43.7% could give correct answer for section B questions. Receiver operating characteristic and t-test also establishes the significance of the difference between the pre and post survey responses to the section B category questions.

The large fraction of respondents was also able to arrange factors according to their relative importance. From the pre-experiment survey research it was observed that the respondents could not give the correct answers to questions related to the correct values [extent] of human related measurement errors. This confirmed the fears expressed regarding lack of knowledge about the extent of human related measurement errors among professionals associated with quality. This is the gap that has been addressed by the experimental study presented in chapter III.

In the post experiment phase, based on the experimental analysis, the choices of answers to the questions related to the quantification of errors [Section B] were provided and the survey responses collected. This when analyzed showed that the answers to the questions dealing with extent of errors improved significantly in all categories of respondents. This shows that doing experiments and providing inputs to professionals helps improve their perception (judgment) regarding effect of different factors on extent of human errors in measurement.

Chapter V

Measurement Uncertainties in Sophisticated Instruments

5.1 INTRODUCTION

Quality of measurements has assumed great significance in view of the fact that measurements (in a broad sense) provide the very basis of all control actions. Incidentally, the word measurement should be understood to mean both a process and the output of that process. Uncertainty of measurement is a parameter, associated with the result of a measurement that characterizes the dispersion of the true values, which could reasonably be attributed to the measurand. The parameter may be, for example, the standard deviation (or a given multiple of it), or the half-width of an interval having a stated level of confidence.

Errors in the observed results of a measurement (process) give rise to uncertainty about the true value of the measurand, as is obtained (estimated) from those results [131]. Both systematic and random errors affecting the observed results (measurements) contribute to this uncertainty. These contributions have been referred to as systematic and random components of uncertainty respectively. Random errors presumably arise from unpredictable and spatial variations of influence quantities, such as, the way connections are made or the measurement method employed uncontrolled environmental conditions or their influences, inherent instability of the measuring equipment, personal judgment of the observer or operator, etc. These cannot be eliminated totally, but can be reduced by exercising appropriate controls.

Various other kinds of errors, recognized as systematic, are also observed. Some common types of these errors, are due to those reported in the calibration certificate of the reference standards /instruments used and different influence conditions at the time of measurement compared with those prevalent at the time of calibration of the standard (quite common in length and direct current measurements etc.). It should be pointed out that errors, which can be recognized as systematic and can be isolated in one case, may simply pass of as random in another case.

5.2 CALIBRATION AND UNCERTAINTY EVALUATION

Calibration determines the error associated with a measurement, and, if possible, reduces that error. This means that calibration is more than just adjusting the measurement capability of a device. Instead, the calibration process include three parts namely, verifying that the measurement capability of the measuring device is within specifications, adjusting the device to reduce its measurement error, and verifying the new measurement capability of the device to ensure that, it is operating within specifications. The most basic requirement of a calibration is proof of traceability. The traceability is defined an unbroken chain of comparisons, all having stated uncertainties between the measurement and some national or international standard.

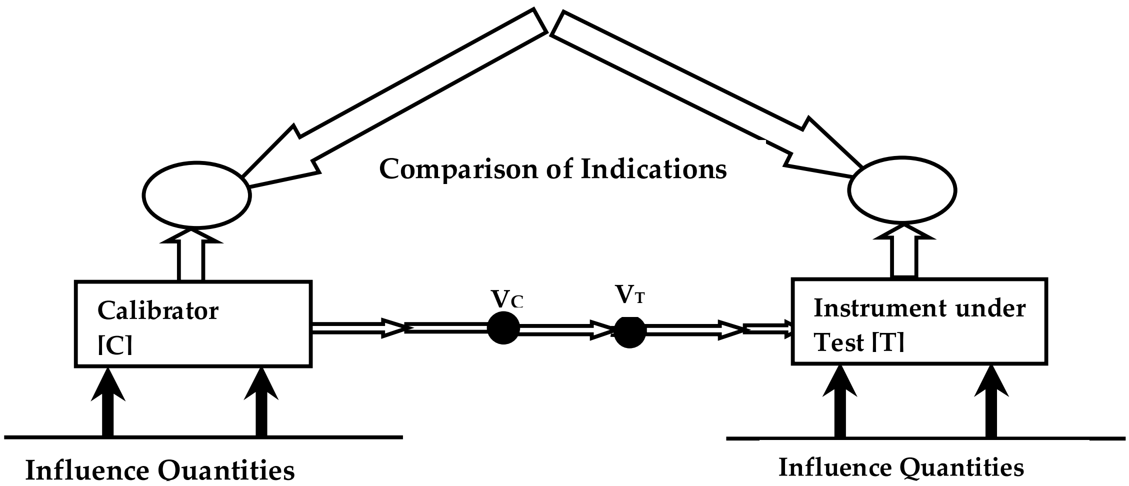


Figure 5. 1 Calibration by Comparison Method

The simplest method of calibration is comparison method. *Figure 5.1* shows the calibration by means of the comparison method, i.e. by comparison of the indication of the instrument under test, and the corresponding indication of appropriate standards. Because standards maintained by national and international bodies are well quantified and maintained, the ability to demonstrate an unbroken chain of comparisons between the measurements and these standards provides advantages such as, ability to trace the measurement uncertainty back to a known and accepted standard, acceptance of the instrument's measurement capabilities between countries and the ability to determine the maximum uncertainty associated with the measurements. Calibration certificates for measuring instruments give the measurement deviation, error correction, and the uncertainty of measurement. Only this combination characterizes the quality of the relation of the measurement result to the appropriate (SI) unit. This research work focuses on (1) characterization of time dependant variations in sophisticated instruments using calibration data (2) method for determining the uncertainty budget from calibration data (3) development of error prediction models.

5.3 CHARACTERIZATION OF CALIBRATION DATA OF SOPHISTICATED INSTRUMENTS

In order to understand the time dependant (in use) variation in the measurement errors of sophisticated instruments it was decided to study the calibration data for these from the records of accredited test and calibration laboratories. The data from Sophisticated Test and Instrument Calibration (STIC) centre, Kochi, Kerala and Fluid Control and Research Institute (FCRI), Palakkad, Kerala has been used for the study. Calibration data of a few selected sophisticated instruments, Digital Multimeter (DMM), Digital Thermometer (DT), Cathode Ray Oscilloscope (CRO), Signal Generator (SG) and Pressure gauge (PG) were analyzed for finding out the effect of ageing, instrument to instrument differences and the evaluation of the response of the instrument to different applied input for calibration. The calibration data was collected for number of instruments and time as given in *Table* 5.1

Table 5.1 Characterization of Calibration Data of Sophisticated Instruments

Name of instrument	No. of instruments	No. of years of data
Digital Multimeter	10	5
Digital Thermometer	10	5
Cathode Ray Oscilloscope	10	10
Signal Generator	10	5
Pressure Gauge	10	5

Calibration data for ten years of ten different Cathode Ray Oscilloscopes was collected and analyzed. The analysis was done with ten year data and five year data. It was seen that there is no significant difference between the two results. Therefore, the data collection was limited to five year calibration data for remaining cases. In both the test centers, calibration is done by comparison method. In comparison method, different input parameters are applied and the outputs are compared with accepted standards.

Statistical analysis was carried out in all the cases to find out the mean and standard deviation of error. The t-test and Post-hoc test were also done to check the significance of error variations.

5.4 UNCERTAINTY BUDGET FROM CALIBRATION DATA

Calibration data of digital multimeter and digital thermometer were used to make a mathematical model, to derive the uncertainties for uncertainty budget of the instrument.

5.5 MODEL FOR ERROR PREDICTION

It was found that the error prediction capability of professionals regarding the change in extent of error in measuring instruments with use (time) was poor. There is need to more accurately predict this, in order to schedule calibration depending on requirement, and not always follow thumb rule like calibrate after twelve months. To enable this using the calibration data collected it was decided to build models for error prediction. Two approaches were taken, one based on regression (time series) and the other using Artificial Neural Network (ANN) These models are discussed in detail in Chapter VI.

5.6 RESULTS AND DISCUSSIONS

Analysis of Calibration data of selected instruments such as DMM, Digital thermometer, CRO, Signal Generator and Pressure Gauge is discussed here. The calibration data of DMM and Digital thermometer are also used to determine uncertainty and hence uncertainty budget. Five year calibration data of all the five types' instruments is used to characterize the measurement error and to find out the effect of ageing, instrument to instrument differences and the error variation within different range of inputs applied for calibration of a particular instrument.

5.6.1 Theoretical Method to Determine Uncertainty Budget

5.6.1.1 Calibration of a 6 ½ digit Digital Multimeter

Calibration of a 6 ½ digit digital multimeter on its 1 Volt AC range at a nominal 0.5 V level at 1 kHz can be done using 0.5 V calibrated thermal voltage converter (TVC) . AC voltage from highly stable AC Calibrator is applied to both digital multimeter (DUC) and the standard (TVC) connected in parallel via a coaxial switch and a Tee adaptor for an indication of 0.500000 V on the digital multimeter and the electromotive force (emf) e_x indicated by the nanovoltmeter is noted. AC Calibrator is replaced by a calibrated DC calibrator and the DUC is disconnected. A DC voltage of positive polarity is applied to TVC and is adjusted so as to repeat a reading of e_x on the nanovoltmeter. The output of the DC calibrator is noted as V^+. The polarity of the DC voltage is reversed and above process is repeated and DC calibrator output voltage V- is recorded. The whole measurement process is repeated several times.

Mathematical model

The mathematical model is $V_{AC} = (V_{DC} + \Delta V_{DC} + \Delta V_{th})(1+\delta)$ (5.1)

VAC is the voltage estimated for an indicated value of 0.500000 V on DUC, VDC=V^++ V^-/2 is average of two polarity DC voltage output of calibrator, δ is AC/DC transfer correction factor of the TVC at the frequency of calibration. ΔVDC is error of DC calibrator due to its three months stability from the manufacturer's data, as the calibrator was calibrated three months before and ΔVth is error due to thermal emf which comes from the fact that the polarity of DC voltage is reversed. This is very small as compared to 0.5 V, being nearly equal to 1 μV and can be neglected. So, finally the equation becomes

VAC = VDC + ΔVDC + δVDC + δΔVDC (5.2)

The product $\delta \times \Delta VDC$ is extremely small since δ=0.008 and is neglected. The assumption is that the drift in the values of δ is small and is also neglected and error of digital multimeter in 1 volt range due to ± 1 count is ± 1μV and is also neglected. The precaution is that the interconnecting leads are coaxial shielded and are kept very small. The reference plane of measurement (mid point of Tee) is brought close to input plane of digital multimeter (DUC). These precautions minimize the loading as well as transmission errors. At frequencies up to 10 kHz, with above precautions taken, the error contributions by above factors are very low (≤ 2 to 3×10^{-6}) and can be neglected.

The inputs are:

1. The DC calibrator is regularly calibrated at intervals of six months. For the range of 1V, the uncertainty in the calibrator from its calibration certificate is a1=$\pm$ 5.8×10^{-6} at 95 % confidence level. Three months' stability data from the manufacturer's specifications is 5.0×10^{-6}.

2. The AC /DC transfer correction factor for the thermal converter is +0.000008. The AC/DC transfer uncertainty is ± 0.01 at 95% confidence level.

The observations are average of two polarity DC voltages. *Table* 5.2 gives the experimental observations.

Table 5.2 Calibration Results of DMM

Serial Number	Readings (V)
1	0.499986
2	0.499982
3	0.499991
4	0.499994
5	0.499993

Uncertainty evaluation

$$VAC = VDC + \Delta VDC + \delta VDC \qquad (5.3)$$

For uncorrelated input quantities, the combined standard uncertainty is

$$u_c^2 = \sum_{i=1}^{n} \left[\frac{\delta f}{\delta x_i} \right]^2 u^2(x_i) \tag{5.4}$$

The components of total measurement uncertainty comprise of

u1 (V) = DC calibrator's applied voltage uncertainty as mentioned in its calibration certificate

u2 (V) = DC calibrator's uncertainty due to its stability

u3 (V) = Uncertainty in the AC/DC transfer and

u4 (V) = Uncertainty due to repeatability and the corresponding sensitivity coefficients are

$$c_1 = \frac{\delta V_{AC}}{\delta V_{DC}} = 1, c_2 = \frac{\delta V_{AC}}{\delta \Delta V_{DC}} = 1, and, c_1 = \frac{\delta V_{AC}}{\delta \delta} = 1 \tag{5.5}$$

Type A evaluation

Mean DC Voltage = 0.499989 V,

Standard deviation SD = 0.0000005 V,

Standard deviation of mean or standard uncertainty

$$s(\bar{q}) = \frac{0.000005}{\sqrt{5}} = 2.23x10^{-6}V \tag{5.6}$$

Degrees of freedom = ∞

$$vi = 5 - 1 = 4 \tag{5.7}$$

Type B evaluation

1. The Uncertainty of DC calibrator from its calibration certificate gives that the distribution is normal and the coverage factor for 95 % confidence level is 1.96.

$$u_1(V) = \frac{a_1}{1.96} xSD = \frac{5.8}{1.96} x0.5\mu V = 1.48\mu V \tag{5.8}$$

Degrees of freedom = ∞

2. It can be observed from DC calibrator's specifications that, uncertainty due to 3 months stability data a2 = ±5.0 × 10^{-6}. For rectangular distribution, the standard uncertainty.

$$u_2(V) = \frac{a_2}{\sqrt{3}} xSD = \frac{5.0}{\sqrt{3}} x0.5\mu V = 1.44\mu V \tag{5.9}$$

Degrees of freedom = ∞

3. From AC/DC transfer at 95 % confidence level a_3 = 100 × 10^{-6}, distribution is normal and coverage factor = 1.96.

Standard uncertainty

$$u_3(V) = \frac{a_3}{1.96} xSD = \frac{100}{1.96} x0.5\mu V = 25.5\mu V \tag{5.10}$$

Degrees of freedom = ∞

Combined standard uncertainty

There is a dominant factor = 25.5 μV.

$$u_c = 25.5 + \sqrt{u_1^2 + u_2^2 + u_4^2} = 25.5 + \sqrt{(1.48)^2 + (1.44)^2 + (2.23)^2} \tag{5.11}$$

$$= 25.5 + 3.04 = 28.5 \ \mu V \tag{5.12}$$

Effective degrees of freedom (v_{eff})

$$V_{eff} = \frac{[u_c]^4}{\frac{(u_1)^4}{\infty} + \frac{(u_2)^4}{\infty} + \frac{(u_3)^4}{\infty} + \frac{(u_4)^4}{4}} = \infty \tag{5.13}$$

Expanded uncertainty

The expanded uncertainty for 95.45 % level of confidence, and for a coverage factor of k = 2, is

$$U = ku_c(V) = 2x28.5 = 57\mu V \quad (5.14)$$

Table 5.3: Uncertainty Budget:

Source of Uncertainty X_i	Estimates x_i V	Limits ± Δ xi μV	Probability Distribution - Type A or B - Factor	Standard Uncertainty u(xi) μV	Sensitivity coefficient c_i	Uncertainty Contribution u_i (y) μV	Degree of freedom v_i
u_1	0.5	2.9	Normal -Type B -1.96	1.48	1.0	1.48	∞
u_2	0.0	2.5	Rectangular -Type B -√3	1.44	1.0	1.44	∞
u_3	0.0	50.0	Normal -Type B -1.96	25.5	1.0	25.5	∞
Repeatability			Normal -Type A			28.5	4
uc(Vac)						28.5	∞
Expanded uncertainty			k=2.0			57.0	∞

Outcome

The measured average AC Voltage corresponding to 0.500000 V indicated by the Digital Multimeter,

VAC = VDC [1 + δ] + ΔVDC = 0.499989 (1 + 0.000008) V ± 57 μV = 0.499993V ± 57 μV (5.15)

5.6.1.2 Calibration of Digital Thermometer

A digital thermometer with a Type K thermocouple was used to measure the temperature inside a temperature chamber. Temperature controller of the chamber was set at 500°C. Digital thermometer has a resolution of 0.1°C and type

K accuracy [1 year] of ± 0.6°C.Type K thermocouple in the thermometer is calibrated every year. The last calibration report provided an uncertainty of ± 2.0°C at confidence level of 99 %. The correction for the thermocouple at 500°C is 0.5°C.

Measurement record

When the temperature chamber indicator reached 500°C, the readings were taken after a stabilization time of half an hour. Ten measurements were taken as recorded in *Table* 5.4.

Table 5.4: Measurements Record

Measurement (i)	T in °C
1	500.1
2	500.0
3	501.1
4	499.9
5	499.9
6	500.0
7	500.1
8	500.2
9	499.9
10	500.0

The mathematical model is represented as follows:

$$T = D + \text{Correction} \tag{5.16}$$

Where

T= Temperature measured.

D = Displayed temperature of the digital thermometer.

Correction = Correction due to the digital thermometer and Type K thermocouple

Analysis of Measurement uncertainty components

Combined standard uncertainty (u_c) includes uncertainties of the repeatability of the displayed readings, the digital thermometer and the thermocouple. This is represented by the equation below:

$$u_c = \sqrt{u_1^2 + u_2^2 + u_3^2} \tag{5.17}$$

where, u_c = combined standard uncertainty in the measurement,

u_1 = standard uncertainty in the repeatability of measured readings,

u_2 = standard uncertainty in the digital thermometer,

u_3 = standard uncertainty in the thermocouple

Type A evaluation

(A) Standard uncertainty in the readings (u_1)

Table 5.4 shows the data obtained from the experiment. The mean value can be calculated from the above experimental data as given below:

$$\overline{T} = \frac{1}{10}\sum_{i=1}^{10} T_i = 500.12 \tag{5.18}$$

Where T_i is the ten measurements taken as listed in *Table* 5.4. The temperature of the chamber after taking into consideration the correction of the thermocouple is 500.5°C.

The variance is calculated as follows:

$$s^2(T_i) = \frac{1}{n-1}\sum_{i=1}^{n}(T_i - \overline{T})^2 = \frac{1}{9}(0.096) = 0.0106°\text{C}^2 \tag{5.19}$$

Standard deviation

$$s(T_i) = 0.103°\text{C} \tag{5.20}$$

Standard deviation of the mean is as follows:

$$ui = s(\overline{T}) = \frac{s(T_i)}{\sqrt{n}} = \frac{0.103°C}{\sqrt{10}} = 0.03°C \tag{5.21}$$

Thus the standard uncertainty (u1) is equal to 0.03°C.

Degrees of freedom (ν_i) = n - 1 = 10 – 1 = 9

Type B evaluation

Standard uncertainty (u_2)

From specifications, the uncertainty in the digital thermometer is ± 0.6°C. Assuming rectangular distribution, the standard uncertainty in the digital thermometer (u_2) is,

$$u_2 = \frac{0.6}{\sqrt{3}} = 0.35°\text{C} \tag{5.22}$$

Degrees of freedom (ν_i) = ∞

Standard uncertainty (u_3)

From calibration report, with a confidence level of 99 % (k = 2.58), the uncertainty in the thermocouple is ± 2.0 °C.

$$u_3 = \frac{2.0}{2.58} = 0.78°\text{C} \tag{5.23}$$

Degrees of freedom (νi) = ∞

Combined standard uncertainty

The value of the combined standard uncertainty is calculated using Equation 5.17

$$u_c = \sqrt{(0.03)^2 + (0.35)^2 + (0.78)^2} = 0.85°\text{C} \tag{5.24}$$

Effective degrees of freedom

$$v_{eff} = \frac{(0.85)^4}{\frac{(0.03)^4}{9}} = \infty \tag{5.25}$$

Expanded uncertainty

$U = k \times u_c = 2 \times 0.85 = 1.7°C$; k=coverage factor = 2 (5.26)

Apportionment of standard uncertainty

Uncertainty analysis for a measurement, called the Uncertainty Budget of the measurement, include a list of all sources of uncertainty together with the associated standard uncertainties of measurement and the methods of evaluating them. For repeated measurements the number n of observations also has to be stated. For the sake of clarity, the data relevant to this analysis is presented in the form of a table given in *Table* 5.5. In this table all quantities is referenced by a physical symbol x_i, or a short identifier. For each of them the least estimate xi, the associated standard uncertainty in measurement $u(x_i)$, the sensitivity coefficient c_i and the different uncertainty contributions $u_i(y)$ is specified. The degrees of freedom are also mentioned. The dimension of each of the quantities is also stated with the numerical values in the table.

Table 5.5 Statement of the Uncertainty Budget

Source of Uncertainty Xi	Estimates xi °C	Limits ±Δ xi °C	Probability Distribution - Type A or B - Factor	Standard Uncertainty u(xi) °C	Sensitivity coefficient ci	Uncertainty contribution ui (y) °C	Deg. of freedom νi
Digital Thermometer	0.6	0.3	Rectangular - Type B - √3	0.35	1.0	0.35	∞
Thermocouple	2.0		Normal - Type B - 2.58	0.78	1.0	0.78	∞
Repeatability			Normal - Type A - √10	0.03	1.0	0.03	9
Combined uncertainty	uc					0.85	∞
Expanded uncertainties	U		k = 2			1.7	∞

5.6.2 Characterization of Errors Observed During Calibration

Digital Multimeter is used to measure direct current voltage [DCV], Alternate Current Voltage [ACV], Direct Current [DCC] and Alternating current [ACC], resistance and capacitance. The calibrations of Digital Multimeter were carried out by applying different inputs as is given in *Table* 5.6.

The DC voltage, AC voltage, DC current, AC current, Capacitance and Resistance calibration data of Digital Multimeter were analyzed separately. The effect of age in calibration error and instrument to instrument difference were studied by analyzing five years calibration data of ten Digital Multimeters.

Table 5.6 Digital Multimeter Calibration

DC Voltage

Group (V)	Range (V)
0.4	.04, 0.2, 0.36
4	.4, 2, 3.6
40	4, 20, 36
400	40, 200, 360
1000	100, 500, 900

AC Voltage

Group (V)	Range (V)
4	.4,2,3.6
40	4, 20, 36
400	40, 200, 360
750	75, 375, 675

DC Current

Group (A)	Range (A)
.004	0.004,0.002, .0036
.4	.0004, .002, .0036
10	1, 5, 9

AC Current

Group (A)	Range (A)
.004	.0004, .002, .0036
.4	.0004, .002, .0036
10	1, 5, 9

Capacitance

Group (nF)	Range (nF)
4	.5, 2, 3.6
40	4, 20, 36
400	40, 200, 360
4000	400, 2000,2600
40000	4000, 20000, 36000

Resistance

Group (kΩ)	Range (kΩ)
.4	0.04, 0.2, 0.36,
4	0.4, 2, 3.6
40	4, 20, 36
400	40, 200, 360
4000	400, 2000, 3600
40000	4000, 20000, 36000

Table 5.7 Error Growth with Ageing

DC Voltage

Input (V)	Error growth (%)
0.4	70
4	91
40	98
400	80
1000	86

AC Voltage

Input (V)	Error growth (%)
4	30
40	89
400	78
750	71

DC Current

Input (A)	Error growth (%)
.004	37
.4	19
10	86

AC Current

Input (A)	Error growth (%)
.004	45
.4	86
10	94

Resistance

Input kΩ	Error growth (%)
0.4	85
4	99
40	97
400	76
4000	96
40000	92

Capacitance

Input (nF)	Error growth (%)
4	39
40	58
400	73
4000	91
40000	81

Mean percentage of error growth after five years is given in *Table* 5.7. It is noticed that the minimum error growth rate is with DC current and the maximum is with Resistance.

5.6.2.1 Digital Multimeter –

a. Direct Current Voltage [DCV]

Direct current voltage calibration of digital Multimeter were carried out by applying, five different groups of voltages as given in *Table* 5.6. In each group, three separate voltages were used for calibration.

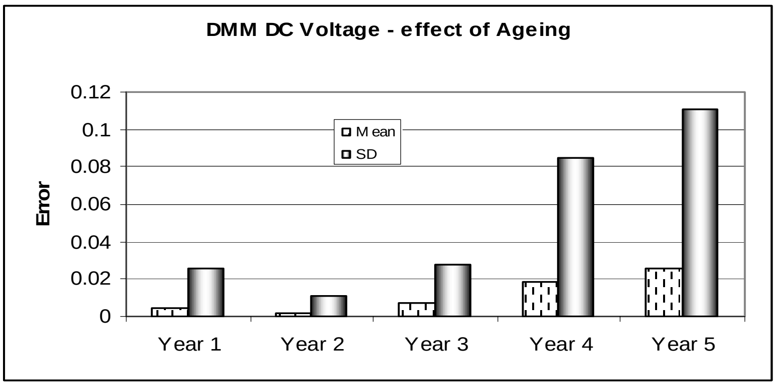


Figure 5.2 Effect of Ageing in Digital Multimeter – DCV mode

The mean error grows by 81% after five years with an average standard deviation of 0.06530 as shown in *Figure 5.2.*

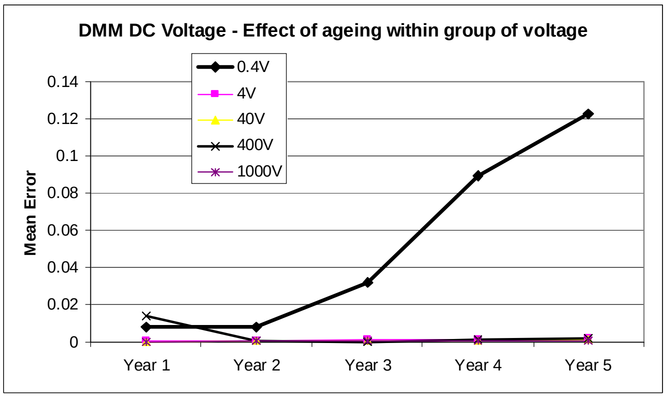


Figure 5.3 Effect of ageing within applied group of voltages for calibration - DMM DCV

Mean error growth after five years is different for different group of applied voltages for calibration. But the mean error growth is seen to be less with higher voltage group compared to smaller voltage group and is significant only in the 0.4V range used for calibration is shown in *Figure 5.3*

Table 5.8 Digital Multimeter - DC Current - ANOVA

	Sum of Squares	Df	Mean Square	F	Sig.
Between Groups	0.009	2	0.004	49.616	<0.001
Within Groups	0.020	222	0.000		
Total	0.029	224			

Table 5.8 showing ANOVA test results, confirms that there is a significant difference between the errors in the groups [five ranges of applied voltages].

b. Digital Multimeter – Alternating Current Voltage [DM – ACV]

Alternating current voltage mode of Digital Multimeter is calibrated by applying four different groups of voltages as is shown in *Table* 5.6. Mean error growth of 67% and an average standard deviation of 0.29271 are observed after five years.

Instrument to Instrument error difference for different groups of applied calibration voltage is significant only for two instruments as can be observed from *Figure* 5.4 (also confirmed by ANOVA test).

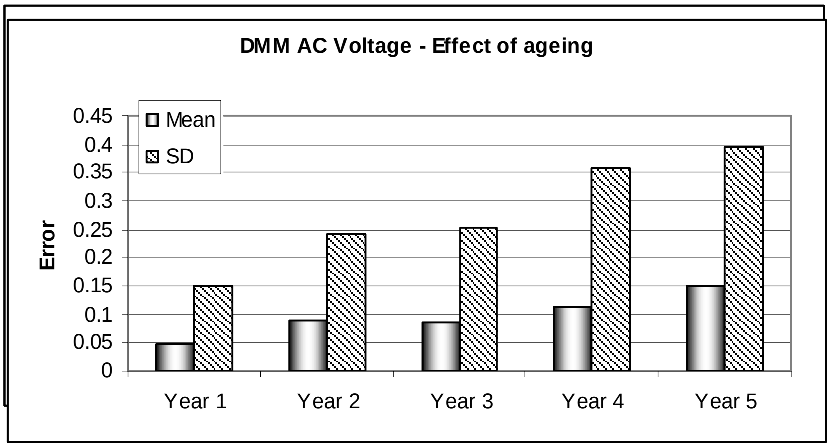


Figure 5.4 DMM AC Voltage – Instrument to Instrument Differences

Table 5.9 Multiple Comparison - AC - Voltage

(I) GROUP	(J) GROUP	Mean Difference (I-J)	Std. Error	Sig.	95% Confidence Interval Lower Bound	Upper Bound
4V	40V	-.01172	.04256	.993	-.1217	.0983
	400V	-.04218	.04256	.755	-.1522	.0678
	750V	-.32889*	.04256	*.000	-.4389	-.2189
40V	4V	.01172	.04256	.993	-.0983	.1217
	400V	-.03046	.04256	.891	-.1404	.0795
	750V	-.31717*	.04256	*.000	-.4271	-.2072
400V	4V	.04218	.04256	.755	-.0678	.1522
	40V	.03046	.04256	.891	-.0795	.1404
	750V	-.28671*	.04256	*.000	-.3967	-.1767
750V	4V	.32889*	.04256	*.000	.2189	.4389
	40V	.31717*	.04256	*.000	.2072	.4271
	400V	.28671*	.04256	*.000	.1767	.3967

* The mean difference is significant at the 0.05 level.

Error is not in the same ratio at different levels of measurement in the working range on a measuring instrument. In order to explore whether the error differences at different levels is significant, multiple comparison Tuskey HSD test was carried out. Multiple comparison Tuskey HSD test results given in *Table*

5.9 shows the significance of error difference between different groups of applied voltage. It can be observed from the *Table* 5.9 that, the higher voltage group [750V] shows a higher mean error difference when compared with other groups.

c. Digital Multimeter -DC current.

Calibration results of DC Current mode of operation of ten Digital multimeter for five years have been analyzed to study the effect of ageing and instrument to instrument differences. In direct current calibration three different groups of currents which are, 0.004A, 0.4A and 10A were used. Within each group, three different currents are used for calibration as can be noted from *Table* 5.6 and the errors are recorded. From the analysis, it is observed that the mean error grows in five years by 79% of the error noted during calibration in the first year.

Table 5.10 Digital Multimeter DC Current - Effect of ageing – ANOVA test

	Sum of Squares	Df	Mean Square	F	Sig.
Between Groups	.002	4	.001	4.101	.003
Within Groups	.027	220	.000		
Total	.029	224			

ANOVA test proves that the error growth is significant at 0.05 confidence level as can be noted from *Table* 5.10. ANOVA test also gives that the errors created by three different groups of DC current used for calibration is significant at 0.05 confidence level. But the difference in error growth between instruments is not significant.

d. Digital Multimeter-AC Current

Calibration of the AC current mode of Digital Multimeter is similar to DC current mode. It is observed that, the error observed during calibration grows after five years in AC current mode of digital Multimeter by 94% and ANOVA test also gives that the error growth is significant at 0.05 confidence level. The error difference within the three different groups [in each group three different currents were used for calibration as can be noticed from *Table* 5.6] is significant as per ANOVA test and the error growth among different instruments is also significant.

e. Digital Multimeter – Capacitance mode

Calibration of digital multimeter in capacitance mode is done by using five different capacitance groups [each group having three different capacitance values] as shown in *Table* 5.6.

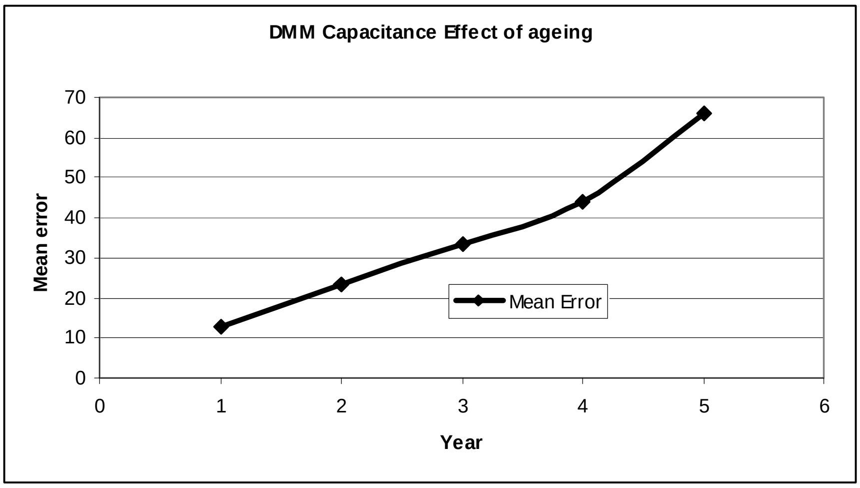


Figure 5.5 Effect of Ageing – DMM Capacitance Mode

Analysis of five years error observed during calibration shows that there is 80% error growth with ageing as it is evident from *Figure 5.5.* The error growth rate is significant only at higher ranges of capacitance values i.e. for 4000 nF and 40000 nF as can be noticed from *Figure* 5.6. It is also observed that, there is error growth difference between instruments.

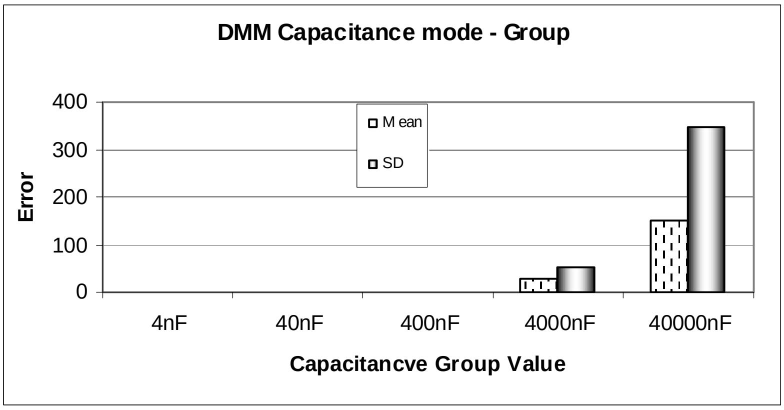


Figure 5.6 Difference in Error Between Different Applied Capacitances for Calibration

f. Digital Multimeter – Resistance Mode

Resistance mode of Digital Multimeter is calibrated by using six different groups of resistances and in each group three resistance values were used as can be seen in *Table* 5.6. The error grows at a rate of 92% after five years and the ANOVA test proves that the error growth is significant.

Instrument to instrument error difference can be analyzed using Post Hoc Test and the multiple comparisons of five equipments. It is observed that, the first equipment shows a significant error difference when compared with 2nd, 3rd and 5th equipments. Similarly, when the fourth equipment is compared with 2nd, 3rd and 5th gives a significant error difference. All other multiple comparison results stand insignificant. The error difference in six different groups of resistances used for calibration is also significant at 0.05 confidence level.

g. Error Correction Techniques

System Cabling Errors

To reduce interference, one must try to minimize the exposure of the system cabling to high-frequency RF sources. If the application is extremely sensitive to radio frequency interference (RFI) radiating from the multimeter, use of common mode chokes in the system cabling helps to attenuate multimeter emissions.

Thermal EMF Errors —

It is a good idea to take the necessary precautions to minimize thermocouple voltages and temperature variations in low-level voltage measurements. For this the best connections are formed using copper-to-copper crimped connections.

Noise Caused by Magnetic Fields —

Noise caused by magnetic fields can be reduced by routing, cabling away from magnetic fields, which are commonly present around electric motors, generators, television sets and computer monitors.

In addition, when operating near magnetic fields, one should be certain that the input wiring has proper strain relief and is tied down securely. Use of twisted-pair connections in multimeter helps to reduce the noise pickup loop area. Another way is to dress the wires as closely together as possible.

Noise Caused by Ground Loops —

The best way to eliminate ground loops is to maintain the multimeter's isolation from earth; do not connect the input terminals to ground. If the multimeter must be earth-referenced, be sure to connect it and the DUT to the same common ground point.

This will reduce or eliminate any voltage difference between the devices. Also, whenever possible, make sure the multimeter and DUT are connected to the same electrical outlet.

Loading Errors Due to Input Resistance —

To reduce the effects of loading errors, and to minimize noise pickup, set the input resistance to greater than 10GΩ for the 100 mVdc, 1 Vdc, and 10 Vdc ranges. The input resistance is maintained at 10 MΩ for the 100 Vdc and 1000 Vdc ranges. The same grounding techniques described for dc common mode problems are recommended to minimize ac common mode voltages.

AC Loading Errors —

It is desirable to use only low capacitance cable when measuring high frequency signals.

Temperature Coefficient and Overload Errors —

This additional error is automatically removed when you remove the overload condition and then change functions or ranges.

5.6.2.2 Digital Thermometer

Digital thermometers are special purpose digital volt meters. Because of the fact that temperature is industry's most measured quantity, therefore digital thermometer is popular. Instrument accuracies are consistent with the quoted resolutions; however, accuracy of the actual temperature measurement will be considerably less, because of deviations caused by the individual thermocouple [TC].

Thermocouple is calibrated by applying different temperatures -50°C, 0°C, 500°C and 950°C. Error growth with ageing is observed as 51% after five years as shown in *Figure* 5.7.

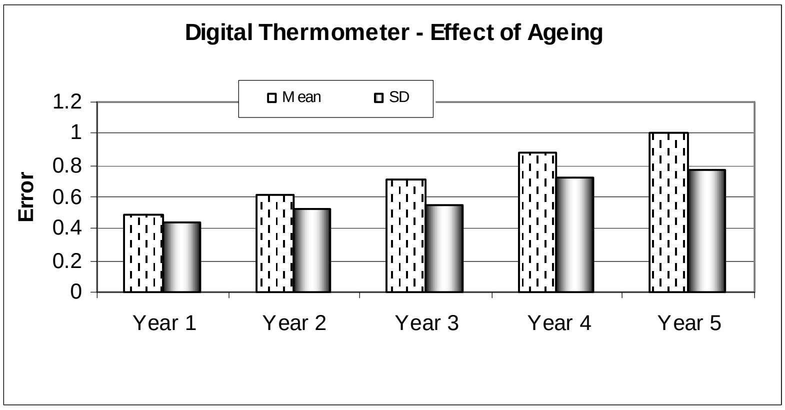


Figure 5.7 Digital Thermometer – Effect of Ageing

It is also observed that, the maximum error growth of a particular instrument is 67% and minimum is 37%. ANOVA test shows that, the instrument to instrument difference is significant. The error variation with respect to different temperature inputs is significant as is evident from *Figure 5.8.*

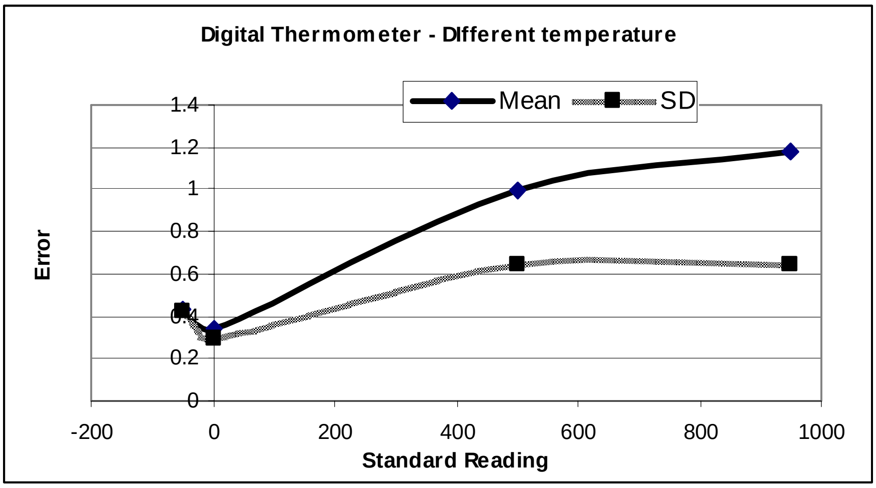


Figure 5.8 Different Temperatures – Digital Thermometer

ANOVA test also gives that, error variation when different temperatures were applied to digital thermometer for the purpose of calibration is significant at 0.05 level of confidence.

a. Error Correction Techniques

In-Homogeneity Error

Experimental tests on digital thermometer can reveal the presence of in-homogeneity, allowing one to discard defective thermocouples. Note that removing a suspected unit and carefully calibrating it in a metrology lab will not correct errors due to in-homogeneity since the lab environment does not duplicate the use environment. An in *situ* calibration, when possible, will correct such errors, but only if the "use" environment (such as immersion length, clamping forces, etc) is not changed.

The in-homogeneity errors problem can be eliminated by using thermocouples consisting of two (or more) materials. The simplest and most common situation is just two materials and such pairs are described by their relative Seebeck coefficient, which is just the difference of their absolute coefficients.

5.6.2.3 Cathode Ray Oscilloscope

a. Vertical Axis Measurements

Error growth with ageing of ten different Cathode Ray Oscilloscopes' has been analyzed. The analysis was carried out by taking ten years calibration data. Calibration of Cathode Ray Oscilloscopes was done by applying DC voltages and square wave of different ranges as is shown in *Table* 5.11.

Table 5.11 Cathode Ray Oscilloscope-Calibration Details

DC Voltage (mV/ div.)	AC Voltage (mV/div.)	Time (μs)	Rise Time at 1V
1, 2, 5, 10, 20, 50, 100, 200, 500, 1000, 2000, 5000	1 kHz Square wave with Amplitude 1, 2, 5, 10, 20, 50, 100, 200, 1000, 2000, 5000	0.005, 0.01, 0.02, 0.05, 0.1, 0.2, 0.5, 1, 2, 5, 10,20, 50, 100, 200, 500, 1mS, 2, 5, 10, 20, 50, 100, 200, 500, 1S	1 mHz ± 25

Cathode Ray-Oscilloscope -DC Voltage Calibration.

Error observed during calibration of Cathode Ray Oscilloscope, when DC voltage of different range is applied, grows with ageing and the error growth in ten years is shown in *Figure 5.9.* Mean error growth after ten years is 74.5% with the standard deviation varying from 44.9524 to 148.011. ANOVA test proves that error growth is significant at 0.05 confidence level.

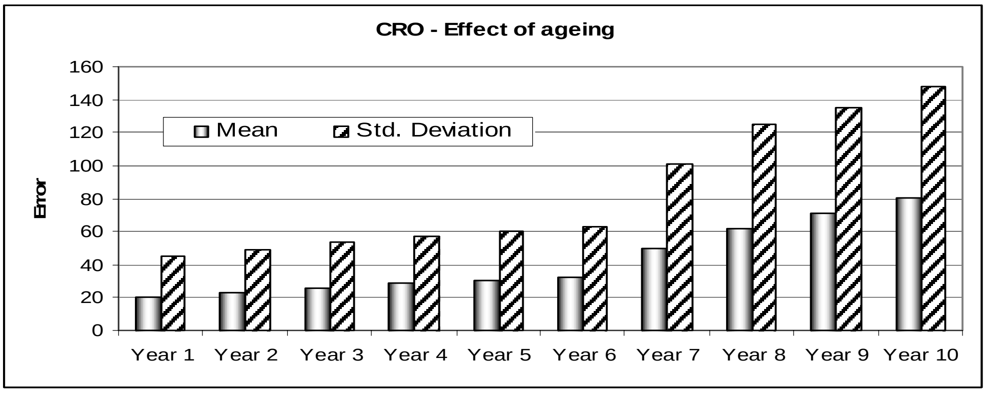


Figure 5.9 Effect of Ageing - CRO

Mean instrument to instrument error growth in ten different Cathode Ray Oscilloscope samples ranges from 33.66 [68%] to 51.74 [85%]. The standard deviation ranges from 69.51 to 127.47.

It can be observed from *Table* 5.12 that, among the different ranges of applied DC voltages for calibration, the middle range [i.e. 50 mV/div. to 1000 mV/div] shows maximum error and the upper range 2000-5000 mV/div shows minimum error.

Table 5.12 Cathode Ray Oscilloscopes-DC Voltage Calibration

Parameter Range	Mean	N	Std. Deviation	Mean of error growth after 10 years
1 mV / Div.	.30481000	100	.262490238	66%
2 mV / Div.	.28054000	100	.275918720	
5 mV / Div.	.39582000	100	.661140129	
10 mV / Div.	.38113600	100	.486291085	
20 mV / Div.	.91533000	100	.828013237	
50 mV / Div.	2.56860000	100	1.742924525	95%
100 mV/ Div.	3.91550000	100	4.902043972	
200 mV / Div.	8.67100000	100	6.219292970	
500 mV / Div.	29.20357000	100	2.121370632E1	
1000 mV /Div.	50.67192000	100	5.624519343E1	
2000 mV / Div.	1.44520000E2	100	6.881441592E1	46%
5000 mV / Div.	2.68100000E2	100	1.493107058E2	
Total	42.49401883	1200	9.358777179E1	

Cathode Ray Oscilloscope - AC Voltage calibration

AC calibration of Cathode Ray Oscilloscope were done with 1 kHz square wave signal having different voltages, ranging from 1mV to 5000 mV as can be observed in *Table* 5.12. Mean error growth after ten years is 65% and standard deviation increases from 1.465 to 3.34. Error growth is significant at 0.05 confidence level as per ANOVA test.

It is observed that, the mean instrument to instrument error growth difference in 10 different Cathode Ray Oscilloscope samples ranges from 33% to 82%. Among the different range of AC voltages applied for calibration, the lowest range, that is from 1mV to 50 mV, shows minimum error growth where as the maximum error growth is with the range staring from 1000 mV to 5000 mV as can be observed in *Table* 5.13.

Table 5.13 Cathode Ray Oscilloscope AC Voltage Calibration – for Different Voltage Ranges

PARAMETER RANGE	Mean	Std. Deviation	Mean % error growth after 10 years
1 mV / Div.	1.58114000	.777081123	31%
2 mV / Div	1.36066000	.816050469	
5 mV/ Div.	1.99418000	1.137552293	
10 mV / Div.	2.27100000	1.449164510	
20 mV / Div	2.47640000	1.247118311	
50 mV/ Div.	2.30837000	1.594374148	
100 mV / Div.	11.14330000	7.206339529	40%
200 mV / Div	11.28930000	8.299401251	
500 mV/ Div.	18.83820000	1.994349089E1	
1000 mV / Div.	1.34240000E2	7.945828205E1	82%
2000 mV / Div	4.60080000E2	2.824649323E2	
5000 mV/ Div.	7.67630000E2	3.428460400E2	
Total	1.17934379E2	2.670082769E2	

b. Horizontal Axis Measurements

Calibration of Time Measuring Parameter of Cathode Ray Oscilloscope

Time parameter of Cathode Ray Oscilloscope can be calibrated by applying constant amplitude square wave with different time period ranging from 0.005μs to 1 second as given in *Table* 5.11. Mean error growth after ten years is 90% as can be observed from *Figure 5.10*. The f*igure* also shows that the error growth is only 45% during the first three years. Error growth with ageing is significant as per ANOVA test. Mean error growth after ten years in ten different Cathode Ray Oscilloscope [instrument to instrument difference] ranges from a minimum value of 0.45725μs to a maximum value of 5.0204μs. Out of the applied signals, with different time period, the upper time period range starting from 500µs to 1 second shows the minimum error growth with ageing.

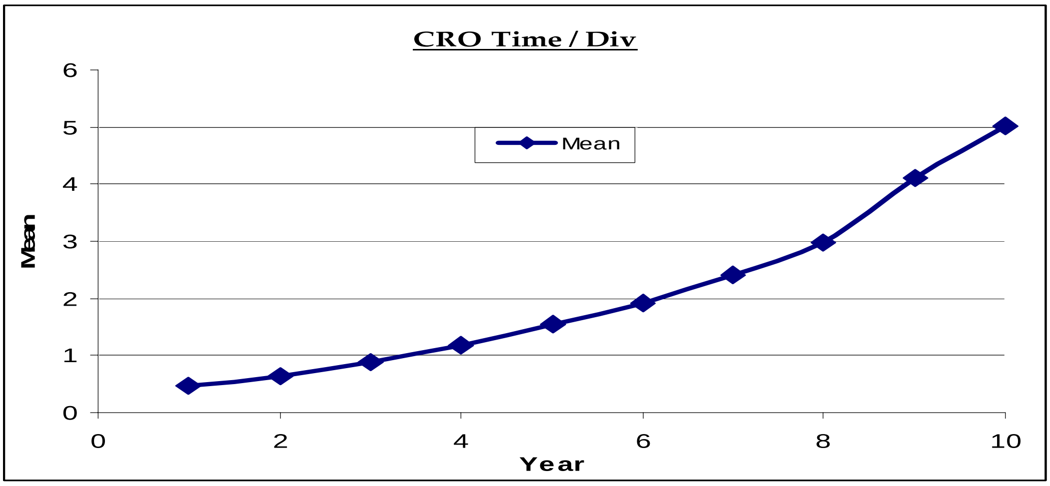


Figure 5.10 Cathode Ray Oscilloscope Time - Effect of ageing

5.6.2.4 Signal Generator

Signal Generator is calibrated by applying square wave of different frequencies. The frequencies are grouped as 0.4 kHz, 4 kHz, 40 kHz, 400 kHz and 2000 kHz. In each group, there are three frequencies, so that fifteen different frequencies are used to calibrate Signal Generator.

Calibration data of ten different Signal Generators for five consecutive years has been analyzed to study the effect of ageing, instrument to instrument difference and the variations in different applied frequencies used for calibration. In the normal calibration procedure, the unit under test [UUT] reading is compared with standard reading for different ranges of frequencies and the deviation for each UUT reading is found out.

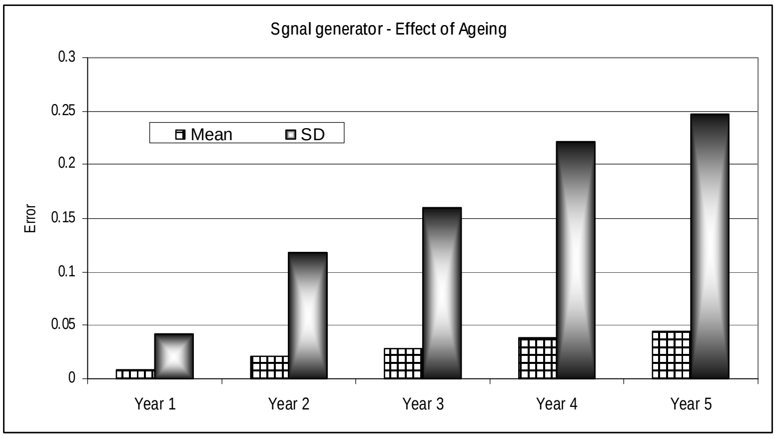


Figure 5.11 Signal Generator – Effect of Ageing

It is observed from the *Figure 5.11* that, the mean error increases by 81% after five years in a Signal Generator. Standard deviation change is from 0.04149 to 0.24681. In this study the deviation also has been found out for different frequency groups such as 0.4 kHz, 4 kHz, 40 kHz, 400 kHz and 2000 kHz. Within each group three different frequencies [eg. 0.4 kHz group, 0.39999 kHz, 0.1999992 and 0.3599986] were compared with standard reading. It is observed from *Table* 5.14 that, error grows with ageing for all the groups and also there are variations in the percentage of error growth among different applied frequencies for calibration.

Table 5.14 - Signal Generator – Difference in Error Growth with Different Applied Frequencies

Signal Generator	
Frequency range	Error growth
0.4kHz	50%
4 kHz	94%
40 kHz	89%
400 kHz	67%
2000 kHz	90%

Post Hoc test (multiple comparison test) was conducted to study the effect of error variations in different groups of frequencies. It is observed that, there is a significant difference when 0.4 kHz is compared with 2000 kHz, 4 kHz is compared with 2000 kHz, 400 kHz is compared with 2000 kHz and also in the comparison of 2000 kHz with all other frequencies. Hence, it can be concluded that the error growth difference is significant, when higher range of applied frequencies for calibration are compared with lower frequencies. Error growth difference with ageing from instrument to instrument was also studied. It is observed, that the maximum error growth is 72% and the minimum error growth is 48%.

a. Error Reduction Techniques

RF Level

Calibration of the RF level can be performed using the manufacturer's adjustment procedure and a power meter. Calibration becomes more complex as the RF level is reduced. When attenuator pads are switched in they are unlikely to be precisely their nominal value. As a result additional correction data is applied to correct for the pads. Each attenuator pad is designed to operate in a perfect 50Ω system but the reality is different. As more attenuator pads are added their mismatches interact and cause errors that are dependent on which combination of pads are in use, and the distance between each of them. These are called 'stacking' errors. It becomes harder to calibrate the attenuator. Therefore the worse the VSWR of the attenuator, the more complex the correction factors that need to be applied and the more complex the test procedure.

User problems increase when a Signal Generator has poor output VSWR but tight RF level accuracy specifications. Calibration of the generator requires increasingly complex test and correction routines to be performed with very demanding load accuracy conditions.

Attenuator Compromises

Attenuators are typically implemented using either electronic or mechanical switches. Generally speaking mechanical attenuators are likely to have better VSWR (and therefore better uncorrected accuracy), are not prone to linearity errors and will be more robust (in terms of accidentally applied reverse power). The mechanical switches can be either implemented using commercially available sealed switch assemblies or using an edge line switch structure. In general, sealed switches provide longer life, while the benefits of edge line structures are higher frequency cover, lower insertion loss and often better repeatability.

Electronic attenuators typically use either PIN diodes or FET's as electronic switches. The FET designs provide much better low frequency cover than PIN diodes, but their performance is rather less predict*able* (especially at low frequency) and they act as fast acting fuses if they are not protected from external power sources. PIN diode designs can be extended to higher frequencies and lower loss than FET's, but require complicated drive arrangements, because of the need for heavy forward current, if non-linear behavior is to be avoided.

The insertion loss of electronic attenuators is generally higher than their mechanical equivalents, and this makes it more likely, that switched high power amplifiers are required, if restrictions in output level are to be avoided (the linearity issues also make this more likely). Electronic attenuators typically have a much longer life than mechanical attenuators, have good repeatability, but are more likely to suffer changes in performance with temperature. The linearity of solid state attenuators and their loss can have a major effect on the design of the Signal Generator, especially when complex modulation schemes or combining systems are deployed.

5.6.2.5 *Pressure Gauge*

Pressure Gauge is calibrated in the testing centre by applying seven different pressures namely 60kg/cm², 100 kg/cm², 200 kg/cm², 300 kg/cm², 400 kg/cm², 500 kg/cm², and 600 kg/cm² in increasing as well as in decreasing order. It can be observed from *Figure 5.12* that there is 70% growth in error with ageing for the increasing order and 57% error growth for the decreasing order after five years.

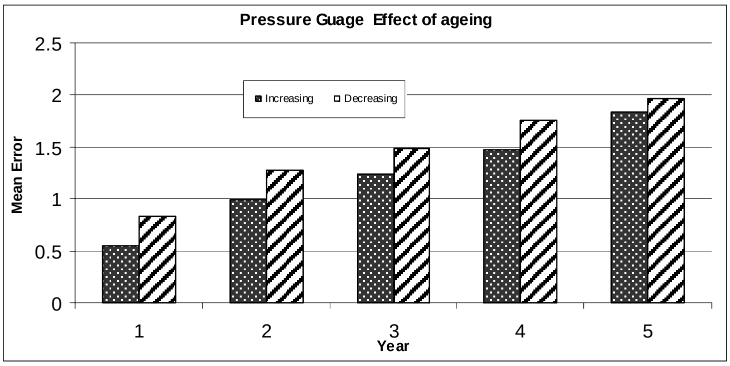


Figure 5.12 Effect of Ageing - Pressure Gauge

ANOVA test also proves the significance of error growth at 0.05 confidence level. It is observed that the error growth with ageing from instrument to instrument is significant.

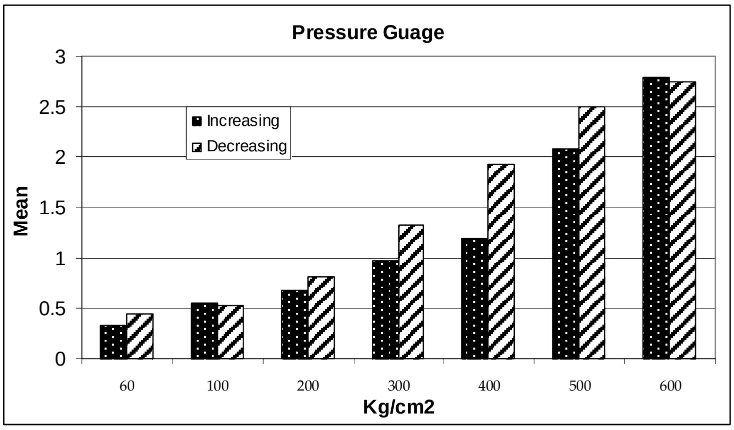


Figure 5.13: Pressure Gauge: Range of Inputs

Error growth with ageing for different input pressures for calibration is different as can be noticed from *Figure 5.13.* It can also be noticed that, the error growth is more for higher level of pressures.

5.7 CONCLUSION

Instrument calibration is the act of determining the uncertainty associated with the instrument's measurement. If possible, the instrument is adjusted to reduce the overall uncertainty associated with the measurement. Calibration quantifies the equipment's time in service, instrument to instrument differences, temperature, humidity, environmental exposure, and abuse and when required and possible, adjusts the device's measurement capability to decrease error. This study describes how the errors in selected measuring instruments change with use (over time). It also discusses how this measurement error data is used to compute the effect of the errors upon the measurement result, and how this effect is reduced or even minimized by rational use of the observations and by experimental design.

In this research work, the calibration data of Digital Multimeter and digital thermometer has been used to determine the uncertainty budget using mathematical model. Six different modes of operation of Digital Multimeter were calibrated separately. On an average there is an error growth of 82% with five years of ageing. Instrument to instrument error growth differences and the error variations in different ranges of applied voltages and currents for calibration are also significant. This shows that these factors have to be considered seriously. Error growth with ageing of digital thermometer is found as 51% after five years and the error growth rate is different for different applied input temperatures for calibration. Vertical axis calibration of Cathode Ray Oscilloscope is done by DC and AC voltages of various amplitudes. It is observed that, the mean error growth is 74.5% and a standard deviation change from 4.495 to 148.011 after 10 years for DC voltages. An error growth of 65% and standard deviation change of 146.52 to 334.047 is noticed for AC voltage calibration after 10 years. Horizontal axis calibration of Cathode Ray Oscilloscope was done by applying 1Volt square wave of different time period starting from 0.005μs to 1 second. The mean error growth is 90% after 10 years. Signal Generator is calibrated by applying different frequencies from 0.4 kHz to 100 kHz and it is found that there is an error growth of 81%. Pressure Gauge is calibrated by applying seven different pressures in the increasing and decreasing order. It is observed that, there is a 70% and 57% error growth with ageing in the increasing and decreasing order respectively of the applied pressure.

Controlling errors is an essential part of instruments and instrumentation systems. Various techniques are available to achieve this objective. Error control begins in the design stages by choosing the appropriate components, filtering, and bandwidth selection, by reducing the noise, and by eliminating the errors generated by the individual subunits of the complete system. In a good design, the errors of the previous group may be compensated adequately by the following groups.

The accuracy of instruments can be increased by post-measurement corrections. Various calibration methods may be employed to alter parameters slightly to give correct results. In many cases, calibration f*igure*s, mathematical equations, tables, the experiences of the operators, and the like are used to reduce measurement errors. In recent years, with the application of digital techniques and intelligent instruments, error corrections are made automatically by computers or the devices themselves.

Chapter VI

Errors and Uncertainties in Type Approval and Verification in Legal Metrology

6.1 INTRODUCTION

Quality of measurements has assumed great significance in view of the fact that measurements [in a broad sense] provide the very basis for trade of goods. These measurements need to be consistent. One method to ensure this is to link all measures from a common [royal] standards and this is called traceability [132]. This is more of a requirement now when trade is more international. In order to reach the objectives of legal metrology, both preventive and repressive measures are needed. Preventive measures are taken before the instruments are placed on the market or put into use and include pattern approval and verification. Market surveillance is an example of a repressive measure, and involves inspection of the instrument at the supplier's, owner's or user's premises. Competent body has to examine at least one instrument, to ensure compliance with the legal requirements. Approval tests and calibrations are carried out, and the results show whether the given requirements are met. It is particularly important to determine whether the maximum permissible errors [MPE's] at rated or foreseeable in situ operating conditions are likely to be met. Sample instrument is also subjected to quality tests which should guarantee its reliability in use.

For reasons of efficiency, verification usually only requires a single measurement (observation) to be carried out. It is therefore important that the spread or dispersion of measured values is determined during the type approval tests. This determination of so-called *apriori* characteristic values forms the justification for the evaluation of the uncertainty of measurement on the subsequent verifications [133].

European harmonization allows the manufacturer to carry out conformity assessment on new instruments as an alternative to verification by a verification body. This leads to the need to harmonize the measuring and testing methods, including determination of the measurement uncertainties and accounting for them in conformity assessments.

This study aims to analyze five years calibration data of various standards used in legal metrology such as Non-Automatic Weighing Instrument (NAWI), weight measures and volumetric measures, to study effect of ageing and instrument to instrument error difference. For the conformity study, initial type approval test and verification data of NAWIs from different manufactures were used. NAWI is selected for this purpose, because the number of applications submitted for type evaluation and approval is the largest in this category among all specified measuring instruments. Therefore the objectives of this part of research study Errors and Uncertainties in Type Approval and Verification in Legal Metrology are:

- To study the effect of ageing on measurement error in selected Working and Commercial Standards using calibration data.
- To study the limits of maximum permissible errors on verification (MPEV) and uncertainty.
- To study methods of calculating the measurement uncertainty based on the statistical interpretation on type evaluation, approval and verification specified by the Measurement Law and to propose to legal metrology, criteria with the uncertainty of measurement for deciding conformity.
- To make a comparison between calibration and verification

6.2 EFFECT OF AGEING ON MEASUREMENT ERROR

6.2.1 Maximum Permissible Errors on Verification and in Service

In many economies with developed legal metrology systems, two kinds of error limits have been defined:

- The maximum permissible errors (MPE) on verification; and
- The maximum permissible errors (MPES) in service.

The latter is normally twice the first. MPES on verification equal "MPE on testing" that are valid at the time of verification. For the measuring instrument user, the MPE in service are the error limits that are legally relevant. The values of the error limits are related to the intended use of the respective kind of instrument and determined by the state of the art of measurement technology [84].

Different instrument standards used in legal metrology are - reference standards, secondary standards, working standards and commercial standards. Verification and testing of a particular standard is done with respect to immediate superior standard, for example working standard is compared with secondary standard, which is preserved at the State Legal Metrology Laboratory [134] and commercial standard is compared with working standard during calibration. Commercial standards are the instruments used in shops and establishments and will have a direct impact on customers. Therefore it was decided to characterize measurement error observed during calibration of working standards and commercial standards as a part of this study.

6.2.2 Working Standard

Working standards in legal metrology are used as the standards to calibrate commercial standards. Maximum permissible errors taken for such working standards are maximum permissible error on service (MPES) and will depend on the type of working standards. In legal metrology laboratory, the working standards of different centers are compared with secondary standards as is shown in *Figure* 6.1.

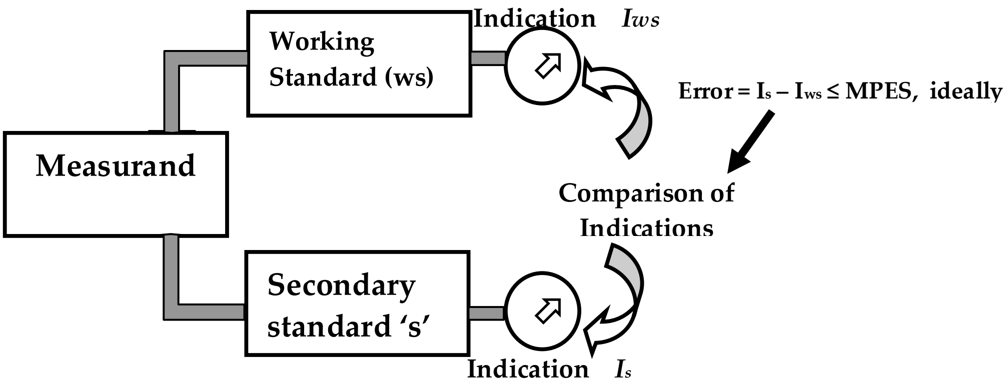


Figure 6.1. Example of Calibration by Comparison Method

Verification and test data of working standards from the legal metrology laboratory for five consecutive years was collected for the following working standards:

i. Non-Automatic Weighing Instrument(s)

ii. Weights Measures

iii. Volumetric measures

Table 6.1 shows the sample of working standards used for this study.

Table – 6.1 Samples of Working Standards

Working Standards	*Denomination of measures*	*No. of Items*
NAWI	--	20
Weight measures (in gms)	20000, 10000, 5000, 2000, 1000, 500, 200, 100., 50, 20, 10, 5, 2, 1, .500, .200, .100, .050, .020, .010, .005, .002, .001	20each
Volumetric measures (in ml)	10000, 5000, 2000, 1000, 500, 200, 100, 50, 20, 10	20each

6.2.2.1. Non-Automatic Weighing Instrument (s)

Non-automatic Weighing Instrument(s) [NAWI] used in the study is the electronic balance. Electronic load cell type balance uses the principle that a force applied to an elastic element produces a measurable deflection. The elastic elements used are specially shaped and designed. The block shapes can be cylindrical, rectangular ring proving frame, parallelogram-cut proving frame and octagonal-cut proving frame. The design aims are to obtain a linear output relationship between the applied force and the measured deflection and to make the measurement insensitive to forces which are not applied directly along the sensing axis. Displacement transducer, strain-gauge, is used to measure the deflection of the elastic elements. This instrument is to be recalibrated from time to time otherwise it will lead to significant measurement errors in the form of a bias on all readings [76].

Test procedure during verification and inspection of NAWI [134] are:

Evaluation of error

At a certain load L, the indicated value, I is noted. Additional weights of say 1/10e are successively added until the indication of the instrument is increased unambiguously by one scale interval (I+e). The additional load ΔL added to the load receptor gives indication P, by using the formula:

$P = I + \frac{1}{2} e \; \Delta L$, where *e* = readability

The error is: $E = P - L = I + \frac{1}{2} e - L - \Delta L \leq mpe$ ……………… (6.1)

<u>b. Weights</u>

Standard weights used for verification of an instrument shall not have an error greater than ⅓ of the maximum permissible error of the instrument for the applied load.

i. Weighing tests

Apply test loads from zero up to and including maximum and similarly remove the test loads back in steps to reach zero. Test loads selected shall include maximum, minimum and values at or near those at which the maximum permissible error occurs. When loading or unloading, the weights shall be progressively increased or decreased. If the instrument is provided with an automatic zero-setting device, it shall remain in operation during test. Error is calculated using the formula 6.1.

ii. Eccentricity Test

Large weights should be used in preference to several small weights. Load shall be applied centrally in the segment if several weights are used. The location of the load shall be marked on a sketch in the report. Automatic zero-setting device shall not remain in operation during the test.

In the case of instruments with a load receptor having more than four points of supports, the load shall be applied over each support on an area of the same order of magnitude as the fraction of 1/n of the surface area of the load receptor, where n is the number of points of support.

6.2.2.2. Weights

Weights starting from 20kg to 1mg, (23 different weights within the range of 20kg to 1mg) have been utilized for the study. Each of these is compared with the secondary standard weights and the difference is found out. Difference so obtained is the error. Corrections have been done if the error exceeds maximum permissible error (MPE).

6.2.2.3 Volumetric measures

Volumetric measures starting from 10 liters to 10 milliliters (10 different volumetric measures within 10 liter to 10ml) have been utilized for the study. Various measures are compared with secondary standards and the errors in each case were found out. If it exceeds MPE, corrective measures have been taken.

6.2.3 Commercial Standard

Table – 6.2 Sample of Commercial Standards

Commercial Standards	*Denomination of measures*	*No. of Items*
NAWI	--	25
Weight measures (in gms)	20000, 10000, 5000, 2000, 1000, 500, 200, 100., 50, 20, 10, 5, 2, 1, .500, .200, .100, .050, .020, .010, .005, .002, .001	25
Volumetric measures	10000, 5000, 2000, 1000, 500, 200, 100,	25

Commercial standards are calibrated with respect to working standards. Maximum permissible error for such commercial standards is in-fact maximum permissible error on service (MPES) and will depend on the type of commercial standards. MPES of commercial standards are two times compared to the MPES of working standard. In taluk legal metrology laboratory, the commercial standards of different shops and establishments are compared with working standards. Verification and test data of commercial standards for five consecutive years of the following were collected:

1. Non-Automatic Weighing Instrument(s)
2. Weights Measures
3. Volumetric measures

Calibration data of twenty five commercial standards of each category have been collected as given in *Table* 6.2.

Analysis

Five years verification and test data of working standards and commercial standards such as non-automatic weighing instruments, weights and volumetric measures were used for year-wise and instrument-wise analysis.

Descriptive statistics such as mean, standard deviation and coefficient of variations has been found out. ANOVA tests, multiple comparison test and regression analysis on the data has been carried out.

6.3 LIMITS OF MAXIMUM PERMISSIBLE ERRORS ON VERIFICATION AND UNCERTAINTY

According to the VIM [135] measurement uncertainty is a "parameter, associated with the results of a measurement, that characterizes the dispersion of the values that could reasonably be attributed to the measurand". Measurement uncertainty is usually made up of many components, some of which may be determined from the statistical distribution of the results of series of measurements and which can be characterized by experimental standard deviations [79]. The other components, which can also be characterized by standard deviations, are evaluated from assumed probability distributions based on experience or other information [136].

Contributions to the measurement uncertainty are: the standards used, the test equipment used, the measuring methods, the environmental conditions, susceptibility to interference, the state of the object to be measured or calibrated and the person performing the measurement or calibration [80]. The *Guide to the Expression of Uncertainty in Measurement (GUM) ISO* [135] and document [11] gives detailed information on the determination of measurement uncertainties.

6.3.1 Relationship Between Legally Prescribed Error Limits and Uncertainty

If a measuring instrument is tested for conformity with a given specification or with a requirement with regard to the error limits, this test used consists of comparisons of measurements with those resulting from use of a physical standard or calibrated standard instrument.

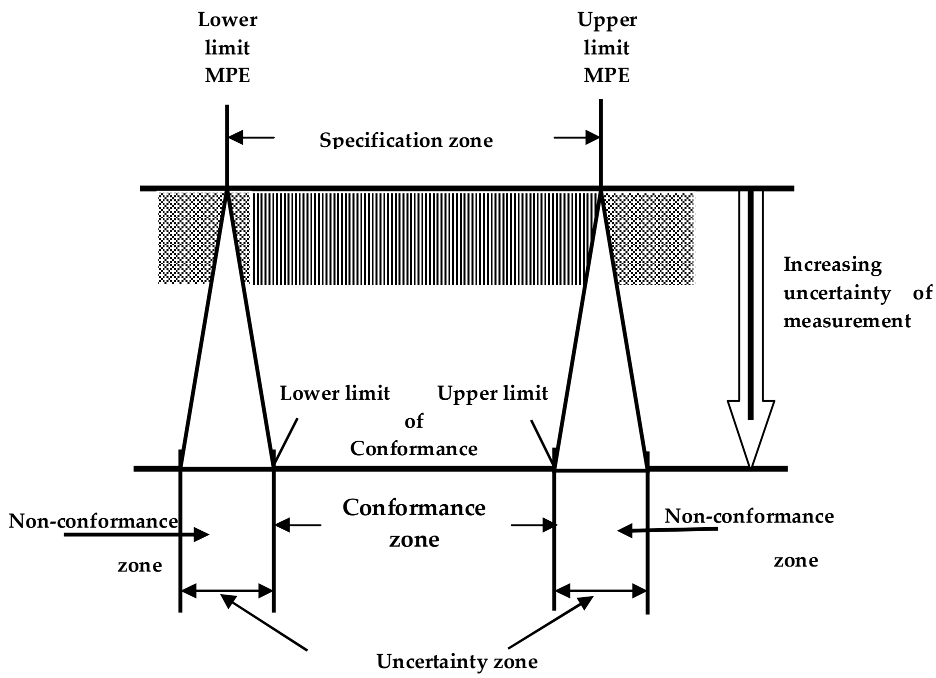


Figure 6.2. Specification And Measurement Uncertainty [According To ISO 14253-1] [91]

Uncertainty of measurement inherent in the measurement process then inevitably leads to an uncertainty of decision of conformity. *Figure* 6.2 (taken from the standard ISO 14253-1) [91] makes this problem quite clear: between the conformance zones and the upper and lower non-conformance zones there is in each case an uncertainty zone whose width corresponds approximately to twice the expanded uncertainty of measurement at the 95% probability level [86]. Uncertainty comprises contributions of the standard(s) used and the instrument under test, as well as contributions that are related to the measuring procedure, and to the incomplete knowledge about the existing environmental conditions. Because of the uncertainty of measurement, measurement results affected by measurement deviations lying within the range of the uncertainty zones cannot definitely be regarded as being, or not being, in conformity with the given tolerance requirement.

6.4 UNCERTAINTY IN TYPE APPROVAL TEST AND CONFORMITY DECISION

If a measuring instrument is tested for conformity with a given specification or with a requirement with regard to the error limits, this test consists of comparisons of measurements with those resulting from use of a physical standard or calibrated standard instrument. The uncertainty of measurement inherent in the measurement process then inevitably leads to an uncertainty of decision of conformity.

6.4.1 Test Requirements of Type Approval and Verification for Weighing Instruments

Tests of type approval and verification on NAWIs stipulated in the current Measurement Law are shown in *Table* 6.3. MPEV is specified in each test. If the measurement errors are within the specified MPEV in every test, the test result is decided as conforming. If the errors exceed MPEV, it is decided as nonconforming. Thus, only a NAWI that passes all tests will be accepted, and one that fails to pass any one of the test will be rejected.

Table 6.3. Test Requirements of Type Approval and Verification

Type Approval	*Verification*
Weighing test	Weighing test
Repeatability test with series of loading	Repeatability test with series of loading
Eccentricity test	Eccentricity test
Discrimination test	
Creep test	
Tilting error test	
Temperature characteristics test	
Zero error test just after switching on	
Instrumental error test	

There are different methods for estimating the measurement uncertainty. Two ways of estimating the measurement uncertainty are discussed as follows:

[1] Uncertainty of performance of NAWIs [Eq.6.2]

[2] Uncertainty in type approval tests on NAWIs [Eq. 6.3]

$$u = \sqrt{Vr + Vd + Vs + (Ve + Vt).w^2} \quad \text{….. … (6.2)}$$

$$u = \sqrt{Vd + Vs} \quad \text{……... (6.3)}$$

where:

u: standard uncertainty in each test

Vr: dispersion of repeatability

Vd: dispersion of rounding error

Vs: dispersion of the mass of weight

Ve: relative dispersion by eccentric load

Vt: relative dispersion by temperature characteristic

W: load on a receptor

In type approval tests, every factor of error related to the performance of NAWIs is tested one by one. Other factors are small enough to be ignored when estimating the total uncertainty under the standard condition. As shown in Equation (6.3), error factors of the uncertainty of NAWIs are the dispersion of rounding errors and that of the mass of weights.

The reason why the dispersion of rounding errors should be considered as a factor of the measurement uncertainty is because if repeating the measurement six times gave the same values, each value might fall in one scale interval but could be spread within the interval. Therefore the dispersion within one scale interval should be considered. In this respect, dispersion of rounding errors should be included in the factors of the uncertainty of the type approval test. However, if all the values measured six times are totally different, rounding errors are naturally included and they do not need to be included in the uncertainty of the type approval test. Here, assuming that all the values are almost equivalent, rounding errors should be included in the factors of uncertainty of the type approval test.

Table 6.4 Relationship Between Accuracy Class of Domestic NAWIs and Uncertainty

Accuracy class	Scale Intervals(e)	No. of scale intervals [n=max/e]	Influence of uncertainty
I	0.001gm≤e	50000≤n	Large
II	0.001gm≤e≤0.05gm	100≤n≤100000	
	0.1gm ≤ e	5000≤n≤100000	
III	0.1gm ≤ e ≤2gm	100≤n≤10000	
	5g ≤e	500≤n≤10000	
IV	5g ≤e	100≤n≤1000	small

Regarding the specified measuring instruments whose conformity is decided based on the average of measurement results or standard deviation, the dispersion of repeatability is included in the factors of the uncertainty. But in the case of NAWIs, it is not included instead a series of six times of measurement is taken and the results of each time are assessed separately. In other words, even if one of the six measurement result fails to pass the test, the instrument will be rejected as a nonconforming item. Therefore, the average or standard deviation of measurement results is not used for assessment of NAWIs. This is why the dispersion of repeatability is not included as a factor of the measurement uncertainty here.

Market surveillance was carried out focusing on the performance of the NAWIs which are practically in use. Classes of NAWIs are shown in *Table* 6.4.

Their accuracy is classified from class I to class IV. A NAWI of class I has the narrowest range of MPEV, whereas one of class IV has the largest. As to NAWIs of class III and class IV, the range of MPEV is larger, while measurement uncertainty is quite small. Therefore, measurement uncertainty does not affect conformity decision of the instruments of class III and class IV. Measurement uncertainty will greatly affect the conformity decision on NAWIs of class I and class II. Therefore NAWIs of class I and class II were taken for the conformity studies. By the voluntary co-operation of 3 companies, the instrumental error tests in type approval were carried out on thirty NAWIs of class I [15 samples] and class II [15 samples].

6.5 RESULTS AND DISCUSSIONS

6.5.1 Effect of Ageing on Measurement Error In Standards Used in Legal Metrology

Analysis results based on the calibration data of selected working and commercial standards are discussed in this section.

6.5.1.1 Working Standards

i. Non-Automatic Weighing Instrument(s)

Verification and test data of twenty different centers from the central laboratory for legal metrology was collected. Five years data of each centre was collected and analyzed. In the laboratory, three different tests such as weighing test, eccentricity test and repeatability test are done as part of the verification and test procedure of non-automatic weighing instruments [NAWIs].

a. Weighing test

Table 6.5 Comparison of Different Tests in NAWI – WS

Year		Weighing Test (mpe = 0.1gm)		Eccentricity Test (mpe=0.2mg)	Repeatability Test (mpe=0.3mg)
		Error up	Error down		
1	Mean	0.01051	0.1177	0.0053	0.0013
	SD	0.00597	0.0673	0.00389	0.0016
	CV	56.8316	57.1453	73.3962	122.3077
2	Mean	0.01536	0.1581	0.0066	0.0017
	SD	0.00849	0.0894	0.0038	0.0017
	CV	55.2539	56.5465	57.5758	101.7647
3	Mean	0.05812	0.05857	0.0078	0.0030
	SD	0.011149	0.010883	0.00463	0.0018
	CV	91.8324	85.8101	59.3590	58.6667
4	Mean	0.06833	0.07077	0.0097	0.0054
	SD	0.11090	0.11414	0.00493	0.0026
	CV	62.3065	61.2873	50.8247	47.7778
5	Mean	0.08521	0.08828	0.0116	0.0070
	SD	0.12789	0.12921	0.0044	0.0017
	CV	50.0833	46.3672	37.9310	23.8571

In this test, standard load from 5 gm to 6200gm is applied and the error is found out (up error). Then the load is removed from the receptor one by one to get the down error. The various weights that were used for calibration are 5gm, 500gm, 2000gm, 5000gm and 6200gm. It is observed from the analysis that, the error exceeds maximum permissible error (MPE), error increases with ageing and varies from instrument to instrument. The MPE on service (MPES) is 0.1gm. It can be observed from *Table* 6.5 that the error is maximum with weighing test followed by eccentricity test and repeatability test. In the weighing test the error down is 3% more than error up. Weighing tests error is 98% more compared to the eccentricity test and 99% more compared to repeatability test. Instrument to instrument error difference is evident from *Figure* 6.3.

It is seen that, there is an annual average growth (AAG) of error of 91% and 87.831% for weighing test-up and weighing test-down respectively. It can also be observed from *Table* 6.5 that there is a significant error growth with respect to year. One-way ANOVA, two way ANOVA and multiple comparison test were also conducted to confirm the significance of the error difference found. Two-way ANOVA results are given in *Table* 6.6.

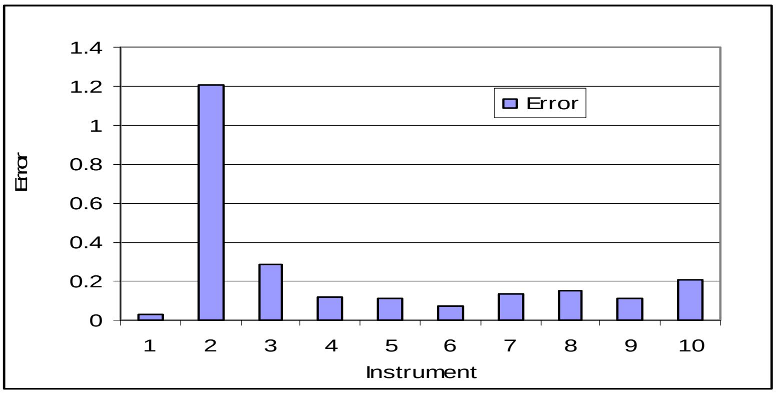


Figure 6.3 Graph Showing Instrument To Instrument Error Differences

Table 6.6 Two –way ANOVA

w-error (UP)

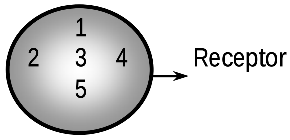
Source of Variation	DF	SS	MS	F	P
year	4	4.371	1.093	2.524	0.058
ws	19	21.486	2.387	5.514	<0.001
Residual	76	15.587	0.433		
Total	99	41.444	0.846		

w-error (DOWN)

Source of Variation	DF	SS	MS	F	P
year	4	4.596	1.149	2.54	0.056
NAWI	19	21.242	2.36	5.218	<0.001
Residual	76	16.282	0.452		
Total	99	42.12	0.86		

b. Eccentricity test

In the eccentricity test, for calibration, large weights are usually used in preference to several small weights. Same load [in this case 2000gm] is placed at different locations of the receptor of the electronic balance as shown below.



Maximum permissible error on service [MPES] as per the rule is 0.2gm. Mean eccentricity test error exceeds MPES by 4%. Eccentricity error varies from instrument to instrument [typical mean error variation of 0.0034 to 0.0158 is observed]. Annual Average Growth [AAG] is 21.9% for eccentricity test.

Eccentricity test error growth with ageing is evident from *Table* 6.5 and Two-way ANOVA test also confirms the significance of error growth. Multiple comparison tests also validate the difference between various instruments. Eccentricity test error is 55% more compared to the repeatability test error.

c. Repeatability test [r-test]

In repeatability test two series of weighing shall be performed, one with a load of about 50% and one with a load close to 100% of maximum. Readings shall be taken when the instrument is loaded, and when unloaded. Here 3000gm and 6000gm have been used. Each weight is loaded and unloaded 5 times.

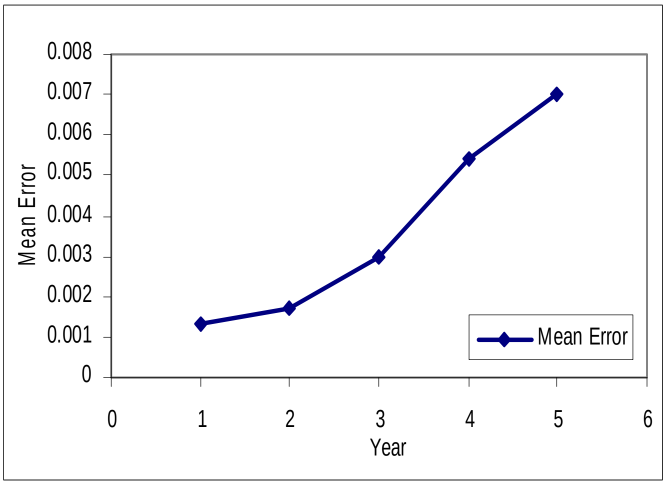


Figure 6.4 Graph Showing Error Growths with Ageing – r- Test for NAWI

Maximum permissible error on service [MPES] of repeatability test is 0.3. *Figure* 6.4 shows that error grows with ageing. There is an instrument to instrument variation. MPES prescribed by the department is maximum with repeatability test and minimum with weighing test. It is observed that the mean error increase is minimum with repeatability test.

Average annual growth [AAG] of repeatability test error is 54.81%. Two-way ANOVA test confirms the significance of error growth with ageing.

ii. Weights

Working standard of solid weights starting from 20kg to 1mg having 23 different dominations within the range was yearly subjected to verification and testing and the error were found.

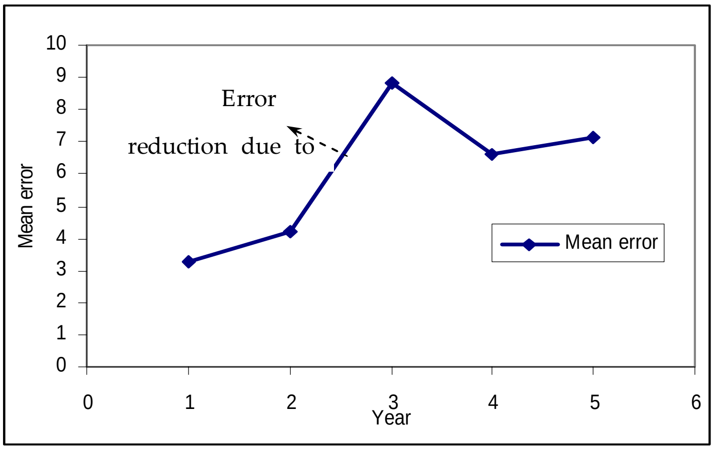


Figure6.5 Graph Showing Error Growth with Ageing – Weight measures

If the error exceeds maximum permissible error on service [MPES] ,corrective measures have been taken to keep it within MPES. Five years verification and test data of twenty different working standards, each having 23 denominations has been analyzed. It is observed that, the mean calibration error increases with year as it is evident from *Figure* 6.5. Mean error variations from working standard to working standard ranges from 2.1194 to 11.9431. It is also observed that there is an average annual average growth of 30.31% in the case of weight measures. If the observed error exceeded MPES, the same was corrected.

iii. Volumetric Measures [v-error]

In this study five years calibration data of 20 conical measures has been collected and analyzed. Conical measures are fabricated from galvanized steel sheets, aluminum alloy sheets, copper sheets, brass sheets, stainless steel sheets or tin plate. The measures are so designed that, when they are tilted 120 degree from the vertical, they become empty.

Twenty working standards of liquid measures have been taken for the study. Each working standard consists of ten different liquid measures which is capable of measuring from 10000ml to 10ml. *Table* 6.7 shows the MPES and error of a typical working standard.

Table 6.7 Typical Error with Respect to MPE for Volumetric Measures

Volume (ml)	Error (ml)	MPE(s) (ml)
10000	7	8
5000	3	4
2000	1.7	2
1000	1.2	1.5
500	0.9	1
Volume (ml)	Error (ml)	MPE(s) (ml)
200	0.65	0.8
100	0.36	0.6
50	0.45	0.5
20	0.35	0.4
10	0.18	0.2

Mean error with ageing of volumetric measures is shown in *Figure* 6. 6. Annual average growth of error is 23.82% and it is minimum compared to weight measures and non-automatic weighing instrument. One-way ANOVA confirms the significance of error increase with year.

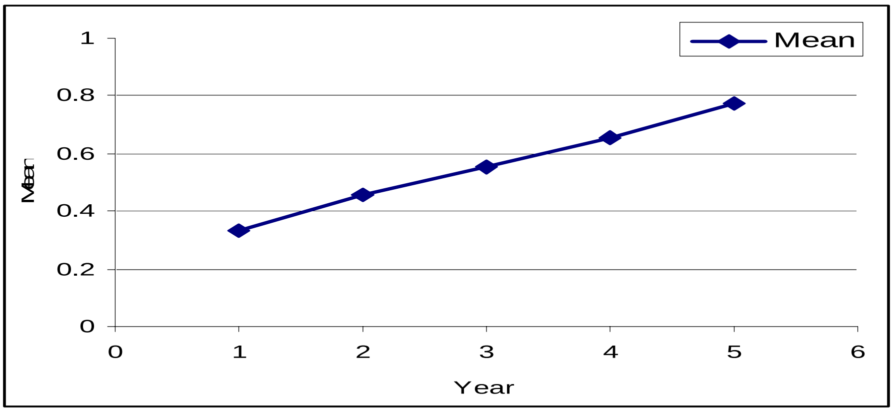


Figure6.6 Graph Showing Error Growth With Ageing- Volumetric

It is also observed that, there is a significant error difference from working standard to working standard. Mean error from instrument to instrument ranges from 0.2323 to 1.2893.

6.5.1.2 Commercial Standards

Analysis of the calibration data of selected commercial standards such as Non-Automatic Weighing Instrument (NAWI), weight measures and volumetric measures were carried out on in the same manner as that of working standards.

i. Non-Automatic Weighing Instrument(s)

Verification and test data of 25 commercial standards for five years were collected and analyzed. Three different tests weighing test, eccentricity test and repeatability test were done as part of the verification and test procedure of NAWIs.

a .Weighing test

Table 6.8 shows that, the error is maximum with weighing test followed by repeatability test and eccentricity test. Weighing test error is 98.9% more compared to the eccentricity test and 98.6% more compared to repeatability test. Between instruments there is a minimum error difference of 2% and a maximum of 15%. Maximum permissible error on service (MPES) of commercial standard NAWI is 0.2gm.

Table 6.8 Comparison of Different Tests In NAWI - CS

Year		Weighing Test (mpe = 0.1gm)		Eccentricity Test (mpe=0.2mg)	Repeatability Test (mpe=0.3mg)
		Error up	Error down		
1	**Mean**	**0.04572**	**0.03741**	**0.0066**	**0.00802**
	SD	0.01517	0.08972	0.0029	0.0034
	CV	38.56	24.72	45.01	43.33
2	**Mean**	**0.06439**	**0.04733**	**0.0098**	**0.0078**
	SD	0.02074	0.01049	0.0023	0.0026
	CV	32.887	21.4743	31.0412	34.02
3	**Mean**	**0.05829**	**0.05722**	**0.0056**	**0.0076**
	SD	0.01589	0.01203	0.00256	0.0022
	CV	26.784	20.498	48.1808	28.44
4	**Mean**	**0.05572**	**0.06845**	**0.00604**	**0.008**
	SD	0.01651	0.01477	0.00264	0.0019
	CV	29.97	21.216	44.63	24.92
5	**Mean**	**0.06278**	**0.0787**	**0.0056**	**0.0092**
	SD	0.05498	0.01556	0.0026	0.0022
	CV	25.31	19.479	47.3869	24.18

It is also observed that, there is annual average growth (AAG) of error of 21.12% and 21.5% for weighing test-up and weighing test-down respectively. Significance of error variation with respect to time and the commercial standards in different centers has been separately studied and found that, in all cases it was significant.

b. Eccentricity test

In the eccentricity test, for calibration, large weights are usually used in preference to several small weights. Same loads [in this case 2000gm] are placed on different locations of the receptor of the electronic balance. Maximum permissible error on service [MPES] is 0.4gm. Mean eccentricity test error growth after five years is 98.9%. Error observed during eccentricity test varies from instrument to instrument [typical mean error variation of 11.9% to 66.27 % was observed]. Annual average growth [AAG] of error is 33.87% for eccentricity test. Error variation with respect to year and commercial standard to commercial standard is observed as significant. Multiple comparison tests also validate the difference between various instruments. Eccentricity test error is 98.9% more compared to the weighing test error.

c. Repeatability test [r-test]

In repeatability test two series of weighing are performed, one with a load of about 50% and one with a load close to 100% of maximum. Readings are taken when the instrument is loaded, and when unloaded. Here, 3000gm and 6000gm have been used for calibration. Each weight is loaded and unloaded five times.

Maximum permissible error on service [MPES] of repeatability test is 0.6gm. There is an instrument to instrument variation with a minimum of 2.4% and a maximum of 29%. MPES prescribed by the department is maximum with repeatability test and minimum with weighing test. Mean error growth with ageing is minimum with eccentricity test. Average annual growth [AAG] of error observed during repeatability test is 20.58%. Two-way ANOVA test shows the significance of error variation. A non-automatic weighing instrument is accepted only if it shows an error less than MPES in all the above three tests.

ii. Weight Measures

Commercial standard solid weights starting from 20kg to 1mg having 23 different dominations within the range are subjected to verification and testing. Five years data regarding errors observed during test has been taken for the study.

If the error exceeds maximum permissible error on service [MPES] corrective measures are taken to keep it within MPES. Five years verification and test data of weight measures from 5 different taluk-centres , each having 23 denominations were collected for the study. From each taluk centre, 5 sets of weight standards were collected and analyzed. Mean error increase after five years is observed as, between 46.4% and 91.6% for weight standards from different taluk centers. Mean error variations from commercial standard to commercial standard ranges from a minimum of 0.2% to a maximum of 20.1%. It is also observed an average annual average error growth of 176.56. Multiple comparison and two-way ANOVA test also supports the above observations.

iii. Volumetric Measures [v-error]

In this study five years verification and test data of 25 conical measures has been collected and analyzed from 5 different taluk centers. Each standard consists of 10 different liquid measures which are capable to measure from 10000ml to 10ml.

Error growth in volumetric measuring standard after five years is shown in *Figure* 6.7. Annual average growth of error is 51.18%. Error variation with respect to instrument ranges from 2.4% to 32.3%. Significance of the observations is confirmed by two-way ANOVA test.

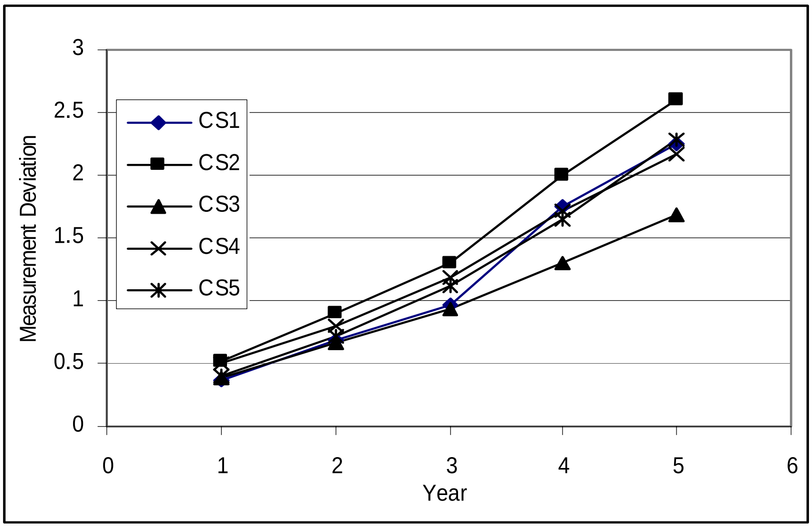


Figure 6.7 Graph Showing the Effect of Ageing – Volumetric Measure

6.5.2 Limits of Maximum Permissible Errors on Verification and Uncertainty

Measuring instruments are considered to comply with the legal requirements for error limits if:

- the absolute value of the measurement deviations is smaller than or equal to the absolute value of the legally prescribed MPES on verification when the test is performed under prescribed test conditions; and
- the expanded uncertainty of measurement of the previous quantitative metrological test, for a coverage probability of 95 %, is small compared with the legally prescribed error limits .

The expanded measurement uncertainty at the 95% probability level, $U0.95$, is usually considered to be small enough if the following relationship is fulfilled:

$$U_{0.95} \leq \frac{1}{3}.\mathrm{MPEV} \ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots \quad (6.4)$$

where *MPEV* is the absolute value of the MPE on verification.

U_{max} [$U_{0.95}$] is, therefore, the maximum acceptable value of the expanded measurement uncertainty of the quantitative test.

6.5.2.1 Relationship upon Testing of Working Standards

In legal metrology, working standards are the standards that are used routinely to verify measuring instruments. In several economies, some of the working standards used in legal metrology must be tested or verified according to special regulations. MPES of such working standards depend on their intended use. In general, they should be significantly lower than the expanded uncertainties that are required by equation (6.3).

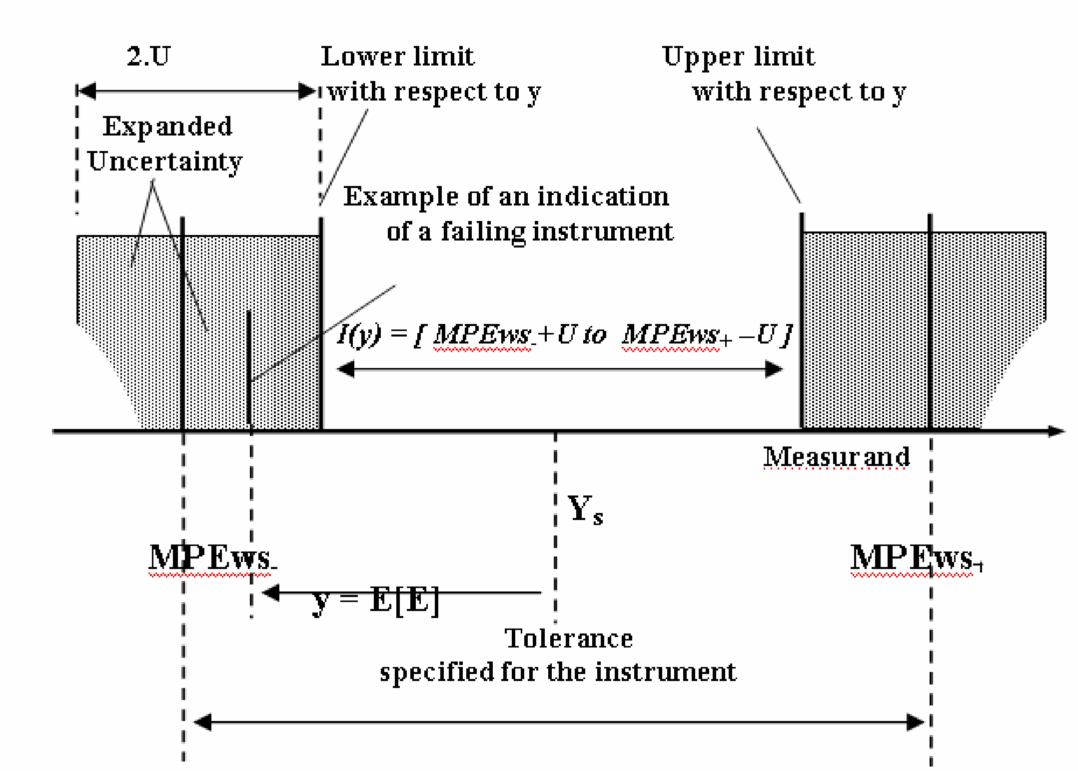


Figure 6.8 Relationship Between MPES and Measurement Uncertainty Upon Conformity Evaluation in Calibration [137] and Testing of Working Standards.

Ys : conventional true value
Y: best estimate of the measurement deviation, E
MPEws- : Lower limit of maximum permissible error of working standard
MPEws+ : Upper limit of maximum permissible error of working standard
I[y]: Acceptance interval with respect to the measurement deviation y.

Usually, a working standard, e.g. non-automatic weighing instrument is considered to comply with the respective requirements for legal error limits if the difference between its indication, or measured value, and the corresponding value realized by a reference standard is equal to or less than the difference between the prescribed error limits, *MPE*ws, and the expanded uncertainty of measurement, $U_{0.95}$:

$|I_{ws} - y_s| \leq MPE_{ws} - U_{0.95}$ …………………………………. (6.5)

where:

I_{ws} = the indication of the working standard under test;

and

y_s = the value provided by a reference standard.

In practice, this means that with respect to measurement deviations, a tolerance band is defined that is significantly reduced when compared with the range between the legally prescribed error limits [MPE_{ws-} to MPE_{ws+}] (see *Figure* 6.8). The magnitude of this tolerance band may be described by the interval [$MPE_{ws-} + U$ to $MPE_{ws+} - U$].

This approach is consistent with the prescribed procedures for statements of conformity on calibration certificates.

6.5.2.2 Uncertainty Contribution of Verified Instruments

In practice, it is often desirable or necessary to determine the uncertainty of measurements that are carried out by means of legally verified measuring instruments. If only the positive statement of conformity with the legal requirements is known, for example in the case of verified instruments without a certificate, the uncertainty of measurements for such instruments can be derived only from the information available about the prescribed error limits (on verification and in service) and about the related uncertainty budgets according to the requirements established by equation 6.4 and section 6.6.2 earlier.

On the assumption that no further information is available, according to the principle of maximum entropy, the following treatment is justified:

- The range of values between the MPES on verification can be assumed to be equally probable.

- Due to uncertainty in measurement, the probability that indications of verified instruments are actually beyond the acceptance limits of the respective verification, declines in proportion to the increase in distance from these limits. A trapezoidal probability distribution according to *Figure* 6.9 can, therefore, reflect adequately the probable dispersion of the deviation of verified measuring instruments.
- Immediately after verification, the indications of measuring instruments may exceed the MPES on verification by the maximum value of the expanded uncertainty of measurements at most.
- After prolonged use and under varying environmental conditions, it can be assumed that the expanded measurement uncertainty, compared with its initial value, may have increased significantly.

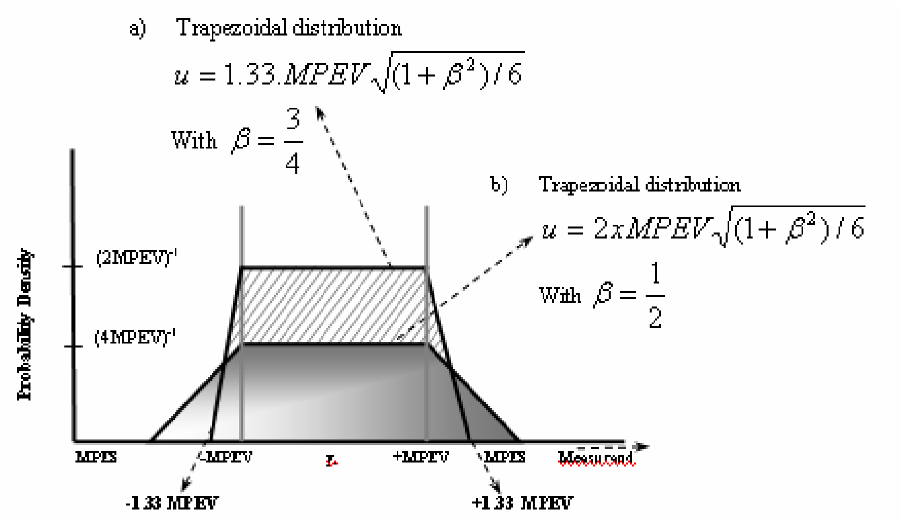


Figure6.9 Suggested Probability Distributions for Evaluating the Standard Uncertainty Contribution of Verified Measuring Instruments.

a : immediately after verification
b : after prolonged use
MPEV : absolute value of the maximum permissible error on verification
MPES : absolute value of the maximum permissible error in service
Yi, m : indication of the verified instrument when using for measurements

In particular, the following evaluation of the uncertainty contribution of verified instruments seems to be appropriate:

(i) Immediately after verification, the trapezoidal probability distribution of the errors [138][139] according to plot (a) of *Figure 6.9* can be taken as a basis for the determination of the uncertainty contribution of the instruments.

The following may, therefore, be assumed for this standard uncertainty contribution U_{INSTR} :

$$U_{instr} = a.\sqrt{(1+\beta^2)/6} \approx 0.70|MPEV| \text{ .... } (6.6)$$

where $a = 1.3\overline{3}.|MPEV|$ $\qquad \beta = \frac{3}{4}$

MPEV is the absolute value of the MPES on verification.

(ii.) After prolonged use and under varying environmental conditions, it can be assumed that, in the worst case, the measurement error extended by the measurement uncertainty will reach the values of the MPES in service. The resulting trapezoidal distribution could more or less be represented by plot (iii.) of *Figure 6.9*. In this case, the following may be assumed for the standard uncertainty contribution:

$$U_{\text{instr}} = a.\sqrt{(1+\beta^2)/6} \approx 0.90|MPEV| \text{ .... } (6.7)$$

where: $\quad a = 2\ |MPEV|$, $\quad \beta = \frac{1}{2}$

6.5.3 Uncertainty in Type Approval Tests and Conformity Decision

Conformity study and the influence uncertainty on conformity decision of newly manufactured Non Automatic Weighing Instrument (NAWI) were discussed in this section.

6.5.3.1 Assessment of Compliance

In practice, measuring instruments are considered to be in compliance with the legal regulations - if the value indicated is smaller than or equal to the maximum permissible error on verification [MPEV] when the test is performed by a verification body under unified test conditions. Type approval test data of 30 samples of NAWIs were analyzed. *Figure* 6.10 shows the relationship between measurement result, uncertainty, maximum permissible error on verification (MPEV) and criteria of conformity of 6 typical cases out of the observed data.

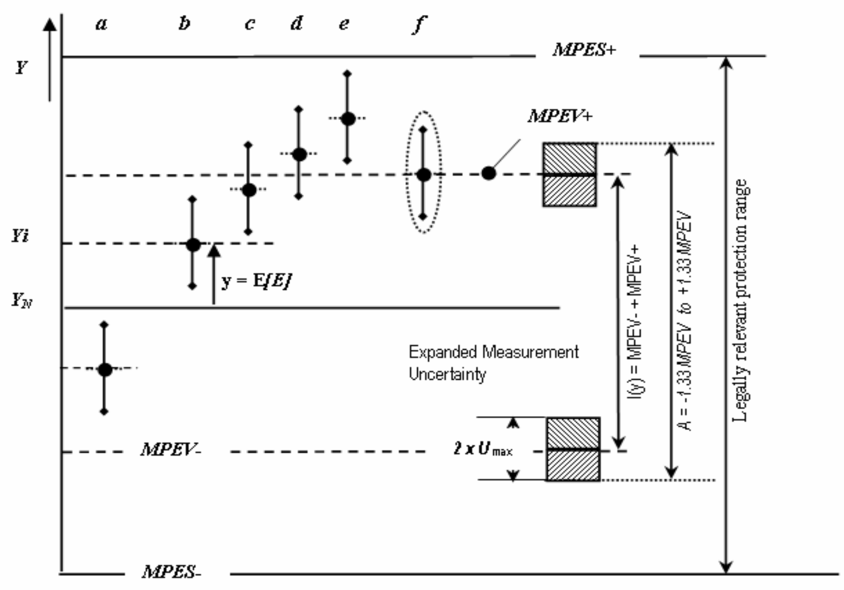


Figure 6.10 Illustration of the Criteria for the Assessment of Compliance

MPEV : absolute value of the maximum permissible error on verification
MPEV– : lower maximum permissible error on verification
MPEV+ : upper maximum permissible error on verification
MPES– : lower maximum permissible error in service
MPES+ : upper maximum permissible error in service
U_{max} : upper permissible limit of the expanded uncertainty of measurement according to equation
y : best estimate of measurement deviation, *E*
I(*y*) : acceptance interval with respect to the measurement deviation, *y*
A = range of possible occurring measurement deviations

The following observations can be made from the *Figure 6.10* :

Case a and b: both measurement result and uncertainty lie within the MPEV. In this case, the result is obviously conforming.

Case c: measurement result lies within the MPEV, but the uncertainty lies outside. In the current legislation, the result is conforming, since the conformity is decided only by the measurement value. If the measurement uncertainty is considered, however, the result can be decided as non-conforming.

Case d: Measurement result lies outside the MPEV and it is not conforming in current legislation. But if the measurement uncertainty is considered, since the measurement result lies within the MPEV, and it is decided as non-conforming

Case e: Both measurement result and uncertainty lie outside the MPEV. In this case, the result is obviously non-conforming.

Case f: Measurement result is equal to the upper limit of the MPEV and hence the result is conforming. If the measurement uncertainty is considered, however, the result can be decided as non-conforming

Regarding case c and case d, a question about how to decide the conformity considering the measuring uncertainty would arise.

In the ISO14253-117, the criteria for deciding conformity of measurement result including the measurement uncertainty are provided as follows: First, discuss how to consider the measurement uncertainty and achieve a consensus about the standard of acceptance among the parties involved. If a consensus cannot be achieved, accept only cases a and b, as conforming for consumer protection. In type approval and verification, the relevant parties are the government (or organizations designated by the government) and private companies that apply for type approval. Their relationship is not equal, as the position of the applying customer tends to be weaker. Since it is difficult to reach a consensus through discussion between them, only cases a and b will be accepted according to the ISO14253-1.

Regarding the maximum permissible error on verification (MPEV), it is a concept which was introduced at the time when the uncertainty of measurement had not been considered. To introduce the concept of the measurement uncertainty, the MPEV should also be reviewed. It is possible to extend the range of the MPEV in accordance with the measurement uncertainty. Nevertheless, the measurement uncertainty will be reduced by the development of technology in the future. It is not a worthwhile idea to review the range of MPEV periodically in accordance with the measurement uncertainty.

In today's legal metrology, the measurement uncertainty is usually considered to be small enough if the so-called "one-third uncertainty budget" (U (95%) ≤ ⅓ .MPEV) is not exceeded. Consequently, this measurement uncertainty becomes an uncertainty of the conformity decision. The criteria for the assessment of compliance are illustrated in *Figure* 6.10. Compliance with the requirements of the verification regulations is given in cases a*, b, c* and f. Cases d and e will result in rejection, although all the values, including the uncertainty of measurement, lie within the tolerances fixed by the maximum permissible errors in service [MPES].

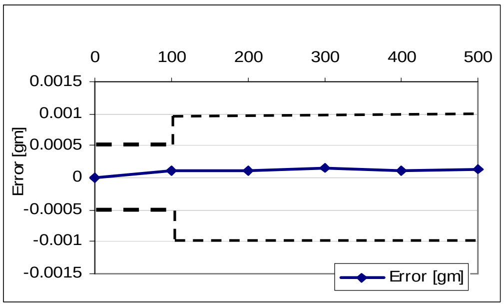


Figure 6.11. Domestic NAWI-1 of Class-II.

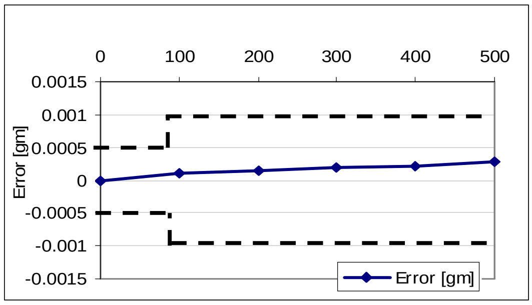


Figure 6.12 Domestic NAWI-2 of Class-II

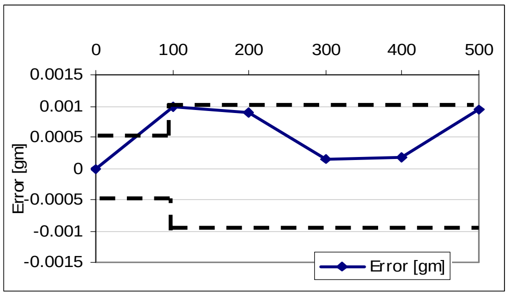


Figure 6.13 Domestic NAWI-3 of Class-II.

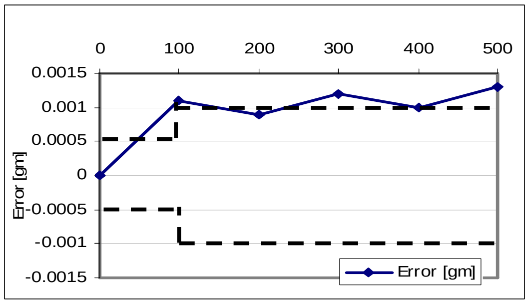


Figure 6.14 Domestic NAWI-4 of Class-I.

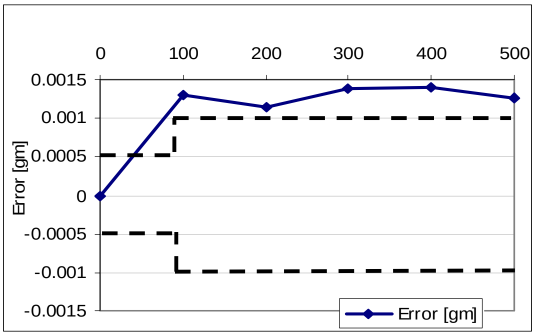


Figure 6.15 Domestic NAWI-5 of Class-I.

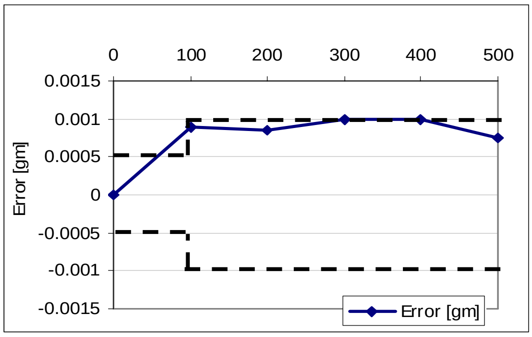


Figure 6.16 Domestic NAWI-6 of Class-I.

Consequently, the maximum permissible error on verification of a newly verified measuring instrument will in the worst case be exceeded by 33%. However, as the legally prescribed maximum permissible errors in service are valid for the instrument users, there is therefore no risk in the sense that no measured value - even if the measurement uncertainty is taken into account - will be outside this tolerance band.

The relationship between measurement results, measurement uncertainty and the MPEV of Classes I and II is shown in *Figures 6.11* to *6.16*. On the horizontal axis of each f*igure*, the mass of weights loaded on the NAWIs are shown. On the vertical axis, the deviation between the indication of the NAWIs and the mass of a weight is given. The dashed line in each f*igure* shows the range of MPEV. If the deviation is within the dashed line, the instruments will be decided as conforming.

In the case of *Figure 6.11*, weights of 100gm, 200gm, 300gm, 400gm and 500gm were loaded. When a 100 gram was loaded the indication was 100.0001, with a deviation of 0.0001gm which is well below the MPEV 0.001gm. In this manner the non-automatic weighing instrument of class II category NAWI-1, NAWI-2 and NAWI-3 [*Figure 6.11 to 6.13*] can be decided as conforming. The class I instrument, NAWI-6 [*Figure 6.16*] is also conforming. Since the measurement deviation observed is equal to or less than MPEV. But in the case of class I, NAWI-4 [*Figure 6.14*] for the mass of weights 100gms, 300gms and 500gms, the measurement deviation exceeds the MPEV. If the measurement uncertainty is considered, in this case, the instrument cannot be accepted because the proportion of the measurement uncertainty to the MPEV is less than one-third. This value of one-third is based on the experience. In the case of class-I, NAWI-5 [*Figure 6.15*] the measurement deviation exceeds MPEV in all applied loads. Therefore as per the present legal metrology rule NAWI-5 is non-conforming and would be rejected. If the uncertainty of measurement error is taken into account, still it is nonconforming because 1.33 MPEV [here it is 0.00133gm] in two loads namely 300gm and 400gm exceeds the limit of 0.00133gm. It can be observed that, when 300gm was loaded the indicated value was 300.00138 and when 400gm was loaded, it was 400.0014. Therefore NAWI-5 is non-conforming and therefore rejected.

6.5.3.2 *Manufacturer's risk*

When performing conformity assessments of measurement results to meet legal requirements, the uncertainty of measurement leads to uncertainty of decision. This leads to two effects which each have a definable probability,

- A small percentage of instruments which are actually acceptable , fail the verification process.

- A small number of instruments which despite having (true) intrinsic measurement errors which exceed the maximum permissible errors on verification, (*MPEV*), pass verification.

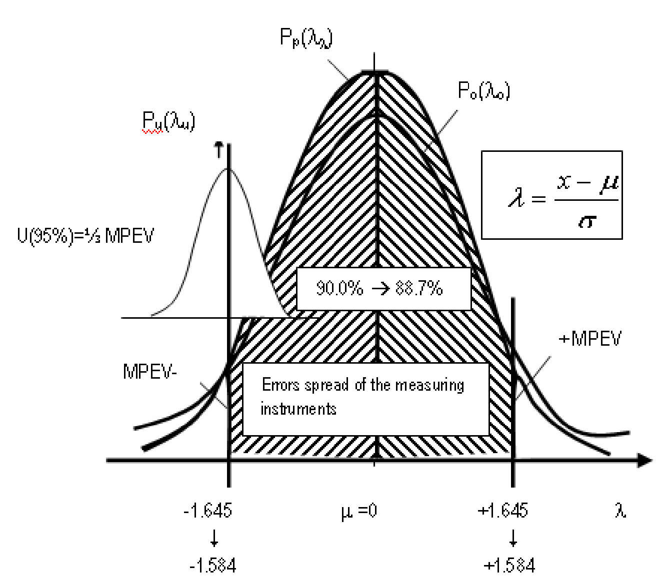


Figure 6.17 Illustration Of The Change In The Probability Distribution.

Depending on the error spread of the instrument population to be tested, the first effect can be expected to happen more often than the second. In view of the large number of instruments to be verified, this involves an economic risk for manufacturers, since it might appear that too many instruments fail the legal requirements. It is the combination of these two effects which leads to the risk of a financial loss for the manufacturer.

To obtain a reasonable estimate of the first effect, the (unknown) real error spread of the respective instrument population can be compared with the observed output scatter of the verification process. The probability distribution to be expected for verified instruments, $P_o(x)$, can be calculated by convolution of the actual distribution of the instrument population to be tested, $P_p(x)$, and the measuring process, $P_u(x)$:

$$P_o(x)=P_p(x) * P_u(x) = \int P_p(z).P_u(x-z)dz \quad \ldots\ldots\ldots\ldots \quad (6.8)$$

Assume that both input probability plots have a Gaussian distribution and the above equation applies [*Figure* 6.17].Calculation shows that for a population of instruments for which the theoretical "true" acceptance rate should be 95%, the actual acceptance rate will be 93.7%. This means that less than 1.5% of instruments have been "unfairly" rejected due to the influence of the measurement uncertainty. Consequently, the legal verification system involves only a low risk that good quality products may suffer unwarranted rejection at verification.

6.5.4 Comparison of Calibration and Verification Systems

Table 6.9 Comparison of Calibration and Verification Systems

Characteristics	*Verification*	*Calibration*
Bases	Legal requirements	Technical rules, norms, demands of customers
Objective	Guarantee of indications within MPE in service range during the validity period	Relation between indication and conventional true value (at nominated defined accuracy level)
Prerequisite	Admissibility for use in regulated area	Broad recognition of calibration results
	Admissibility for verification directly or with type approval if required	Instrument should be able to be calibrated
Validity of the results	Within the period fixed for subsequent verification(as regards MPE in service)	At the moment of calibration under specific calibration conditions
Evaluation of the results	By the verifying body	By the user of the measuring instruments
Traceability	Regulated by the procedure	Calibration laboratory to provide evidence
Uncertainty of measurement	U ≤ ⅓ MPEV	Depending on the technical competence of the laboratory and of the instrument performance.

Technical procedures followed in verification is equivalent to those used in calibration and provides confidence in the correctness of indication of verified instruments although no expert knowledge by the instrument's user is required. Verification therefore, may be considered a strong tool in both legal metrology and quality assurance, when large numbers of measuring instruments are involved. In particular, it excels as a single means by which enforcement can be realized, and because the user is only affected by the MPES in service, it provides a high degree of confidence over a long period of time. One disadvantage in verification is that the influence of uncertainty on a decision of conformity of a measuring instrument to specific requirements is not completely clear. In comparison, traditional calibration is considered an important basic procedure for legal metrology activities and also for fundamental measurement application in scientific and industrial metrology. It is practically not limited as far as the measurement task is concerned, but does require sound expert knowledge on the part of the instrument's user in carrying out and evaluating measurements.

6.6 CONCLUSION

The state metrological compliance system was developed to provide, through legislative requirements, trust and confidence in the measurements and minimize disputes and transaction errors. The different standards used in legal metrology are reference standards, secondary standards, working and commercial standards. Each of these standards is compared with immediate superior measuring standards during periodical calibration and verification.

In this research work five years verification and test data of various working standards such as non-automatic weighing instrument, weights and volumetric measures were analyzed for time variance to check the error growth with ageing. A study was conducted to find out the relationship between legally prescribed error limits and uncertainty, impact of inclusion of uncertainties in conformity decision of NAWI. Comparison of calibration and verification is also done.

In the verification and test data analysis, it is observed that the error on an average increases with ageing of the instrument and there was a significant error variation from instrument to instrument. Three tests namely weighing test, eccentricity test and repeatability test were conducted in the verification procedure of non-automatic weighing instrument. Annual Average Growth [AAG] of error is maximum with weighing test and is 89.8% followed by repeatability test where AAG is 54.81% and AAG is 21.9% in eccentricity test. Verification results of twenty sets of working standards of weight measures, each having 23 different solid weights ranging from 20kg to 1mg were analyzed. It is observed that there is an AAG of error of 30.31%. It is also observed that AAG of error is minimum with volumetric measure compared to weights and non-automatic weighing instrument and is only 23.82%. One-way, two-way ANOVA and multiple comparison tests confirm the significance of the findings. Commercial standard being a standard used for calibrating the measures used in shops and establishments directly, similar time based characterization of the observations from periodical calibration were carried out. Regression models were also developed out of the time based characterization results of the observations from periodical calibration (details are given chapter VII). Regression analysis curve can be used as a benchmark for the error growth with ageing of the instrument. Verifying authorities may fit their observed values in the regression curve and if these values are well above the regression curve, the instruments can be rejected. More accurate ANN based prediction model is also discussed in chapter VII.

The maximum permissible error is a concept which does not take into account the uncertainty of measurement. Type approval test and verification are basic rules provided by the measurement law and related regulation. The study proposes the following criteria for conformity decision: The measurement uncertainty in the type approval tests is different from the measurement uncertainty in performance of non-automatic weighing instruments (NAWIs). The factors of measurement uncertainty in type approval tests are dispersion of rounding error and dispersion of the mass of weights. If conformity decision in type approval tests is based on the average or standard deviation of the measurement values, the dispersion of repeatability should also be evaluated as one of the factors of the measurement uncertainty. In type approval tests, if the measurement uncertainty is equal to or less than one-third of the MPEV, the measurement uncertainty should be considered in conformity decision. In this case, if the measurement results including the measurement uncertainty lie within the MPEV, it should be decided as conforming, otherwise declared nonconforming. There is therefore, a certain risk for the manufacturer that acceptable instruments might be rejected. In this study, the probability of this has also been calculated from the uncertainty of measurement and the spread of the errors of measurement of the measuring instrument population to be tested. Before applying these criteria, the consensus among the manufacturers of each instrument shall be achieved. In type approval tests, if the measurement uncertainty exceeds one-third of the MPEV, the measurement uncertainty should not be considered in conformity decision

The legal metrology department is handling all type of working standards, which includes the working standard for weighing precious metals to weigh bridge weighing tons. They follow uniform procedures for the conformity study of various types of working standards. The findings of this study of uncertainty can be incorporated for the unsurpassed conformity, at least for the instruments for low value. Efforts should be made to reduce the measurement uncertainty. Incorporation of measurement uncertainly as stated above in legal metrology testing scheme will improve its quality.

Chapter VII

Models for Prediction of Measurement Errors Using Regression Analysis and Artificial Neural Network

7.1 INTRODUCTION

Characterization of calibration errors of selected instruments (details given in Chapter V and VI) shows that calibration errors grow with ageing. It is also noticed that the error growth is different for different types of instruments. Test and calibration engineers usually follow the prescribed rules for deciding the period between calibrations, since they do not know the extent of error growth of instruments in use before getting it calibrated. It was also found that the error prediction capability of professionals, regarding the change in extent of error in measuring instruments with use (time) was poor. Discussions with professionals revealed that, an accurate prediction of errors would help them to schedule calibration depending on requirement, and not always follow thumb rule like calibrate after twelve months. To enable this, it was decided to build models for error prediction using the calibration data collected. Two approaches were taken, one based on regression (time series) and the other using Artificial Neural Network (ANN) [140]. Time dependant errors observed during calibration of selected sophisticated instruments, and some working standards used for calibration in legal metrology, were used to build the models to predict future value of such calibration errors in this research work. Artificial Neural Network [ANN] technique also has been utilized to predict human related measurement errors.

7.2 PREDICTION OF THE ERRORS USING REGRESSION ANALYSIS

Regression is the determination of a statistical relationship between two or more variables. In simple regression there are only two variables, one variable (defined as independent) is the cause of the behavior of another one (defined as dependent variable). Regression can only interpret what exists physically, that is, there must be a physical way in which independent variable 'x' can affect dependent variable 'y'. The regression analysis is a statistical method to deal with the formulation of mathematical model depicting relationship amongst variables which can be used for the purpose of prediction of the values of dependent variable, given the values of the independent variable.

i. Simple Linear Model

Linear model can be derived when one variable, y called the dependent variable is 'driven by' some other variable x called the independent variable. In addition, suppose that the relationship between y and x is basically linear, but is inexact: besides its determination by x, y has a random component, u, which is called the 'disturbance' or 'error'. In this research, the independent variable is time and the dependant variable is measurement error of the instrument. The regression model therefore becomes a time series model in this case.

Let *i* index the observations on the data pairs(x, y). The simple linear model formalizes the ideas just stated:

$$y_i = \beta_0 + \beta_1 x_1 + u_i \qquad [7.1]$$

The parameters β_0 and β_1 represent the y-intercept and the slope of the relationship, respectively [141].

Measurement Errors observed during Calibration of selected sophisticated instruments and working standards used in legal metrology, over five years were used as the dependant variable, time as the independent variable, for each type of instrument and models were developed. Details of the five years calibration data of the sophisticated instruments, is shown in *Table* 5.1 of chapter V and the same for working standards is shown in *Table* 6.1 of chapter VI. These were used for development of the models for prediction of errors that can be observed during calibration.

7.2.1 Details of the Regression Based Models Developed

Using the five year data discussed above, Regression equations, curves and R^2 values were found out in each case. These equations can be used for prediction.

7.2.1.1 Sophisticated Equipments

Regression based models of selected sophisticated instruments such as Digital Multimeter, Digital Thermometer, Cathode Ray Oscilloscope, Signal Generator and Pressure gauge were discussed in this section.

i. Digital Multimeter –DMM

DMM is basically used in six different modes of measurement. They are DC voltage, AC voltage, DC current, AC current, Resistance and Capacitance. Separate models were made for measurement error prediction in each mode. For each model, five years error observed during calibration (calibration data) of 10 instruments were used.

a. DMM – Direct Current Voltage(DCV)

DC Voltage mode can measure five different ranges of input voltage as is given in *Table* 5.6 of chapter V. Five year's calibration data of ten different instruments is used for regression analysis. Regression equations, curves and R^2 values have been found and are given in *Figure* 7.1

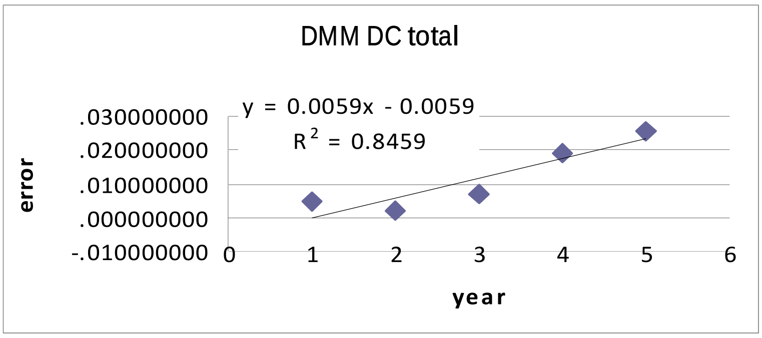


Figure 7.1 Graph Showing Regression Curve, Equation and R^2 Value

Similar set of regression equations, curves and R^2 values were found out for all operating ranges of voltage.

b. DMM Alternating Current voltage (ACV)

Calibration of ACV mode of DMM is done by using four different groups of AC Voltages and within each group; three different voltages were applied to the DMM as given in *Table* 5.6. Regression analysis equation with the R^2 value for a single group is given below as a sample.

$Y = 0.0002\,x + 0.0003$; $R^2 = 0.891$ for 4V group

Mean residue (difference between actual and predicted measurement error value) for the four different ranges of ACV operation of DMM is obtained as 0.0008424.

c. DMM DC Current

Calibration of DC Current mode of DMM is done by applying three different ranges of current as shown in *Table* 5.6. Regression equation with R^2 value for 0.004A range of operation is as follows:

$Y = 0.00000016\,x + 0.0000016$; $R^2 = 0.9817$ for 0.004A

d. DMM AC Current

DMM in alternating current mode is calibrated in the same manner as that of DC Current mode by applying different currents given in *Table* 5.6. Regression equation with R^2 value for 0.004A is as follows:

$Y = 7E\text{-}07\, x + 3E\text{-}06$; $R^2 = 0.8278$ for 0.004A

e. DMM – Capacitance mode

Capacitance mode of DMM is calibrated by applying five different capacitance groups as given in *Table* 5.6. Regression equation and R^2 value for 4nF group is given below:

$Y = 0.0664\, x + 0.1593$; $R^2 = 0.9157$, for 4nF

e. DMM – Resistance mode

Resistance mode of operation of DMM is calibrated by applying six different groups of resistance. Regression equation for 0.4kΩ of applied input resistance for calibration is given below:

$Y = 0.0002\, x - 0.0001$; $R^2 = 0.9891$ for 0.4kΩ

ii. Digital Thermometer

Digital Thermometer using thermocouple is calibrated by applying four different temperatures. Regression equations and R^2 value for the applied temperatures are given below:

$Y = 0.1309x + 0.0433$; $R^2 = 0.9567$ for -50°C
$Y = 0.0869x + 0.0758$; $R^2 = 0.9324$ for 0°C
$Y = 0.1289\, x + 0.6115$; $R^2 = 0.9963$ for 500°C
$Y = 0.1751x + 0.6557$; $R^2 = 0.9782$ for 950°C

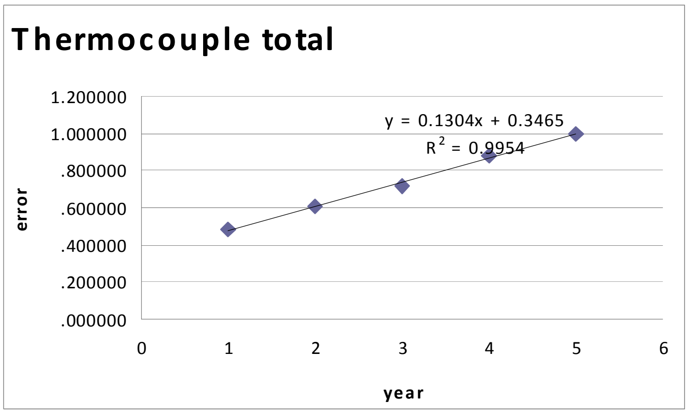


Figure 7.2 Regression Curve, Equation and R^2 value - Digital Thermometer

Cumulative regression equation, regression curve and R^2 value is given in *Figure* 7.2 Actual and predicted values of errors are very close with a minimum residue as can be observed from the above f*igure*. It can also be observed that the predicted error values are more accurate for higher range of applied input temperature for calibration.

iii. Cathode Ray Oscilloscope

Cathode ray Oscilloscope is calibrated in the vertical axis, horizontal axis or frequency axis mode by applying different voltages and square wave with different frequencies and voltages as given in *Table* 5.11 of chapter V.

a. Vertical Axis Mode Calibration

In vertical axis mode, calibration is done for DC and AC voltages separately.

CRO- DC Voltage Calibration

It is observed that, prediction of errors using regression model for different dc voltages used for calibration is accurate with R^2 value in all the cases close to 0 .9.

CRO AC Voltage Calibration

AC Voltage calibration of CRO is done by applying 1 kHz square-wave of different amplitudes as given in *Table* 5.11.

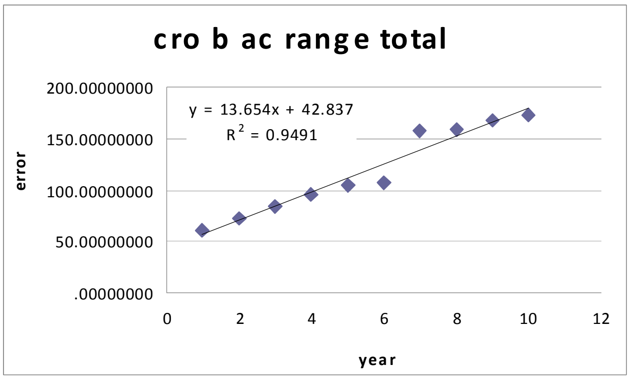


Figure 7.3 Graph Showing Regression Curve, Equation and R^2 Value-CRO

Total regression analysis equation, R^2 value and regression curve are shown in *Figure* 7.3. It can be observed from the *Figure* that the predicted errors almost follow the actual errors. R^2 value is also close to one.

b. Horizontal axis calibration

CRO Time Calibration

In the time-calibration 26 different frequencies are applied as is given in *Table* 5.11. Residue of error estimation is more for the horizontal axis calibration in low time period or high frequency range of applied inputs. R^2 value is also less than 0.5 starting from .005μS to 1μS.

iv. Signal Generator

Signal generator is calibrated by applying four different groups of frequencies which are 0.4 kHz, 4 kHz, 40 kHz, 400 kHz and 2000 kHz. It is observed that for the lower frequency ranges 0.4 kHz and 4 kHz, the residue is more and the R^2 value is less than 0.9. In the higher range of applied input frequencies the R^2 value is close to one and the residue is less and thus predicted errors are close to the actual value of errors. This indicates that the use of regression model for error prediction will yield more accurate results when applied frequency range is higher.

v. Pressure Gauge

Pressure gauge is calibrated by applying different pressures in the increasing and decreasing order. Different pressures applied are 60kg/cm², 100 kg/cm², 200 kg/cm², 300 kg/cm², 400 kg/cm², 500 kg/cm², and 600 kg/cm².

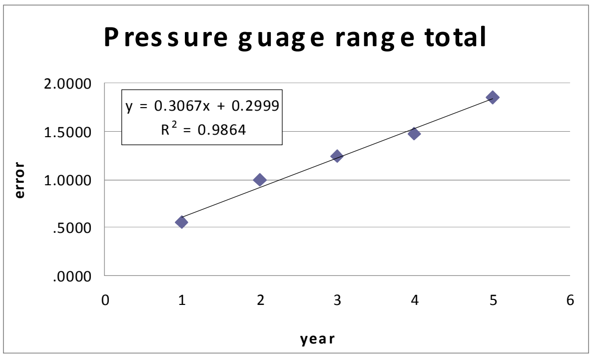


Figure 7.4 Regression Curve, Equation and R^2 Value-Pressure Gauge - Increasing

Total pressure gauge regression equation, R^2 value and regression curve for increasing order of applied input pressure for calibration is shown in *Figure* 7.4. Mean residue is 0.044036.

7.2.1.2 Working Standards Used in Legal Metrology

Errors observed during calibration of selected working standards such as Non Automatic Weighing Instrument, Volumetric Measures and Weight Measures were used for error modeling.

i. Non Automatic Weighing Instrument (NAWI)

Regression equation, regression curve and R^2 value which predicts the time dependant error for the three different tests, the weighing test [up and down], eccentricity test [e-test] and repeatability test [r-Test] used for calibrating NAWI is shown separately in *Figure* 7.5. *Figure* shows that the residue is minimum with R-test [0.0036] and maximum with weighting test [0.453]. R^2 value is close to one except for W-test (down).

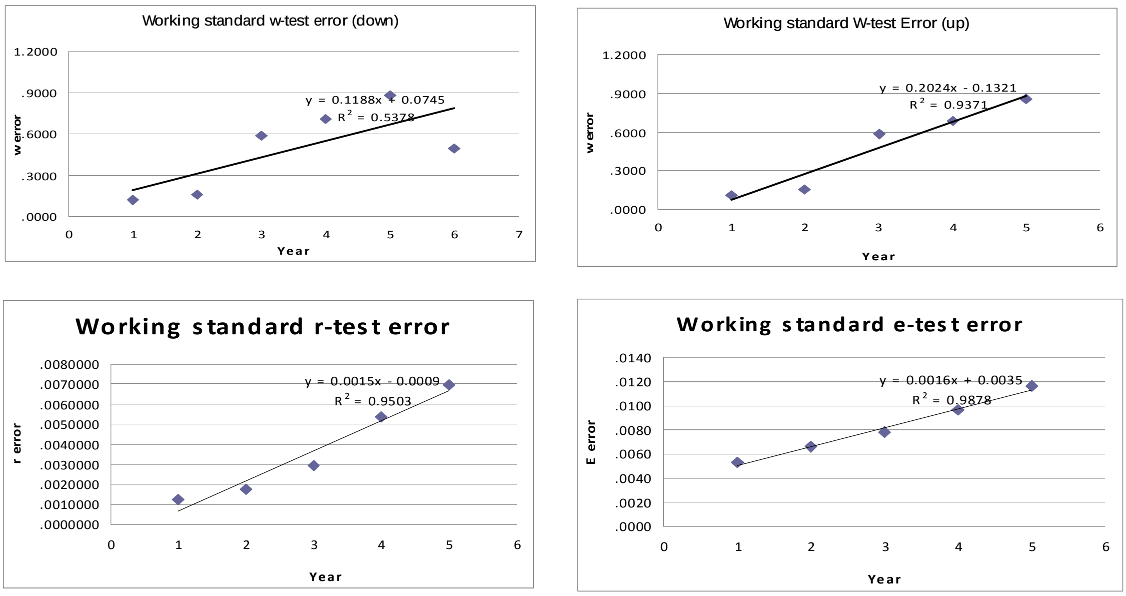


Figure 7.5 Regression Curve, Regression Equation and R^2 value - NAWI

ii. Volumetric Measures

Future possible errors that can be observed during calibration of working standards used for volumetric measurement are given in *Figure* 7.6. *Figure* shows that the predicted values are close to the actual values with an R^2 value of 0.9974.

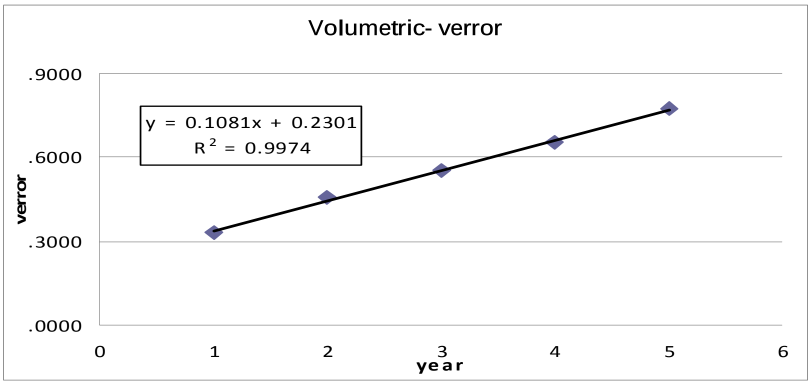


Figure 7.6 Regression Curve, Equation and R^2 Value-Volumetric Measure

iii. Working Standard for Weight measurement

The regression parameters of the prediction model for weight measure are given in *Figure* 7.7. The R^2 value is 0.9257.

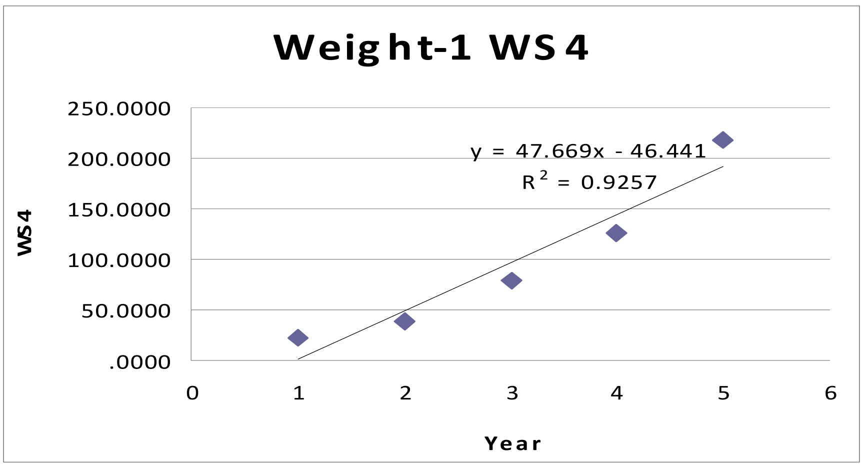


Figure 7.7 Regression Curve, Equation and R^2 value-Weight Measure

7.3 PREDICTION MODELS USING ARTIFICIAL NEURAL NETWORKS

Advances have been made in developing tools for “intelligent" computer programs, a category of which have been inspired by biological neural networks, these are called artificial neural networks (ANNs). Researchers from various scientific disciplines are using artificial neural networks (ANNs) to solve a variety of problems in decision making, optimization, prediction, classifications and control [142]. The second set of measurement error prediction model, developed in this thesis uses ANN.

7.3.1 Modeling of Time Dependant Errors Using ANN

The recent increase in usage of artificial neural networks has proven that neural networks have powerful pattern classification and prediction capabilities [143]

Numerous efforts to develop “intelligent” programs based on Von Neumann’s centralized architecture have not resulted in general-purpose intelligent programs [144]. Inspired by biological neural networks, ANNs are massive parallel computing systems consisting of an extremely large number of simple processors with many interconnections [145]. ANN models attempt to use some “organizational” principles believed to be used in the human brain [146].

Function approximation: Given a set of n labeled training patterns (input-output pairs), $\{(x_1; y_1); (x_2; y_2); \ldots\ldots. (x_n, y_n)\}$, generated from an unknown function $\mu(x)$ (subject to noise), the task of function approximation is to find an estimate, say μ, of the unknown function μ [*Figure* 7.2]. Various engineering and scientific modeling problems require function approximation [147].

Prediction/forecasting: Given a set of n samples {$y(t_1)$; $y(t_2)$,….$y(t_n)$g in a time sequence, t_1; t_2; ….. t_n, the task is to predict the sample $y(t_{n+1})$ at some future time t_{n+1}.

Prediction/forecasting have a significant impact on decision making in business, science and engineering. Stock market prediction and weather forecasting are typical applications of prediction/forecasting techniques.

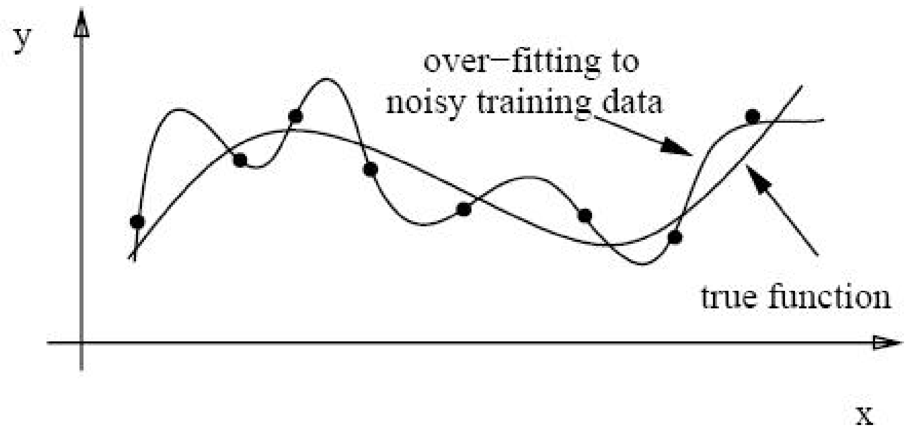


Figure 7.8 Function Approximation

7.3.2 Computational Models of Neurons

McCulloch and Pitts proposed a binary threshold unit as a computational model for a neuron [148]. This mathematical neuron computes a weighted sum of its n input signals, x_j; j = 1; 2; …… n, and generates an output of 1 if this sum is above a certain threshold u and an output of 0 otherwise. Mathematically,

$$y = \theta\left(\sum_{j=1}^{n} w_j x_j - u\right) \qquad [7.2]$$

where θ(.) is a unit step function at zero, and w_j is the synapse weight associated with the j^{th} input. For simplicity of notation, it is often considered the threshold u as another weight w^0 = -u which is attached to the neuron with a constant input, $x^0 = 1$. Positive weights correspond to excitatory synapses, while negative weights model inhibitory ones [149].

McCulloch and Pitts proved that with suitably chosen weights let a synchronous arrangement of such neurons is, in principle, capable of universal computation. There is a crude analogy here to a biological neuron: wires and interconnections model of axons and dendrites, connection weights represent synapses, and the threshold function approximates the activity in a Soma [150].

The model of McCulloch and Pitts contains a number of simplifying assumptions, which do not reflect the true behavior of biological neurons.

The McCulloch-Pitts neuron has been generalized in many ways. An obvious generalization is to use activation functions other than the threshold function, e.g., a piecewise linear, sigmoid, or Gaussian. The sigmoid function is by far the most frequently used function in ANNs [151]. It is a strictly increasing function that exhibits smoothness and has the desired asymptotic properties. The standard sigmoid function is the logistic function, defined by

$$g(x) = 1/(1 + \exp(-\beta x)); \qquad [7.\ 3]$$

where β is the slope parameter.

7.3.3 Network Architecture

An assembly of artificial neurons is called an Artificial Neural Network (ANN). ANNs can be viewed as weighted directed graphs in which nodes are artificial neurons and directed edges (with weights) are connections from the outputs of neurons to the inputs of neurons [152]. Feed-forward network architecture has been used in this research [153]. The most common family of feed forward networks is a layered network in which neurons are organized into layers with connections strictly in one direction from one layer to another.

Generally speaking, feed-forward networks are static networks, i.e., given an input, they produce only one set of output Neural Networks values, not a sequence of values [154]. Feed-forward networks are memory-less in the sense that the response of a feed-forward network to an input is independent of the previous state of the network. Recurrent networks are dynamic systems. Upon presenting a new input pattern, the outputs of the neurons are computed [155].

7.3.4 Learning [training of neural network]

Ability to learn is a fundamental trait of intelligence. Although a precise definition of learning is often difficult to state, a learning process in the context of artificial neural networks can be viewed as the problem of updating network architecture and connection weights so that a network can efficiently perform the specific task for which it is designed [156]. Most of the time, the network must learn the connection weights from the available training patterns. Improvement in performance is achieved over time through iteratively updating the weights in the network [157].

The ability of artificial neural networks to automatically learn from examples makes them very attractive and exciting. Instead of having to specify a set of rules, ANNs appear to learn them from the given collection of representative examples [158]. This is one of the major advantages of neural networks over traditional expert systems.

In order to understand or design a learning process, one must first have a model of the environment in which a neural network operates, i.e., what information is available to the neural network. This model is referred as a learning paradigm [159]. Second, one must understand how weights in the network are updated, i.e., what are the learning rules which govern the updating process. A learning algorithm refers to a procedure in which learning rules are used for adjusting weights in the network.

Supervised learning paradigm has been used in this study. In supervised learning or learning with a teacher, the network is provided with a correct answer to every input pattern. Weights are determined so that the network can produce answers as close as possible to the known correct answers.

7.3.5 Algorithm in MATLAB for Error Models

In the supervised learning paradigm, the network is given a desired output for each input pattern. During the learning process, the actual output, y, generated by the network may not equal the desired output, d. Error- correction rules are followed in this algorithm. The basic principle of error-correction learning rules is to use the error signal (d - y) to modify the connection weights such that this error will be gradually reduced [160].

The well-known perceptron learning rule is based on this error-correction principle. A perceptron consists of a single neuron with adjus*table* weights, w_j; j = 1,2,….n, and threshold u. Given an input vector x = (x1; x2; …… xn) t, the net input to the neuron (before applying the threshold function) is

$$v = \sum_{j=1}^{n} w_j x_j - u \qquad [7.4]$$

The output y of the perceptron is +1 if v > 0, and 0 otherwise. In a two-class classification problem, the perceptron assigns an input pattern to one class if y = 1, and to the other class if y = 0. The linear equation $\sum_{j=1}^{n} w_j x_j - u = 0$ defines the decision boundary (a hyperplane in the n-dimensional input space) which divides the space into two halves.

Note that learning occurs only when an error is made by the perceptron. Rosenblatt proved that if the training patterns are drawn from two linearly-separable classes, then the perceptron learning procedure will congregate after a finite number of iterations [161]. This is the well known perceptron convergence theorem. In practice, one does not know whether the patterns are linearly separable or not. Many variations of this learning algorithm have been proposed in the literature [162] [163]. Other activation functions can also be used, which lead to different learning characteristics. However, a single layer perceptron can only separate linearly separable patterns, as long as a monotonic activation function is used.

Back-propagation learning algorithm, which is based on the error-correction principle [164], is used in this research study.

Back-Propagation Algorithm

1. Initialize the weights to small random values;
2. Randomly choose an input pattern $x^{(\mu)}$;
3. Propagate the signal forward through the network;
4. Compute δ_i^L in the output layer $(o_i = y_i^L)$

   $$\delta_i^L = g^1(h_i^L)\left[d_i^\mu - y_i^L\right],$$

   where h_i^l represents the net input to the i^th^ unit in the ℓ^{th} layer, and g' is the derivative of the activation function g.
5. Compute the deltas for the preceding layers by propagating the errors backwards;

   $$\delta_i^l = g'(h_i^l \sum_j w_{ij}^{l+1}\delta_j^{l+1}, \qquad \text{for } l = (L-1), \ldots, 1$$
6. Update weights using

   $$\Delta w_{ji}^l = \eta \delta_i^l y_j^{l-1}$$
7. Go to step 2 and repeat for the next pattern until the error in the output layer is below a pre-specified threshold or a maximum number of iterations is reached.

ANN based models were developed for doing the following, in this research work:

i. Prediction of time dependant error, in selected sophisticated measuring instruments and, in working standards used in legal metrology.

ii. Prediction of human error in measurement with given combination of factor values. .

A typical feed-forward back propagation neural network used in this model consists of three layers, the input layer, a hidden layer and the output layer as shown in *Figure* 7.9 The feed-forward fully connected network was trained in supervised manner. The input layer consists of twenty four neurons, hidden layers consist of twenty neurons and the output layer consists of a single neuron. The input and target samples are automatically divided into training, validation and test sets. Training set was used to teach the network. Training was continued as long as the network continues improving on the validation set. Test set provided a completely independent measure of network accuracy. The information moves in only one direction, forward, from the input nodes, through the hidden nodes and to the output nodes. There are no cycles or loops in the network.

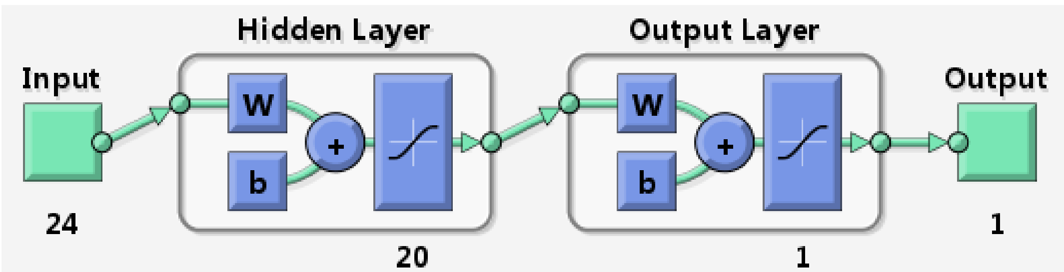


Figure 7.9. Error Prediction Model Based on Neural Network.

The above network allows signals to travel one way only; from source to destination. The hidden neurons are able to learn the pattern in data during the training phase and mapping the relationship between input and output pairs. Each neuron in the hidden layer uses a transfer function to process data it receives from input layer and then transfers the processed information to the output neurons for further processing using a transfer function in each neuron. The output of the hidden layer can be represented as equation

$$Y_{N \times 1} = f(W_{N \times M} X_{M \times 1} + b_{N \times 1}) \qquad [7.5]$$

where, Y is a vector containing the output from each of the 'N' neurons in a given layer, W is a matrix containing the weights for each of the 'M' inputs for all 'N' neurons, 'X' is a vector containing the inputs, 'b' is a vector containing the biases and 'f(.)' is the activation function[165][166]

7.3.6 Prediction Model for Calibration Errors

Prediction of calibration errors (errors that is observed during calibration) in selected sophisticated instruments and in selected working standards which are used for calibration in legal metrology, were done using Artificial Neural Network [ANN]. The sophisticated instruments selected are digital multimeter, digital thermometer, cathode ray oscilloscope, signal generator and pressure gauge. Working standards selected for the study include, non-automatic weighing instruments, weight measures and volumetric measures. As in the case of regression based models, errors observed during calibration(five years) of the sophisticated instruments given in *Table* 5.1 of chapter V and the errors observed during calibration(five years) of working standards listed in *Table* 6.1 of chapter VI were used for modeling. Residues [difference between actual and predicted values] of prediction errors were also found out. A comparison of the residue of prediction of regression analysis and artificial neural network is also given.

Back propagation algorithm used for developing code in MAT LAB neural network for prediction of errors can be further interpreted as follows:

i. The errors observed during calibration for year1 to year 4 is the training data.

ii. Year 5 data is the target data.

iii. 80% of the data is taken for training the neural network and 20% for testing the neural network.

iv. As the errors observed during calibration increase in ascending order, the training data and testing data were selected at random and not in sequential order.

v. In order to compare one machine with another machine separate folders were created for each machine and the .xls file inside the folders were given identical file names and sheet names.

vi. Math-Lab built-in functions are used to train and test the neural network.

7.3.6.1 Sophisticated Instruments

i. Digital Multimeter

Prediction of errors using ANN was carried out for all the six modes of operation of DMM. Out of the six different modes of operation of DMM, the least residue of prediction result is with DMM-DC current (0.0000293) and the maximum with DMM Capacitance (0.157).

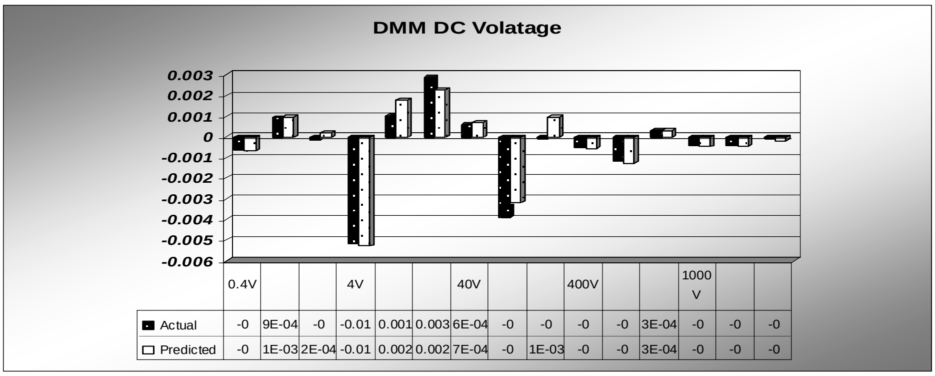


Figure 7.10 Graph Showing Actual Vs Predicted Error Values Using ANN-DMM

In DMM- DC voltage mode, the maximum residue is with 40V range and the residue is 0 with 400V and 1000V range as can be noticed from *Figure* 7.10 Minimum residue of DMM- AC Voltage is 0.0003533 and the maximum residue is 0.0006067. It can also be observed from the *Figure* 7.10 that the maximum residue is with 40V range and the minimum residue is with 400V range.

In DMM-DC current mode, the minimum residue is with 0.004A (ie.0.0000246) and the maximum is with 0.4A (0.00003463).In DMM-AC current mode, the 10A range simulation gives minimum residue (0.00002) and maximum is with 0.4A (0.0003115). Residue is minimum with 4kΩ (0.0014894) and is maximum with 40kΩ (0.006733) in the resistance mode of operation of DMM. Prediction of error in capacitance mode shows that, minimum residue of prediction is with 40000nF (0.03485) and the maximum residue is for 4000 nF (0.343419).

ii. Digital Thermometer

Figure 7.11 shows the comparison of actual and predicted values of errors with digital thermometer. Residue is minimum with the calibration temperature of 950°C out of the four different temperatures used for calibration and the value is 0.016. Maximum residue is 0.2633 and is for the input -50°C

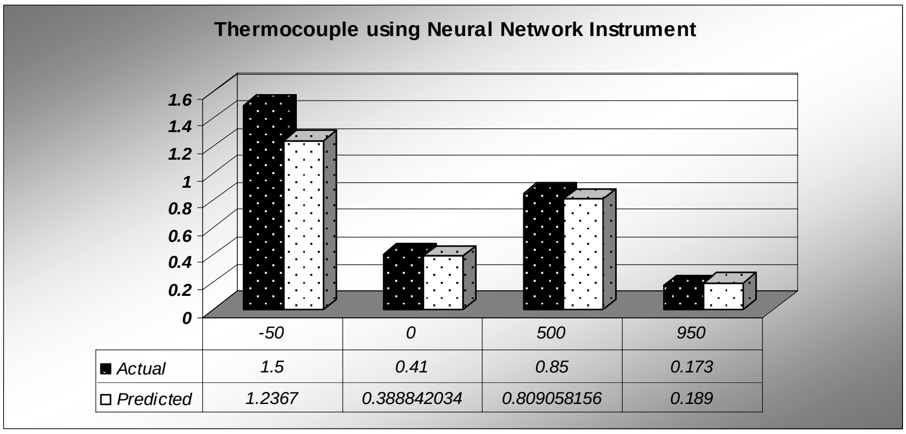


Figure 7.11 Actual Vs. Predicted Values of Errors for Digital Thermometer

iii. Cathode Ray Oscilloscope

a. Vertical Axis calibration-DC Voltage

Figure 7.12 shows the actual and predicted values of errors for twelve different ranges of voltage used for calibration. Residue is minimum with 10 mV range (0.01044) and the maximum residue with 100mV range(2.5972).

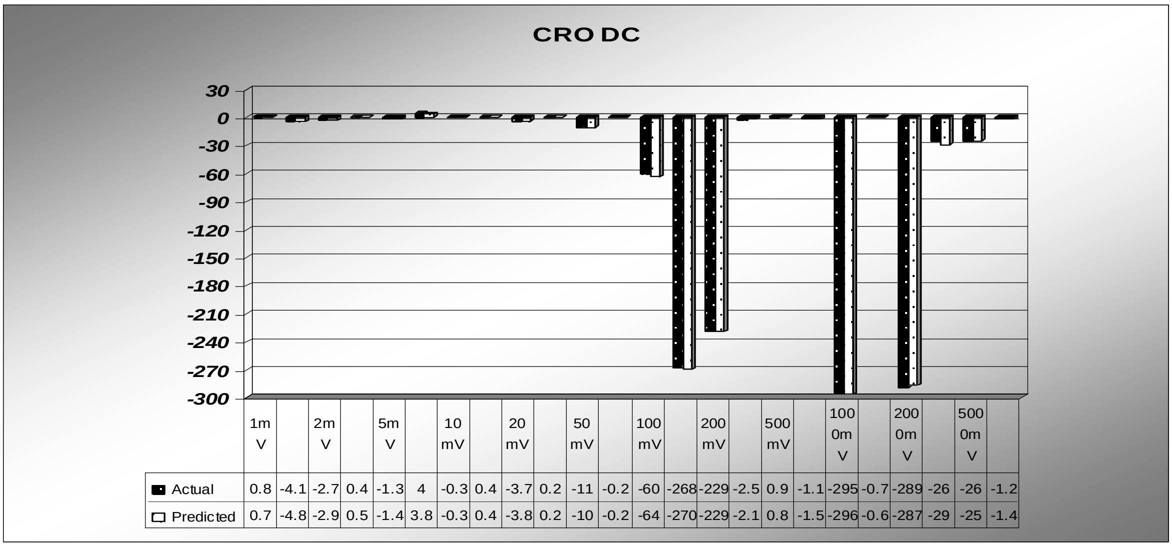


Figure 7.12 Actual Vs. Predicted Values of Errors for CRO

AC Voltage

In AC voltage mode of calibration, the minimum residue is with 5V range (0.083074) of applied input for calibration and maximum is with 200V (2.04263). Generally it is observed that the prediction error (possible errors that can be observed during calibration) values are close to actual error values.

b. Horizontal Axis Calibration

CRO Time Mode Prediction

Prediction of errors that can be observed during calibration in the time mode of CRO shows that the residue is minimum with 20μs (0.0000044) and maximum with 0.1μs (0.9056). It is also observed that the predicted errors using ANN is almost equal to actual errors.

iv. Signal Generator

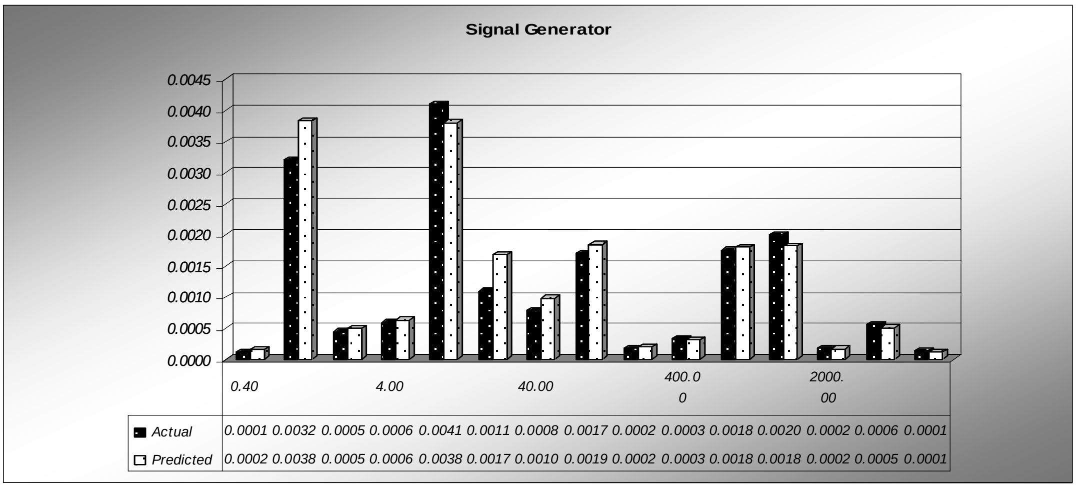
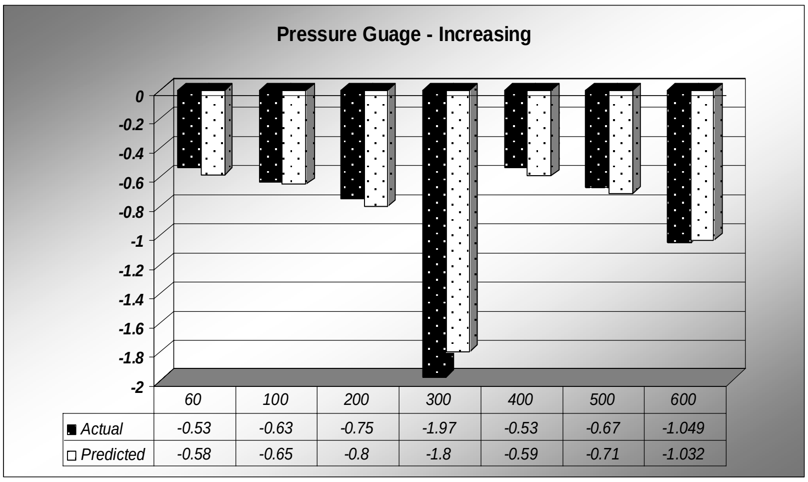


Figure 7.13 Actual Vs. Predicted Values of Errors for Signal Generator

Signal generator is calibrated by applying signals with five different frequencies as is given in *Figure* 7.13. It is observed that the average residue is 0.00015 and in majority of the cases, residue is zero. It is also noticed that the residue is minimum with signal generator out of the five different sophisticated instruments studied in this research work.

v. Pressure Gauge

Pressure gauge is calibrated by applying seven different pressures in the increasing and decreasing order. *Figure* 7.14. shows that, the difference between actual and predicted values of errors are generally less when the pressure is applied in the increasing order for calibration rather than the decreasing order .



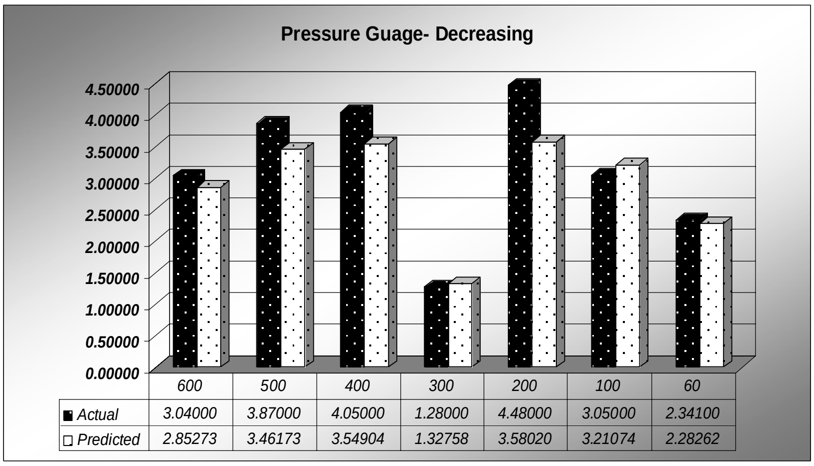


Figure 7.14 Graph Showing the Actual Vs. Predicted Values of Errors for Pressure Gauge

Average residue of increasing order is 0.058 and for decreasing order, the average is 0.3233.

7.3.6.2 Working Standards Used in Legal Metrology

Neural network was trained using the errors observed during calibration of working standards of non automatic weighing instrument, volumetric measures and weight measures. Initially five years error observed during calibration of ten different standards were used to train neural network. Then to check the effect of residue on the predicted errors, the data size has been increased to ten years data of 50 working standards each from NAWI, volumetric and weight measures. Increased data size reduces residue and is also included in the following discussion.

i. Non Automatic Weighing Instrument (NAWI)

Neural network was trained, first with, errors observed during calibration of ten NAWI and then with fifty NAWI for prediction of errors. Actual and predicted values of errors observed during calibration when training is done with ten instruments are given in *Figure* 7.15. It is observed that the average residue with ten instruments is 0.002324 and with fifty instruments is .0015152.

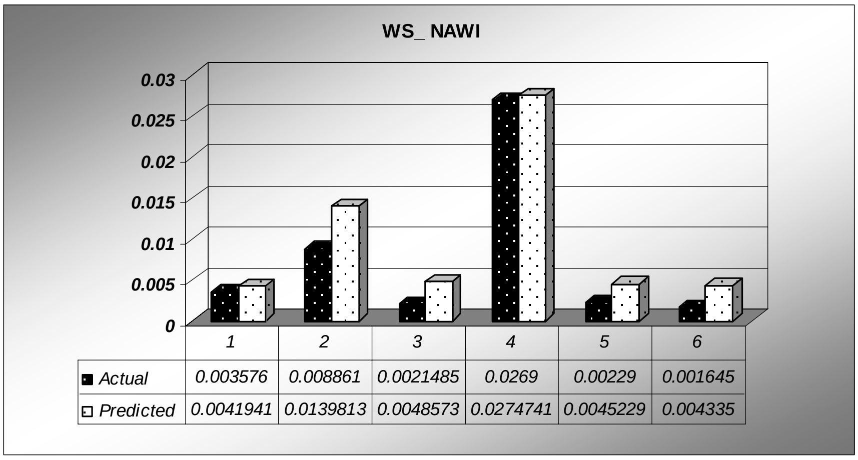


Figure 7.15 Actual Vs. Predicted Values of Errors Using ANN for NAWI

It is also observed that, the residue with regression model is more than ANN model by 0.0494 in the case of NAWI.

ii. Volumetric Measures

Errors that can be observed during calibration of the working standard, volumetric measures were predicted by training neural network using five years error observed during calibration of ten instruments. To study the effect of increased data size, the training data size was increased to ten years calibration data (error observed during calibration) of fifty instruments. Average residue is 0.21257 for 10 instruments and when the neural network is trained with the calibration data of 50 instruments the average value of residue is reduced to 0.05969. It is observed that the residue with ANN model is more compared to regression model and the difference is 0.072298.

iii. Weight Measures

Error prediction model was trained, first with five years error observed during calibration of ten weight measures standard and then training is done with the calibration data (error observed during calibration) of 50 instruments as before. Average residue with weight measures using ten instruments is 0.49342 and for fifty instruments it has reduced to 0.40226. Residues with regression analysis are more by 0.563 than that of ANN.

7.4 COMPARISON OF REGRESSION AND ANN MODEL FOR ERROR PREDICTION

Errors observed during calibration of selected sophisticated instruments were utilized for predicting errors using Regression analysis technique and Artificial Neural Network. Residue, which is an indication of error of model output, is compared for regression and artificial neural network based models. The results of the comparison are shown in *Table* 7.1.

Table 7.1 Comparison of Residues in Regression and ANN models

Instruments	*Residues*	
	Regression Analysis	*Artificial Neural Network*
DMM – DCV	0.0003288	0.00028
DMM – ACV	0.0008424	0.00046
DMM – DCC	0.0001232	0.0000293
DMM – ACC	0.000600	0.00018
DMM – Resistance	0.0696	0.00614
DMM – Capacitance	0.2764	0.157
Digital Thermometer*	0.025764*	0.0854*
CRO – DC	0.682	0.654
CRO – AC	0.8099	0.678
CRO – Time	0.376	0.17925
Signal Generator	0.00148	0.00015
Pressure Gauge – Increasing	0.044036	0.058
Pressure Gauge – Decreasing	0.05248	0.03233

It can be observed from the residue comparison of sophisticated instruments that except for Digital thermometer (*) in all other cases residue with artificial neural network technique is less than regression analysis technique. This indicates the superiority of ANN model in terms of accuracy of prediction.

Comparison of the prediction results of regression based model and ANN based model for sophisticated instruments and working standards used in this study shows that, the ANN based model is more accurate. At the same time, one requires the MATLAB package and the model developed using ANN on a computer, to predict the measurement error (this is expensive and requires greater expertise to use). But in the case of regression model, just by substituting the value of independent variable x in the regression equation will yield the dependent variable Y, which is the measurement error to be estimated. For example in the regression equation of a digital thermometer given below:

$Y = 0.1289\ x + 0.6115;\qquad R^2 = 0.9963$ for 500°C

$Y = 1.7716$ for $x=9$. That means the digital thermometer will give a reading of 501.7716°C for a true temperature of 500°C after 9 years.

Hence it is worth noting that the regression model is simple and easy to use. Therefore, based on the requirement an appropriate model needs to be chosen. If the relationship between x and y is non-linear, regression analysis can only be successfully applied if prior knowledge of the nature of the non-linearity exists. Because, the range of study has to be divided into sections of linear segments, and regression equations made for each such segment. On the contrary, this prior knowledge of the nature of the nonlinearity is not required for ANN. In ANN, the degree of non-linearity can be also changed easily by changing the transfer function and the number of hidden layer nodes. Hence when linear relationship of the change in error with time is not seen it is safer to use ANN model for prediction.

7.5 HUMAN RELATED MEASUREMENT ERROR PREDICTION MODEL

ANN model used for predicting calibration errors were modified and prediction models for estimating human related measurement errors were developed. ANN model is developed to predict the probable errors due to the identified factors on human related measurement errors. Factors that have been used for modeling are experience, intelligent quotient, age, instrument differences and time of work. Feed-forward neural networks are used as before. Weight-age factors for the output layer of the error prediction model for human related measurement errors are given in *Table* 7.2 (only half of the *table* is shown).

Table 7.2 Weight-Age Factors for the Output Layer [20 by one is reduced to 10 by one]

Experienced	-0.29955	0.257678	-0.31	0.128832	-0.15527	0.574401	0.681581	0.23159	0.004816	-0.64912
IQ	-0.34275	0.325156	-0.21047	-0.11961	-0.54294	0.651265	0.36477	0.496829	0.492116	-0.61665
Age	-0.14374	0.135488	0.095413	-0.65776	-0.25275	-0.00641	0.291201	0.838691	1.0926	-1.47869
Instrument Differences	-0.39471	0.290607	-0.12279	-0.16093	-0.57177	0.512582	0.301061	0.514837	0.711673	-0.8694
Time of work	0.585989	-0.37959	-5.15308	0.62126	1.037867	1.430809	0.266766	1.188384	1.432843	-3.17839

Weight-age factor in this model depends on the number of neurons. Since the number of neurons in the input layer is twenty four (M=24) and the number of neurons in the hidden layer is (N=20), then the weight-age matrix of the hidden layer is 24 x 20. Similarly the weight-age matrix of the output layer is 20 x 1. Part of the 20 x 1 weight-age matrix [10 x 1 as shown in *Table* 7.2] of output layer of five different features were shown in the *Table 7.2* as a sample.

7.5.1 Details of Prediction Model of Human Related Measurement Errors Using ANN

Prediction model for human related measurement errors was developed using the observations from experiment on human related measurement errors (already discussed in Chapter III). Observations (data) from the experiments on measurement errors occurring with subjects, with known factor values for, experience, age, IQ, instrument difference and time of work, who did the measurement, was used in the model. Training and testing of artificial neural network model developed using MAT-LAB is able to predict the human related measurement errors due to selected human factors such as experience, age, IQ, instrument difference, time of work. *Table* 7.3 shows the actual and predicted values of human related measurement errors.

Table 7.3 Human Related Measurement Errors Actual (From Experimental Results) Vs Predicted Using ANN

Human Errors - ANN											
	Experienced Vs. Inexperienced [%]		Intelligent Quotient [%]			Instrument Differences [%]		Age [%]		Time of Work [%]	
	ET	*IE*	*IQ1*	*IQ2*	*IQ3*	*Analog*	*Digital*	*31-40*	*41-50*	*FN*	*AN*
Actual	0.000925	0.0017	0.003617	0.00555	0.00789	0.003617	0.00064	0.000653	0.000954	0.00565	0.008383
Predicted	0.000928	0.00175	0.003718	0.00568	0.008045	0.003718	0.000659	0.000665	0.000969	0.00571	0.009489
Residue	3E-06	5E-05	0.000101	0.00013	0.000155	0.000101	1.9E-05	0.000012	0.000015	6E-05	0.001106

i. Experience

A three layer feed-forward network with twenty four inputs and twenty sigmoid hidden neurons and a linear output neuron was created. The data set contains 2400 samples. 1920 samples were used in training the network while 480 samples were used in testing the network. Training is done using scaled conjugate gradient back propagation network. The scaled conjugate gradient algorithm (SCG) developed by Moller (1993) was designed to avoid the time consuming line search.

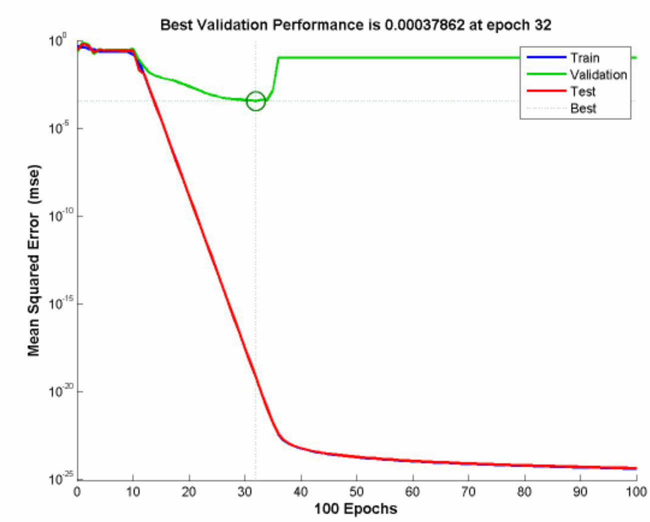


Figure 7.16 Prediction of Human related errors on Measurement - Experienced and Inexperienced

This algorithm combines the model-trust region approach with the conjugate gradient approach. Results of the error prediction model for predicting errors due to experienced and inexperienced subjects, showed very good abilities of the network to learn the patterns.

Results are good as can be seen from *Figure*s 7.18 and 7.19. Best validation performance is 0 .00037862 [mean square error] at epoch 32 as shown in *Figure* 7.16. Mean squared error (MSE) is the average squared difference between outputs and the target. Lower values are better while zero means no error.

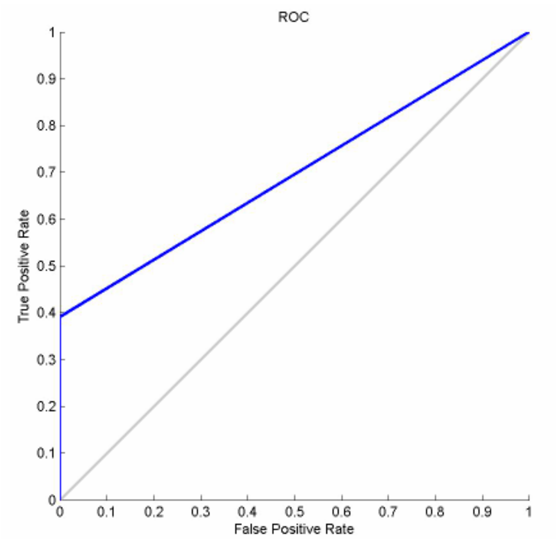


Figure 7.17 Receiver Operating Characteristic Curve

Receiver operating characteristic is used to inspect the prediction performance more closely as shown in *Figure* 7.17. By definition, ROC curve shows true positive rate Vs false positive rate (equivalently, sensitivity Vs 1-specificity) for different thresholds of the predicted output. One can use ROC, for example, to find the threshold that maximizes the prediction accuracy or to assess, in more broad terms, how the model performs in the regions of high sensitivity and high specificity.

Data sets containing a total of 2400 samples each were used for training and testing ANN to predict human related measurement errors in the case of experienced technicians (ET) and inexperienced (IE) subjects. Out of that 1920 samples (80% of the total samples) each were used in training the network while 480 samples (20%) were used in testing the network. It can be observed from *Table* 7.3 and *Figure* 7.18 and 7.19 that the human error prediction model could accurately predict the values of errors in the case of both experienced technicians and inexperienced subjects. Residue in ET is only 0.000003 and in the case of IE it is 0.00005.

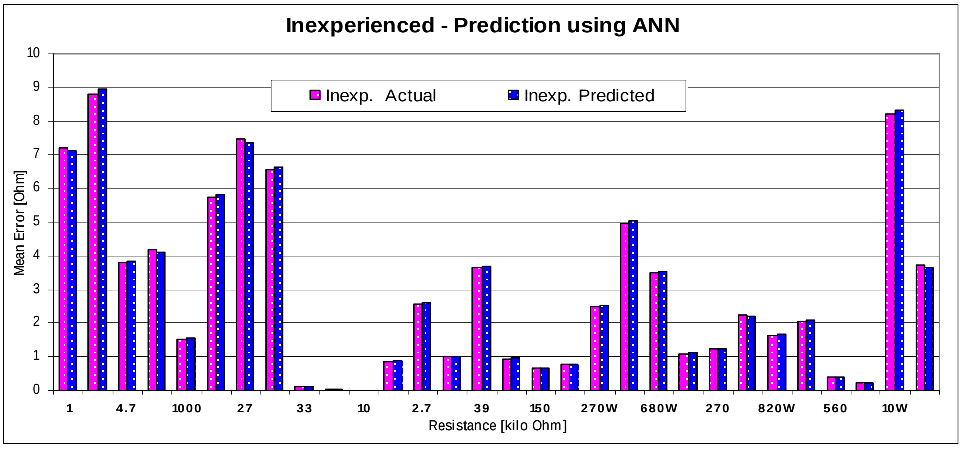


Figure 7.18 Prediction of human related measurement errors – Inexperienced subjects

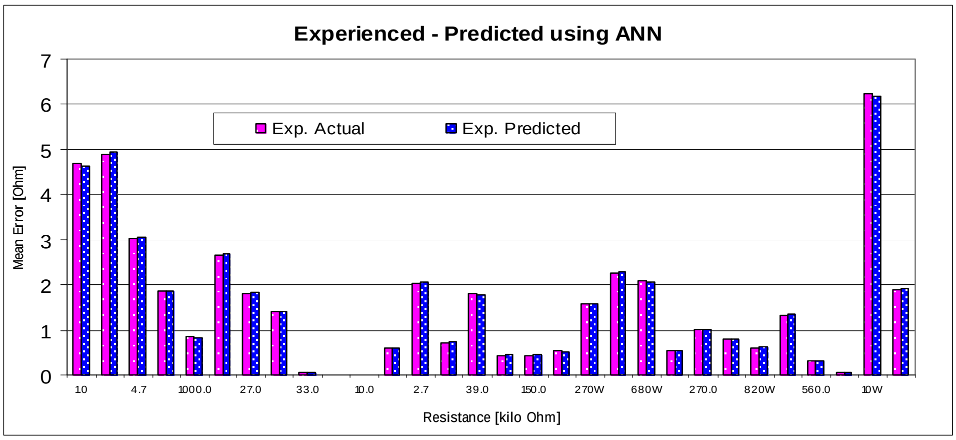


Figure 7.19 Prediction of human related measurement errors – Experienced Technicians

ii. Intelligent quotient

In this case the ANN prediction model was trained to predict the human related measurement errors in subjects with above average IQ (IQ1), average IQ (IQ2) and below average IQ (IQ3). Data set contains a total of 3600 samples. 2880 samples were used for training the ANN while 720 samples were used fro testing the network. *Table* 7.3 shows that, the prediction results are reasonably accurate with residues 0.000101, 0.00013 and 0.000155 for IQ1, IQ2 and IQ3 categories respectively. It is observed that best validation performance is 0.079537 at epoch 7.

iii. Instrument Differences

Experimental analysis on one of the work related human factor, instrument difference, was carried out by asking the subjects to make measurements using analog and digital devices (details are given in chapter III). In the prediction error model, the ANN was trained using 2880 samples and 720 samples were used for testing the network. Training of the neural network was very effective as is evident from *Table* 7.3 Residue is 0.000101 for measurement using analog instrument and it is only 0.000019 for measurement using digital instruments. Best validation performance is 2.7139e-25 at epoch 100. Since the MSE which is the mean squared difference between outputs and the expected outputs is very close to zero the result is very accurate.

iv. Age

Estimation of human related measurement errors in subjects who belongs to different age groups is done by training and testing ANN prediction model using the experimental data of subjects in the age group of 31 to 40 years and 41 to 50 years. Difference between actual and predicted values of human related measurement errors (residue) is only 0.000012 for the age group 31 to 40 years were as it is 0.000015 for the subjects in the age group of 41 to 50 years. Best validation performance is 8.0189e-25 at epoch 100. Since the mean square error is close to zero the result is close to the actual value. This shows that the prediction model is quite efficient.

v. Time of Work

Transient qualities of human being such as fatigue, boredom etc. directly depend on his or her time of work. Prediction error model of ANN was trained and tested firstly by using forenoon session measurement error data then, using the afternoon session measurement errors. *Table* 7.3 gives the residue for the forenoon session as 0.00006 and the residue for afternoon session is 0.001106. Best validation performance is 0.057823 at epoch 65.

7.6 CONCLUSION

In order to make the data collected more useful, Regression and Artificial Neural Network [ANN] based models have been developed to predict error [extent] for instrument type and standard types studied. Prediction models were also developed for human related measurement errors. These can be used as ready reference by practitioners wanting to estimate or predict measurement errors. The comparison of the prediction results of regression based model and ANN based model shows that the ANN based model is more accurate. It is also worth noting that the regression model is simple and easy to use. Therefore, based on the requirement an appropriate model needs to be chosen.

Chapter VIII

Summary, Conclusion and Scope for Future work

8.1 SUMMARY

Instruments used in daily course of work are designed to provide reliable and accurate (though extent of precision depends on requirement and the measurand) measurement while at the same time being rugged enough to withstand the harsh operating environment. Test engineers and professionals in the field of measurement face a unique set of challenges, when they perform measurement. A systematic (clearly defined process) and systemic (all encompassing) approach is needed to identify every source of error that can arise in a given measuring system. It is then necessary to determine their magnitude and impact under the prevailing operational conditions. Measurement related errors can only be defined in relation to a real specific measurement task. Systematic error or bias is a permanent deflection in the same direction from the true value. It can be easily corrected. Bias and long term variability are controlled by monitoring measurements against a check standard over time. Systematic errors or bias are repeatable errors existing with the specified source; these can be adjusted out or compensated for, by periodical calibration of the instruments. Random error is a short-term scattering of values around a mean value. It cannot be corrected on an individual measurement basis. Random errors are studied by statistical methods and expressed in statistical terms.

Measuring devices are of different types such as mechanical, electrical, electronic etc. The electronic measuring instrument can be analog or digital. Measurement error, which is the difference between the distorted information and the undistorted information about a measurand, caused due to the measuring device and/or the method and/or the person involved in the measurement. However, research work on the experimental analysis of the effect of various factors on human errors was not found. Survey among test engineers and professionals in the area of measurement revealed that, they know that human related measurement errors exist and there is a possibility for calibration error growth with ageing, but they do not have an idea about the factors that influence human related measurement errors and the extent of human related measurement errors and the quantum of error growth with ageing in sophisticated instruments and working standards used in legal metrology.

Survey of literature on measurement errors showed that the application of error reduction techniques has enabled the manufacturing and service industries to gain better control over the processes and expenses, and also to improve customer satisfaction and thus strengthen their level of competitiveness. Therefore, it is evident that a clear understanding of the critical factors of measurement errors in the method and the measurement would help the implementation of management practices to improve productivity and enable better control over processes, quality and costs. This research has brought out the critical factors of human related measurement errors and also contributed a model for prediction of human related measurement errors using ANN. The contributions are given in detail in the next section.

8.2 RESEARCH CONTRIBUTIONS

The contributions of this research work are summarized as follows:

- Professionals and practitioners know that, there exist human related measurement errors. They also agree with the general 'say' about human related measurement errors, for example, human related measurement error possibility is more with analog measuring instrument than digital, fatigue will be more during afternoon, performance of human being depend on their

IQ etc. From the pre-experiment survey research studies, it is quite clear that the practitioners and professionals could not quantify the possible extent (values) of human related measurement errors or error reduction or error increase due to the influence of various factors. Based on the experimental analysis, it was also observed that when the modified survey instrument, wherein the choices of answers given to the questions related to the quantification of errors, it helped to improve the perception of the practitioners and professionals to a great extent.

- Human related error in measurement is an ever present, generally significant but not a much empirically studied phenomenon. Human related measurement errors depend on many factors. In this experimental research work, an attempt has been made to identify these factors and characterize them as person related and work related factors. The upshot of a manual measuring system is a combination of the results produced by the measuring device and what was observed and noted by the human subject involved in the manual measurement exercise. This study focused on the effect of a few person related factors such as experience, training, intelligent quotient, age and gender and work related factors such as type of instrument used, time of work, task, time pressure and environmental differences, when measurement was done.

- Mean percentage of human related measurement error in each of the above said factors were found out from the experimental study. It was necessary to study and bring out these values, since knowing the values would help in dealing with the errors when they are significant in the system. This study brings out values of human related measurement errors under different person and work related factor combinations. These may be taken as expected error values under those conditions for work system design. The practitioners can use these results to design measurement systems (people included) with the error within the given acceptable limits.

- In the decreasing order of influence on human related measurement errors, the person related factors are: Experience, Training, Intelligent Quotient, Age and Gender. The ascending order of influence of work related factors on human related measurement errors are instrument differences [digital, analog], working environment, time pressure, time of day and type of work. Identification of these factors and an assessment of the extent of errors they produce will be of use to practitioners who rely on measurement for research, control, production, and quality control. This work has also demonstrated a simple methodology that can be used for such work. This work was carried out with only few parameters, since taking too many parameters all at a time would make the experiment very difficult to conduct and control. It is hoped that this work will help users of measurement in practice to better understand and manage the phenomena of Human related measurement errors in measurement.
- In the time based characterization of errors observed during calibration of sophisticated instruments and the standards used in legal metrology, Annual Average Growth (AAG) of error and instrument to instrument error difference were found out. It is observed that instrument error grows with ageing and there is instrument to instrument error growth rate difference. Instrument calibration is the act of determining the uncertainty associated with the instrument's measurement. If feasible, the instrument is adjusted to reduce the overall uncertainty associated with the measurement. Calibration quantifies the equipment's time in service, instrument to instrument

differences, temperature, humidity, environmental exposure, and abuse and when required, adjusts the device's measurement capability to decrease error. The study describes how the errors in different equipments are described today. It also discusses how this description is used to compute the effect of the errors upon the measurement result, and how this effect is reduced or even minimized by rational use of the observations and by experimental design.

- In this research work, the calibration data of Digital Multimeter and digital thermometer were used to determine the uncertainty budget through a mathematical model. Controlling errors is an essential part of instruments and instrumentation systems. The error control begins in the design stages by choosing the appropriate components, filtering, and bandwidth selection, by reducing the noise, and by eliminating the errors generated by the individual sub-units of the complete system. In a good design, the errors of the previous group may be compensated adequately by the following groups. The accuracy of instruments can be increased by post-measurement corrections. Various calibration methods may be employed to alter parameters slightly to give correct results. In many cases, calibration graphs, mathematical equations, tables, the experiences of the operators, and the likes are used to reduce measurement errors. The application of digital techniques and intelligent instruments, error corrections are made automatically by computers or the devices themselves.
- Estimation of errors was done using regression analysis. Regression Analysis on the calibration data of various sophisticated instruments and working standards used in legal metrology were done and regression curves were plotted and the regression equations give the estimate of errors. The statistical

significance of the regression analysis has been verified by determining the R^2 value in each case. The regression analysis curve can be used as a benchmark for the error growth with ageing of the instrument. Verifying authorities can fit their observed values in the regression curve and if these values are well above the regression curve, the instruments can be rejected.

- ANN based models were also used for the estimation of errors. The code written with MATLAB –NEURAL NET software has been used to predict errors of several sophisticated instruments and working standards used in legal metrology. Training has been optimized by varying calibration data size. ANN model was developed for estimating human related measurement under different factor combinations. These can be used as ready reference by practitioners wanting to estimate or predict measurement errors. The comparison of the prediction results of regression based model and ANN based model shows that the ANN based model is more accurate. However, it is also worth noting that the regression model is simple and easy to use. Therefore, based on the requirement, an appropriate model needs to be chosen. If the relationship between x and y is non-linear, regression analysis can only be successfully applied if prior knowledge of the nature of the non-linearity exists. Because, the study range has to be divided into sections of linear segments, and regression equations were made for each such segment. On the contrary, this prior knowledge of the nature of the nonlinearity is not required for ANN. In ANN, the degree of non-linearity can be also changed easily by changing the transfer function and the number of hidden layer nodes. Hence when linear relationship of the change in error with time is not seen it is safer to use ANN model for prediction.

- Characterization and estimation of measurement errors would also enable to quantify:
  - the instrument's time in service.

    The observations during calibration can be applied in the error models and if the error growth is abnormal, appropriate error correction technique can be used or the instrument can be rejected. Eg: Test engineers can fit the new calibration data in the regression curve and if it exactly follows the regression curve developed in this study using the previous year's calibration data, then the instrument under calibration can be accepted.
  - Period of calibration.

    An accurate estimation of Annual Average Growth (AAG) of errors from the available calibration observations of instruments and also by prediction of errors using the appropriate models would help to schedule calibration depending on requirement, and not always follow thumb rules like calibrate after twelve months.
  - Error reduction / compensation.

    Knowledge of the extent of error growth with instruments and the instrument to instrument error growth differences would help the test engineers and designers to apply appropriate error reduction and compensation techniques in instruments.
- The legal metrology department is handling many types of working standards, which varies from the working standard for weighing precious metals and up to weigh bridge. They follow uniform procedures for the conformity study of various types of working standards. The findings of this

study of uncertainty can be incorporated for the unsurpassed conformity, at least, for the instruments used for sophisticated measurements.

8.3 LIMITATIONS OF THE WORK

- The study was limited to errors in instrument and errors in method and measurement. In error in method and measurement, the critical human factors which influence measurement errors were identified and they were analyzed using simple experiments. The experimental results were used for error modeling using ANN. The study was limited to only certain factor combinations. The other limitation of the study is that limited number of subjects was asked to do only certain measurements using selected type of instruments. In this research work the study was conducted in ordinary and special laboratory environments; industrial environment is not used.
- Characterization of calibration errors in terms of only time was done in this research. Only simple regression and ANN were used for error prediction in this study. Complex time series analysis and other prediction /forecasting tools were not explored in this study. Characterization is done only by collecting the calibration data from selected testing and calibration laboratories. It was beyond the scope of this study to analyze the parameters which creates errors in a particular or some selected instruments. Some error reduction techniques were suggested in this study, but a detailed error reduction techniques study was not conducted especially in the challenging environments.
- Hand book or directory showing complete table of human related measurement errors due to various factor combinations and the complete equations and curves of error prediction models based on regression and

ANN are not brought out, even though the reasonable quantity and quality of data has been collected and conclusions derived in this study.

8.4 SCOPE FOR FUTURE WORK

Study can be extended to identify more human factors like climatic conditions, nationality/race, emotional status/quotient of the person, extent of training/refresher training etc which influence measurement error. Type of instruments and measurements can also be increased in number. The knowledge acquired from the experimental analysis on human related measurement errors and the instrument developed for estimating human errors can be utilized for real time human error study in few other selected industries, so that it would be possible to have a cross verification of the study and thus improve the results. It would be interesting to study the possible improvement in measurement, when the practitioners in industry, who make the measurement are given the feedback (based on the study of experimental analysis on human measurement errors) about the possibility of errors in the given work condition. Errors of instruments in use and errors in practice can be studied by making a comparative study of the measurement data obtained through the calibrated standards and the actual calibrated instrument (instrument used by that particular industry for measurement) in an industrial environment.

The reason and the extent of errors in various instruments can be studied and simulated. For example, stacking error study in attenuator pads of signal generator and its optimization would reduce calibration errors in signal generator. The goal of system calibration is to quantify and compensate for the total measurement error in the system. Manual calibration can be time-consuming, costly and error prone due to human interventions. Automated system can be developed and its effect can be studied. Techniques to reduce measurement errors can also be developed, for example a comparative study of centralized approach of measurement and distributed measurement system can be done, and an ideal method of measurement can be suggested. While this study has brought out much new knowledge and insights, there is a need to carry out multi-dimensional and time-series studies to create universal models and precise prediction models.

REFERENCES

1. A. Messer, Philippe Bernadat, Guangrui Fu, Deqing Chen, Zoran Dimitrijevic "Susceptibility of commodity systems and software to memory soft errors," *IEEE transactions on Computers*, Vol. 53, No. 12, pp 1557 to 1568, 2004.
2. Carter J.A., Jr., "A taxonomy of user – oriented functions ", *International Journal of Man – Machine studies*, 24, pp 270 – 286, 1986.
3. B. J. Hoopes, K. P. Triantis, "Efficiency Performance, Control Charts, and Process Improvement: Complementary Measurement and Evaluation," *IEEE Trans. on Engineering Management*, vol. 48, no. 2, pp. 239–253, 2001.
4. M. Wang, P. D. Hale, K. J. Coyakley, and T. S. Clement, "Uncertainty of Oscilloscope time based Distortion Estimate", *IEEE transactions on Instrumentation and measurements*, Vol. 51 No. 1, pp 53 -58, 2002.
5. Parasuraman, R., Sheridan, T. B and Wickens., (2000). "A model for types and levels of human interaction with automation," *IEEE Transactions on Systems, Man, and Cybernetics*, A30 (3), pp. 286-295, 2002.
6. Chesher, A., "The effect of measurement error", *Biometrika*, Vol. 78, No.3, PP. 451 – 462, 1991.
7. Cannon M. D. , and Amy C. Edmondson, "Confronting failure: Antecedents and consequences of shared beliefs about failure in organizational work groups", *Journal of Organizational Behavior*, 22, PP 161 – 177, 2001.
8. Clarke, S., "Organizational Factors Affecting the Incident Reporting of Train Drivers". *Work & Stress* 12**:** 6-16, 1998.
9. Davies, J.B., Wright, L., Courtney, E. & Reid, H., "Confidential incident reporting on the UK railways: The CIRAS system", *Cognition, Technology & Work* 2: pp.117-125, 2000.
10. Helmreich, R. & Foushee H., "Why crew resource management? Empirical theoretical bases of human factors training in aviation". In E. Wiener, B. Kanki, & R. Helmriech (Eds), *Cockpit resource management*. pp 3-45 San Diego, CA: Academic, 1993.
11. International Vocabulary of Basic and General Terms in Metrology*: BIPM, IEC, IFCC, ISO, IUPAC, IUPAP, OIML, 1993*
12. J. E. Decker, R. Schodel, G. Bonsch, "Considerations for the evaluation of measurement uncertainty in interferometric gauge block calibration applying methods of phase step interferometry," Metrologia 41, pp. L11–L17, 2004.
13. J. P. Bentley, "Principles of Measurement Systems", 2nd ed., Burnt Mill, *UK: Longman Scientific and Technical*, 1988.
14. J.G.M.Grinten, Confidence levels of measurement based decisions, *Proc. of 11th Flomeko Conf. Groningen, Netherlands*, pp.1-10, May 2003.
15. J. Verspecht and K. Rush, "Individual characterization of board-band sampling oscilloscopes with a nose-to nose calibration procedure", *IEEE transactions in instrumentation and measurements* , vol. 43, no.2, pp 347 -354, 1994.

16. K. Berka, "Are there objective grounds for measurement procedures?" C.W. Savage and P. Ehrlich (Eds.), *Philosophical and Foundational Issues in Measurement Theory*, Hillsdale, NJ: Lawrence Erlbaum, pp. 181-194, 1992.
17. K. G. Birch, "Uncertainties in the Measurement of Gauge Blocks by Interferometry," *National Physical Laboratory (NPL)*, Teddington, Middlesex, UK, Report MOM 29, May 1979.
18. K. P. Birch, F. Reinboth, R. E. Ward, G. Wilkening, "The Effect of Variations in the Refractive Index of Industrial Air upon the Uncertainty of Precision Length Measurement," *Metrologia, no.* 30, pp. 7–14, 1993.
19. Joshi, A., Ramakrishman, N., Houtis, E.N., and Rice, J.R., "On neurobiological, neuro-fuzzy, machine learning, and statistical pattern recognition techniques", *IEEE Transactions on Neural Networks*, no. 8: pp18-31, 1997.
20. Helmreich, R. L., "Safety and Error Management: the Role of Crew Resource Management", *Ashgate Publishing*, Aldershot, England, pp. 107 – 119, 2000.
21. Lee, T. & Harrison, K., "Assessing Safety Culture in Nuclear Power Stations". Safety Science, 34: pp 61-97, 2000.
22. Lippmann ,R.P., "An introduction to computing with neural nets" *IEEE ASSP Magazine*; vol.4, no.2, pp.4 - 22, 1987.
23. Mahnaz Mohammadi, "Diagnosing and correcting systematic errors in Spectral based digital imaging", *13th Colour imaging conference final program and proceedings*, pp 25 to 30, 2008.
24. Masayoshi Koike, "Instrument and Control", 37-5, pp. 312-317, 1998.
25. Rogosa, D, R, and Willett, J. B.), "Demonstrating the reliability of the difference score in the measurement of change", *Journal of Educational Measurement*, no. 20, pp. 335 -343, 1983.
26. S.Bell, "*Measurement Good Practice Guide* No.11", *National Physical Laboratory*, 1999.
27. Toshiro Masui, "Consideration on accuracy for deciding instrumental errors in inspection of verification standard", *Institute Symposium on accuracy JCSS committee,* 1969.
28. F. Williams. P. D. Hale, T. S. Clement and J. M. Morgan, "Calibrating Electro-optic sampling systems", *in IEEE MTT –s Instrumentation – Microwave Symposium*, vol 3, pp 1527 – 1530, 2001.
29. Dylan F. Williams, "Systematic Error of the Nose-to-Nose Sampling - Oscilloscope Calibration", *IEEE transactions on Microwave Theory and Techniques*, vol. 55, No.9, pp 1951 – 1957, 2007
30. Baxter and Bass. "Human Error Revisited: Some Lessons for Situational Awareness," *Fourth Symposium on Human Interaction with Complex Systems*, March 22-24, pp.81-87, 1998.
31. Blanchard R. E., "Requirements, concept, and specification for a navy human factors performance data store", *International Journal of Man – Machine studies*, no. 31, pp 643 – 672, 1973.
32. David Embrey, "Understanding Human Behavior and Error", *Human Reliability Associates 1, School House, Higher lane, Dalton, Wigan, Lancashire*, WN8 7RP, 1996.

33. Frese, M., “Error Management or error prevention: two strategies to deal with error in software design”, in H-J Bullinger (Ed.), *Human Aspects in Computing, Design and use of Interactive Systems and Work with Terminals* (pp. 776 – 782), Amsterdam: Elsevier, 1991
34. Holland P., “Statistics and Causal Interface”, *Journal of the American Statistical Association, v*ol. 81, no. 260, pp 663 – 685, 1986
35. T.J. Hebert, R. Leahy, “Statistic-based MAP image reconstruction from Poisson data using Gibbs prior”, *IEEE Trans. Signal Processing,* 40 (9) 2290–2303, 1992.
36. V.Daniela, and I. Castanheira, “Uncertainty budgets and MPES in refractometry: A project study”, *OIML BULLETIN XLI-4*, pp. 8-12, 2000.
37. A. Kolaczkowski, J. Forester, E. Lois, and S. Cooper, "Good Practices for Implementing Human Reliability Analysis (HRA)”, *NUREG-1792, Washington, DC: US Nuclear Regulatory Commission*, 2005.
38. Zapt D., Brodbeck F. C., “Errors in working with computers: A First Validation of a Taxonomy for Observed Errors in a Field Setting”, *International Journal of Human – Computer Interaction* No. 4, pp 311 – 339, 1992.
39. Dormann, T., and Frese M., “Error training: Replication and the function of exploratory behavior”, *International Journal of Human Computer Interaction*, no. 6, pp. 365 – 372, 1994.
40. Douglas A. Wiegmann, and Esa Rantanen, “Defining the Relationship Between Human Error Classes and Technology Intervention Strategies”, *Aviation Research Lab, Institute of Aviation, University of Illinois at Urbana – Champaign*, 1 Airport Road, Savoy, IL 61874, 2002.
41. Duncan, K. D., “Fault diagnosis training for advanced continuous process installations”. *In: Rasmussen, J., Duncan, K., and Leplat, J. (Eds), New Technology and Human Error*. Chichester: Wiley, 1987.
42. Feyer, A.M. and Williamson, A.M., “Human factors in accident modeling”, *In: Stellman, J.M. (Ed.), Encyclopedia of Occupational Health and Safety*, Fourth Edition. Geneva: International Labour Organization, 1998.
43. Peccei, R. and Thomas, A., “Safety culture and safety performance: British Rail in the aftermath of the Clapham Junction disaster”, *Paper presented at the Bolton business school conference on changing perceptions of risk*, Bolton, 1994.
44. Carmen Simion, “The Study of Measurement Equipment Bias” *University Lucian Blaga from Sibiu,* Faculty of Engineering, Hermann Oberth Romania, 2005
45. Cathy Van Dyck, Michael Frese, Markus Baer and Sabine Sonnentag, “Organizational error management culture and its impact on performance: A two-study replication”, *Journal of applied psychology*, vol. 90, no. 6, 1228 – 1240, 2005
46. Frese, M., “*Error Management in training: Conceptual and Empirical Results*.” In C, Zucchermaglio, S. Bagnara, and S. Stucky (Eds.), “Organizational learning and technological change” pp 112 – 124, Berlin, Germany: *Springer – Verlag*, 1995.
47. C. M. Wang, P. D. Hale, K. J. Coyakley, and T. S. Clement, “Uncertainty of Oscilloscope time based Distortion Estimate”, *IEEE transactions on Instrumentation and measurements*, vol. 51 no. 1, pp 53 -58 , 2002.

48. Douglas A. Wiegmann, et. al., “Human Error perspectives in Aviation”, *The International Journal Of Aviation Psycology* , vol. 11, no. 4, pp 341 – 357, 2000.
49. Gawron V. J., et. al., (1989), “A taxonomy of independent variables affecting human performance” *International Journal of Man – Machine studies*, 31, PP 643 -672, 1989.
50. H. Castrup, “Uncertainty Analysis for Risk Management,” *Proc. Measurement Science Conference*, pp. 338–364, Anaheim, CA–USA, Jan. 1995.
51. D. F. Williams. P. D. Hale, T. S. Clement and J. M. Morgan, “ Calibrating Electro-optic sampling systems”, *IEEE MTT Instrumentation – Microwave Symposium*, vol 3, pp 1527 – 1530, 2001.
52. C.F. Dietrich, “Uncertainty, Calibration and Probability”, *Adam-Hilger, Bristol*, 1991, p.p. 564 .
53. Christian Hone, “Error Propagation after Concealing a lost Speech frame”, *Proceedings on Multicomm*, pp 7 -12, 2006.
54. Wickens C. D., Gordon S. G. and Liu Y., (1998), “An introduction to human factors engineering. New York” : *Addison Wesley Longmann*, 1998
55. Jens Rasmussen, “Skills, Rules, and Knowledge; Signals, Signs, and Symbols, and Other Distinctions in Human Performance Models”, *IEEE Transactions on Systems, Man, and Cybernetics*, vol. SMC-13, no.3, 1983.
56. Mark A. Huselld, Brian E. Becker, “Comment on measurement error in research on human Resources and firm performance: How much error is there and how does it influence effect size estimates?” by Gerhart, Wright, Mc Mahan and Sneha”, P*ersonal Psychology*, Inc. 53, 2000.
57. R.L. Boring, “Improving human scaling reliability,” *Proceedings of the Human Factors and Ergonomics Society* 47th Annual Meeting, pp. 1820- 1824, 2003.
58. Nordstrom, C. R., Wendland D., and Williams K. B., “To err is Human. An Examination of the Effectiveness of Error Management Training”, *Journal of Business and Psychology*, 12 , pp 269 – 282, 1998.
59. R.L. Boring and R.L. West, “Mind as magnitude: Reconsidering information processing in cognitive engineering,” *Proceedings of the 49th Annual Meeting of the Human Factors and Ergonomics Society*, pp. 1826-1830, 2005.
60. Meijman, T. F., Mulder G., “Psychological aspects of workload” in P. Drenth .H. Thierry, and C. De Wolff (Eds.). *Hand Book of Work and Organizational Psychology* (2nd ed., Vol., pp – 5-33). 1998.
61. Hollnagel, E., “Cognitive reliability and error analysis method (CREAM)”, *Oxford, England*, Alden Group, 1998.
62. Shappell, S., & Wiegmann, D.,“Applying Reason: The Human Factors Analysis and Classification System (HFACS)”. *Human Factors and Aerospace Safety, 1,* 59–86, 2001
63. Wickens C., & Flach J., “Information processing”. In E. Wienner & D. Nagel (eds.), *Human factors in aviation*, pp 111-115. San Diego, CA: Academic, 1988.
64. Rasmussen, J., “Human errors: A taxonomy for describing human malfunction in industrial installations”, *Journal of Occupational Accidents*, Vol. 4, pp 311-333, 1982

65. Heinrich, H., Peterson, D., & Roos, N.,"*Industrial accident prevention: A safety management approach"* (1st ed.). New York: McGraw-Hill, 1981.
66. Kayten, P., "The accident investigator's perspective". In E. Wiener, B. Kanki, & R. Helmriech (Eds.), *Cockpit Resource Management* pp. 283- 283-314, San Diego, CA: Academic, 1993.
67. Senders J. W., "Human error: Cause, Prediction and Reduction", Hillsdale N. J. : Erlbaum, 1991.
68. Michael Bair, " Typical Pressure Measurement Uncertainty Defined by an FPG 8601 Forced Balanced Piston gauge.", *Technical note* by D. H. Instruments, pp 41-49, 2004.
69. T. Skwirczynski, "Uncertainty of the calibrating instrument, confidence in the measurement process and the relation between them", *OIML BULLETIN XLII-3*, pp. 5-10, 2001.
70. H.E. Kyburg, "Theory and Measurement", Cambridge, UK: Cambridge University Press, 1984.
71. Xin Li, " A memory soft error measurement on production systems", *USENIX annual Technical Conference*, PP 275- 280, 2007.
72. T. C. May and M. H. Woods, "Alfa – Particle – Induced soft errors in dynamic memories", *IEEE transactions on Electron devices*, vol. 26(1), pp 2 – 9, 1979.
73. D. F. Williams., T. S. Clement, Kate A Remley, P. D. Hale, Frans Verbeyst [2007], "Systematic Error of the Nose-to-Nose Sampling" - Oscilloscope Calibration "*IEEE transactions on Microwave Theory and Techniques*" vol. 55, no.9, pp 1951 – 1957, 2007.
74. Murari, "On the importance of considering measurement errors in a fuzzy logic system for scientific applications with examples from nuclear fusion", *Fusion Engineering and Design*, pp. 1 – 11, 2008.
75. R.L. Boring, D.I. Gertman, J.C. Joe, L.G. Blackwood, H.S. Blackman, and B.M. Brady, "A Simplified Expert Elicitation Guideline," *Proceedings of the 8th International Conference on Probabilistic Safety Assessment and Management*," Paper PSAM-0089, pp. 1-9, 2006.
76. Ernest O. Doebelin, "Measurement Systems Application and Design", *Tata McGraw Hill Publishing Company Ltd*., India, 5th Edn, 2005.
77. N.G. Paulter and D. R. Larson (2003), " Sources of Uncertainty in the nose-to- nose sampler calibration method", *IEEE transactions on instrumentation and measurements* vol. 52, no.4, pp 1618 -1626, 2003.
78. Oldham, Nile; Hetrick, Paul; Zeng, Xiangren, "A Calculable TransporTable Audio Frequency AC Reference Standard", *IEEE Transactions on Instrumentation and Measurement*. vol. 38, no. 2, pp. 368-371, April 1989.
79. M. T. Clarkson, T. Collins, and B. Morgan, "A combinatorial technique for weighbridge verification", *OIML BULLETIN XLIII-2*, pp. 5-12, 2002.
80. T. Lammi, "Calibration of weighing instruments and uncertainty of calibration", *OIML BULLETIN XLII-4*, pp.9-20, 2001.

81. Klaus-Dieter Sommer, Samuel E. Chappel, Manfred Kochsiek "Calibration and verification: Two procedures having comparable objectives and results", vol. 2, no. 13, pp. 117-125, 2000.
82. Leo M. Harvill, "Standard Error Of Measurement", *Journal of Educational Measurement*, pp 181-189, 1998.
83. EA-3/02, "The Expression of Uncertainty in Quantitative Testing, Ed. 1: European Cooperation for Accreditation" (EA), (previously EALG23), August 1996.
84. M. Buzoianu, "Measurement uncertainty and legal limits in analytical measurements", *OIML BULLETIN XLI-1*, pp.5-12, 2000.
85. Stanley, J., C., "Reliability In R. Thorndike, (Ed.), Educational Measurement", (2nd ed., PP 356-442) Washington, D. C. : *American council of Education*, 1971.
86. W. Schulz, K.-D. Sommer, "Uncertainty of Measurement and Error Limits in Legal Metrology", *OIML Bulletin XL* (4) (Oct. 1999) 5 – 15.
87. R. Schodel, J. E. Decker, "Methods to recognize the sample position for most precise interferometric length measurements," *Proc. of SPIE Conference Interferometry, Denver,* Colorado, August 2004.
88. Traub, R. E. & Rowley, G. L., "Understanding reliability Educational measurement : Issues and practice", 10 (1), 37 -45, 1991.
89. "National Institute of Technology and Evaluation, Japan Calibration Service System", "*Guidelines for evaluation of uncertainty in calibration*", 2002.
90. "Guide to the Expression of Uncertainty in Measurement (GUM)", 2nd ed. Geneva: *International Organization for Standardization*, ISO, ISBN 92-67-1018 pp 8-9, ISO 1995.
91. Geometrical Product Specification (GPS) – "Inspection by measurement of work-pieces and measuring equipment, Part 1: Decision rules for proving conformance or nonconformance with specification", ISO 14253 – 1: 1998, *International Organization for Standardization (ISO)*, Geneva, 1998.
92. Qeethara K. Al-Shayea, Ghaleb A. El-Refae, "Evaluating credit risk using Artificial Neural Networks", Global Engineers And Technologists Review, Vol. 1, No. 1. , pp 21 – 28, 2011.
93. Bahrammirzaee, A., Ghatari, A., Ahmadi, P. and Madani, K., "*Hybrid Credit Ranking Intelligent System Using Expert System and Artificial Neural Networks*", Applied Intelligence, Vol. 34 No.1, pp. 28-46, 2011.
94. Marcano-Cedeno, A., Marin-De-La-Barcena, A., Jimenez-Trillo, J., Pinuela, J. and Andina, D., "Artificial Metaplasticity Neural Network Applied to Credit Scoring", *International Journal of Neural Systems*, Vol. 21, No.4, pp. 311-317, 2011.
95. Kashman, A., "A Neural Network Model for Credit Risk Evaluation", International Journal of Neural Systems", vol. 19, no. 4, pp. 285–294, 2009.
96. Bahrammirzaee, A., "A Comparative Survey of Artificial Intelligence Applications in Finance: Artificial Neural Networks, Expert System and Hybrid Intelligent Systems", *Neural Computing & Applications*, vol. 19, no. 8, pp.1165-1195, 2010
97. Anzai, Y., 1992. "*Pattern Recognition and Machine Learning*". Academic Press, Boston, 1992.

98. Holmström, L., and Koistinen, P., "Using additive noise in back-propagation training", *IEEE Transactions on Neural Networks*, vol. 3: pp 24-38, 1992.
99. Tamura, S., and Tateishi, M., "Capabilities of a four-layered feed-forward neural network: four layers versus three", *IEEE Transactions on Neural Networks*, vol. 8: pp. 251- 255, 1997.
100. David Leverington, " A basic Introduction to Feed-Forward Back propagation Neural Networks", *Texas Technical university – Department of Geo-Sciences*. pp 1-33, 2012
101. Basheer,I.A. and M.Hajmeer; "Artificial neural networks : fundamentals , computing, design and application"; Journal of Microbiological Methods; vol.43; no.1, pp 1-12, December 2000.
102. Al-Shayea, Q.K., El-Refae, G.A. and El-Itter, S.F., "Neural Networks in Bank Insolvency Prediction", *International Journal of Computer Science and Network Security*, Vol. 10, No. 5, pp. 240-245, 2010 .
103. Anil K.Jain, Jianchang Mao, K.M.Mohiuddin, "Artificial Neural Networks: A Tutorial", I*EEE Computer*, Vol. 29, No. 3, pp 31- 44, 1996.
104. J. Nocedal and S.J. Wright, "Numerical Optimization," *Springer*, New York, 1999.
105. J. Nunnally, "Psychometric Theory" (2nd Ed.), New York: McGraw-Hill, 1978.
106. "American Psychological Association, Standards for Educational and Psychological Testing", Washington, DC: *American Psychological Association*, 1985.
107. Bishop, C.M., "Neural Networks for Pattern Recognition"; Oxford University Press, New York, 1995.
108. Minsky, M., and Papert, S., "Perceptrons". MIT Press, Cambridge, 1969
109. J. C. Williams, "A Data-Based Method for Assessing and Reducing Human Error to Improve Operational Performance", I*EEE Fourth Conference on Human Factors and Power Plants*, 1988 .
110. R.L. Boring, T.Q. Tran, D.I. Gertman, and A. Ragsdale, "A human reliability based usability evaluation method for safety-critical software," Proceedings of the 5th *International Topical Meeting on Nuclear Plant Instrumentation, Controls, and Human Machine Interface Technology* (NPIC&HMIT), pp. 1275-1279, 2006.
111. Rasmussen, J., "*The human as a systems component*", In: Smith, H.T. and Green, T.R.G. (Eds), Human Interaction with Computers. London: Academic Press, 1980.
112. Reason J., "*Human Error*", Cambridge, England : Cambridge University Press, 1990
113. W.J. Galyean, "Orthogonal PSF taxonomy for human reliability analyses," Proceedings of the 8th International Conference on Probabilistic Safety Assessment and Management," Paper PSAM-0281, pp. 1-5, 2006.
114. Wickens, C. D.; Sandy, D. L.; Vidulich, M., "Compatibility and resource competition between modalities of input, central processing, and output". Human Factors (Santa Monica, CA, ETATS-UNIS: Human Factors and Ergonomics Society) 25 (2): 227–248, 1983.
115. Wiegmann D., and Shappell S, "Human factors analysis of post- accident data: Applying theoretical taxonomies of human error", *The International Journal of Aviation Psychology*, 7, PP 67 – 81, 1997

116. V.J. Gawron, C.G. Drury, S.J. Czaja, and D.M. Wilkins, " A taxonomy of independent variables affecting human performance" *International Journal of Man – Machine studies*, 31, PP 643 -672, 1989.
117. Philip Carter, "IQ and Psychometric Test Workbook", *Kogan Page India Pvt. Ltd*, 2008
118. Website : *http://www.allthetest.com/iq*test
119. Akkerboom, H., and Dehue, F., "The Dutch Model of Data Collection Development for Official Surveys," *International Journal of Public Opinion Research*, Vol. 9, 126-145, 1997.
120. Belson, W.R., "The Design and Understanding of Survey Questions", Aldershot, England: Gower, 1981.
121. Converse, J.M., and Presser, S., "Survey Questions: Handcrafting the Standardized Questionnaire", Newbury Park CA: Sage, 1986
122. Cronbach, L. J., " Coefficient Alfa and the Internal Structure of Tests", Psychometrika, Vol. 16, pp 297-334 , 1985.
123. Kaplan R. M. and D. P. Scauzzo " *Psychological testing : Principles, Applications and Issues*", Pacific Grove, C. A. 1993.
124. Hair, J. F., R. E. Anderson, (Jr), R. L. Tatham, and W. C. Black, "*Multivariate Data Analysis*", Prentice Hall International, New Jersey, USA 1998.
125. Bohrnstedt, G., Measurement. In Rossi, P., J. Wright, and A. Anderson (eds.) "*A hand book of survey research"*, Academy press, San Diego, CA, 1983
126. Beatty, P. (1995). "Understanding the Standardized/Non-Standardized Interviewing Controversy," *Journal of Official Statistics*, 11, 147-160, 1995.
127. Faraggi D, Reiser B. "*Estimating of area under the ROC curve*". Stat Med, no. 21: pp 3093-3106, 2002.
128. Hajian Tilaki KO, Hanley JA, Joseph L, Collet JP. "A comparison of parametric and nonparametric approaches to ROC analysis of quantitative diagnosis tests", *Med Decis Making*, no. 17: pp 94-102, 1997.
129. Hand, D. J. and R. J.,"A simple generalization of the area under the ROC curve to multiple class classification problems", *Machine Learning* 45(2), 171- 186, 2001.
130. Hanley JA, McNeil BJ. "The meaning and use of the area under a receiver operating characteristic (ROC) curve". *Radiology*, 143: pp 29-36, 1982.
131. Evans M, Hastings N and Peacock B., "*Statistical Distributions*" 3rd edn (New York: Wiley), 2000.
132. Study group of traceability of weighing instruments, Japan Measuring Instruments Federation, Guideline for calibration laboratory of weighing instruments to acquire JCCS accreditation, 2002.
133. K.D.Sommer, and M. Kochsiek, "Role of measurement uncertainty in deciding conformance in legal metrology", *OIML BULLETIN XLIII-2*, pp.19-24, 2002.
134. Gupta S.C (1984), "Engineering Metrology", Dhanpat Rai Publications, India
135. Guide to the Expression of Uncertainty in Measurement, (First edition 1995), International Organization for Standardization (ISO), Geneva, 1995, 101 pp.

136. B. Ellis, “Basic Concepts of Measurement”, Cambridge, UK: Cambridge University Press, 1968.
137. EA-4/02, “Expression of the Uncertainty of Measurement in Calibration”, Ed. 1: European Cooperation for Accreditation (EA), April 1997.
138. J Ren_e van Dorp and Samuel Kotz. “Generalized Trapezoidal Distributions”, Metrika, vol. 58, no. 1: pp. 85-97, 2003.
139. J. R. van Dorp, S. C. Rambaud, J. G. Perez, and R. H. Pleguezuelo. “An Elicitation Procedure for the Generalized Trapezoidal Distribution with a Uniform Central Stage. Decision Analysis”, vol. 4 no. 3, pp 56 -166, September 2007.
140. Cross, S. S. and R.F.Harrison and R.L.Kennedy; "Introduction to neural networks" The Lancet; vol.346, no.8982, pp.1075-1079, October 1995.
141. Allin Cottrell, “Regression Analysis Basic Concepts”, *Journal of Official Statistics*, Vol 9. pp 171- 175, 2007
142. Flood, I., and Kartam, N., "Neural networks in civil engineering I: Principles and understanding." *Journal of Computing in Civil Engg.*, ASCE, 8(2), 131-148, 1994.
143. Zhang, G. P. “Neural Networks in Business Forecasting”. Idea Group inc. , 2004
144. Freeman, J.A. and Skapura, D.M. “Neural Networks: Algorithms, Applications and Programming Techniques”, Addison Wesley Longman, 1991
145. Winston, P.H., “*Artificial Intelligence*”, Addison-Wesley Publishing Co., Reading, Mass, 1991.
146. Vemuri, V.R., “Artificial Neural Networks: Concepts and Control Applications”, *IEEE Computer Society Press*, Los Alamintos, California, 1992.
147. S.S. Stevens, “Psychophysics. Introduction to its Perceptual, Neural, and Social Prospects”, New York: Wiley, 1975.
148. Kohonen,T; "An introduction to neural computing", Neural networks, vol.1, no.1; pp.3-16, 1988.
149. Gallant, S.I., “Neural Network Learning and Expert Systems”. *MIT Press, Cambridge,* 1993
150. Kashman, A., “Neural Networks for Credit Risk Evaluation: Investigation of Different Neural Models and Learning Schemes. Expert Systems with Applications”, vol. 37, no.9, pp. 6233-6239, 2010.
151. Bishop, C.M., “Training with noise is equivalent to Tikhonov regularization”, *Neural Computation*, vol. 7: pp. 108-116, 1995.
152. Luger, G.F., and Stubblefield, W.A.,“Artificial Intelligence: Structures and Strategies for Complex Problem Solving”. 2nd Edition, Benjamin/Cumming Publishing, Redwood City, California1993.
153. Yam, J.Y.F., and Chow, T.W.S., 1997. “Extended least squares based algorithm for training feed-forward networks”, *IEEE Transactions on Neural Networks*, 8: pp. 806-811, 1997
154. Karl Nygren, Stock Prediction – “A Neural Network Approach”, *Master Thesis*, Royal Institute of Technology, KTH, 2004.
155. Oh, S.H., 1997, ”Improving the error back-propagation algorithm with a modified error function”, *IEEE Transactions on Neural Networks*, 8: 799-803, 1997.

156. Madan M.Gupta, Liang Jin, and Noriyasu Homma, "Static and Dynamic Neural Network", A John Wiley & Sons, INC., Publication, Hobokon, New Jersey
157. S.J. Lee, S.R. Lee, Y.S. Kim, "An approach to estimate unsaturated shear strength using artificial neural network and hyperbolic formulation", *Computers and Geotechnics* 30 (2003) pp. 489–503, 2003. www.elsevier.com/locate/compgeo
158. Piche, S., Keeler, J., Martin, G., Boe, G., Johnson, D., Gerules, M., "Neural Network Based Model Predictive Control", *Proceedings of Neural Information Processing System Conference*, 2000.
159. Hansen, L.K., and Salamon, P., "Neural network ensembles", *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 12: 993-1001, 1990.
160. Piuri, V. and C. Alippi; "Artificial neural networks"; *Journal of system Architecture*; Vol. 44, No.8; pp.565-567, April 1998.
161. Reed, R.D., and Marks II, R.J., "Neural Smithing", MIT Press, Cambridge, Mass, 1999.
162. Geantă, M., Vaszilcsin, N., "Prediction of Periodic Proprieties of Elements using an Artificial Neural Network", Chem. Bull,*"POLITEHNICA" Univ. Timişoara*, vol. 43(57), pp. 44-50, 1998.
163. *Rich, E., and Knight, K.,* "Artificial Intelligence*". McGraw-Hill, New York, 1991*
164. Geantă, M., Perju, D., "The Main Parameters of Backpropagation Artificial Neural Network", Chem. Bull.*"POLITEHNICA" Univ. Timişoara*, vol. 42 (56), pp. 35-430, 1997
165. Bressloff,P.C. and D.J.Weir "Neural networks"; *GEC Journal of Research*, vol.8, no.3, pp.151-16, 1991.
166. C.Y. Kim, G.J. Bae, S.W. Hong, C.H. Park,H.K. Moon, H.S. Shin, "Neural network based prediction of ground surfacesettlements due to tunneling", *Computers and Geotechnics*, no. 28, pp. 517–547, 1991. www.elsevier.com/locate/compgeo

LIST OF PUBLICATIONS [INTERNATIONAL JOURNAL] BASED ON THIS THESIS

1. *Vinodkumar Jacob, M. Bhasi, R. Gopikakumari*, "Impact of Person related variables on human errors in measurement", Canadian Journal on Electrical and Electronics Engineering, January 2012, Vol. 3, No. 1., Pp. 17 – 24.

2. *Vinodkumar Jacob, M. Bhasi, R. Gopikakumari*, "Influence of Work related variables on human errors in measurement", International Journal of Electrical and Electronics Engineering Research, December 2011., Pp. 1 - 11.

3. *Vinodkumar Jacob, M. Bhasi, R. Gopikakumari*, "Impact of Human Factors on Measurement Errors", International Journal of Measurement Technologies and Instrumentation Engineering, Oct. – Dec. 2011, Vol 1(4), Pp. 28 – 44.

4. *Vinodkumar Jacob, M. Bhasi, R. Gopikakumari*, "Error in Measuring Standards, Uncertainty in Type Approval and Verification in Legal Metrology" Accepted for Publication by the Mediterranean Journal of Measurement and Control.

5. *Vinodkumar Jacob, M. Bhasi, R. Gopikakumari,* "A Study of Measurement Errors and uncertainty in Legal Metrology", under review in Elsevier.

APPENDIX - I

SANWA - Analog multimeter

max. 1 000 V, max. 0.25 A | YX360TRF

Drop shock proof meter, Null (zero center) meter ±5 / ±25 in DCV, High resistance up to 200MΩ with low voltage
Protective body cover Capacitance, dB, Li measurement, Bandwidth : 30～100kHz (AC10V)

EZ Digital DM-333 Digital Multimeter

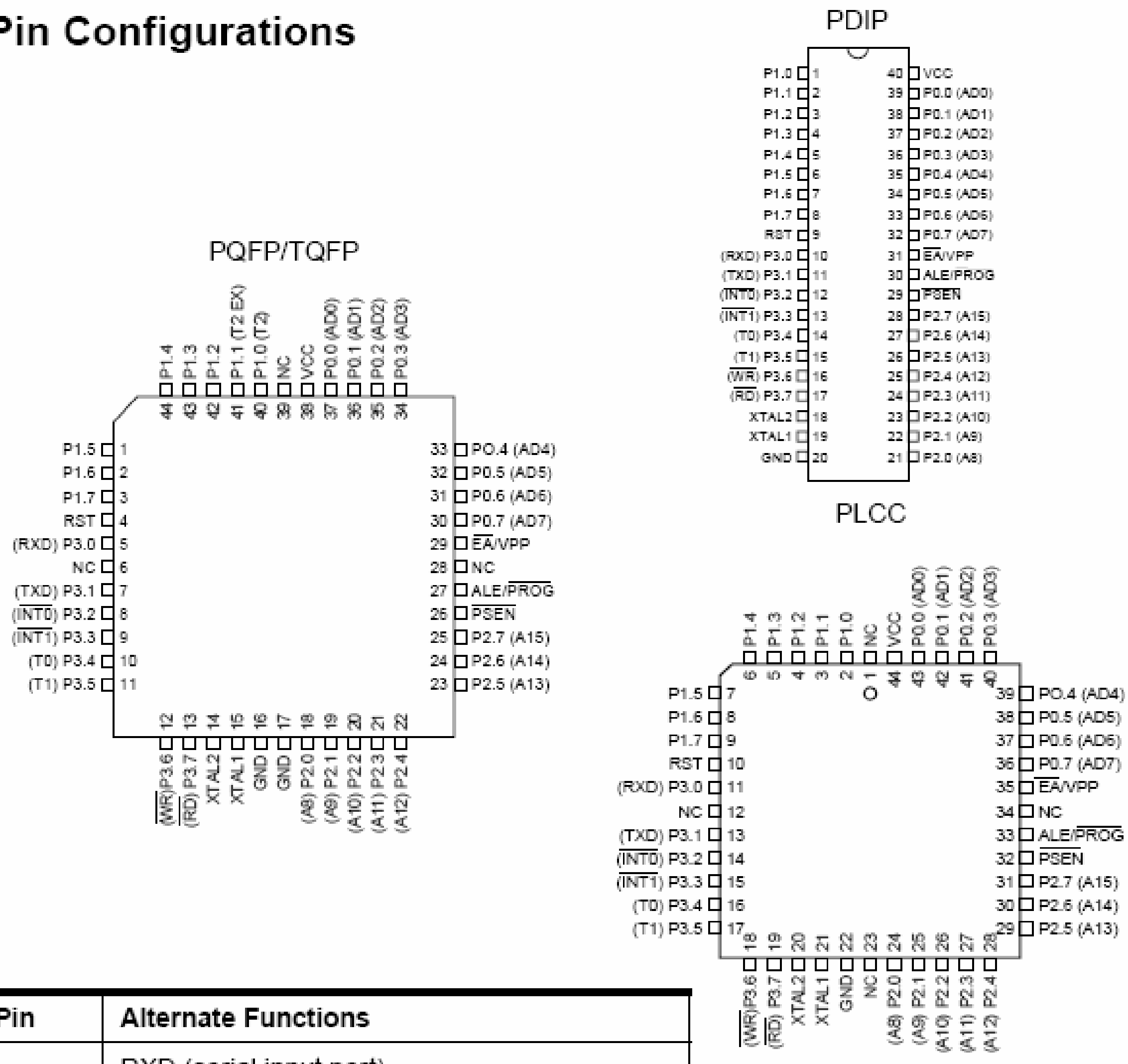
Features:

3 and 3/4 digit display, 3200 count, 0.5% basic accuracy on DC volts, Auto/manual range select type
AC/DC volts/amps with analog bar graph, Resistance/Diode/Continuity check, Data hold, Low battery indication/Auto power off, Overload protection/Safety rubber holster, Wrong position protector in the 10A input terminal.

	Range	Resolution	Accuracy
DC Voltage	320mV 3.2V	0.1mV 1mV	±(0.5%+2dgt)
	32V 320V 1000V	10mV 100mV 1V	±(1.0%+2dgt)
AC Voltage	3.2V 32V 320V 750V	1mV 10mV 100mV 1V	±(1.2%+4dgt)
DC Current	320μA 32mA	0.1μA 10μA	±(1.0%+2dgt)
	3200μA 320mA	1μA 0.1mA	±(2.0%+2dgt)
	10A	10mA	±(2.5%+5dgt)
AC Current	320μA 3200μA 32mA 320mA	0.1μA 1μA 10μA 0.1mA	±(1.0%+2dgt)
	10A	10mA	±(2.5%+5dgt)
Resistance	320Ω 3.2kΩ 32kΩ	0.1Ω 1Ω 10Ω	±(1.0%+2dgt)
	3.2MΩ 32MΩ	1kΩ 10kΩ	±(3.5%+5dgt)

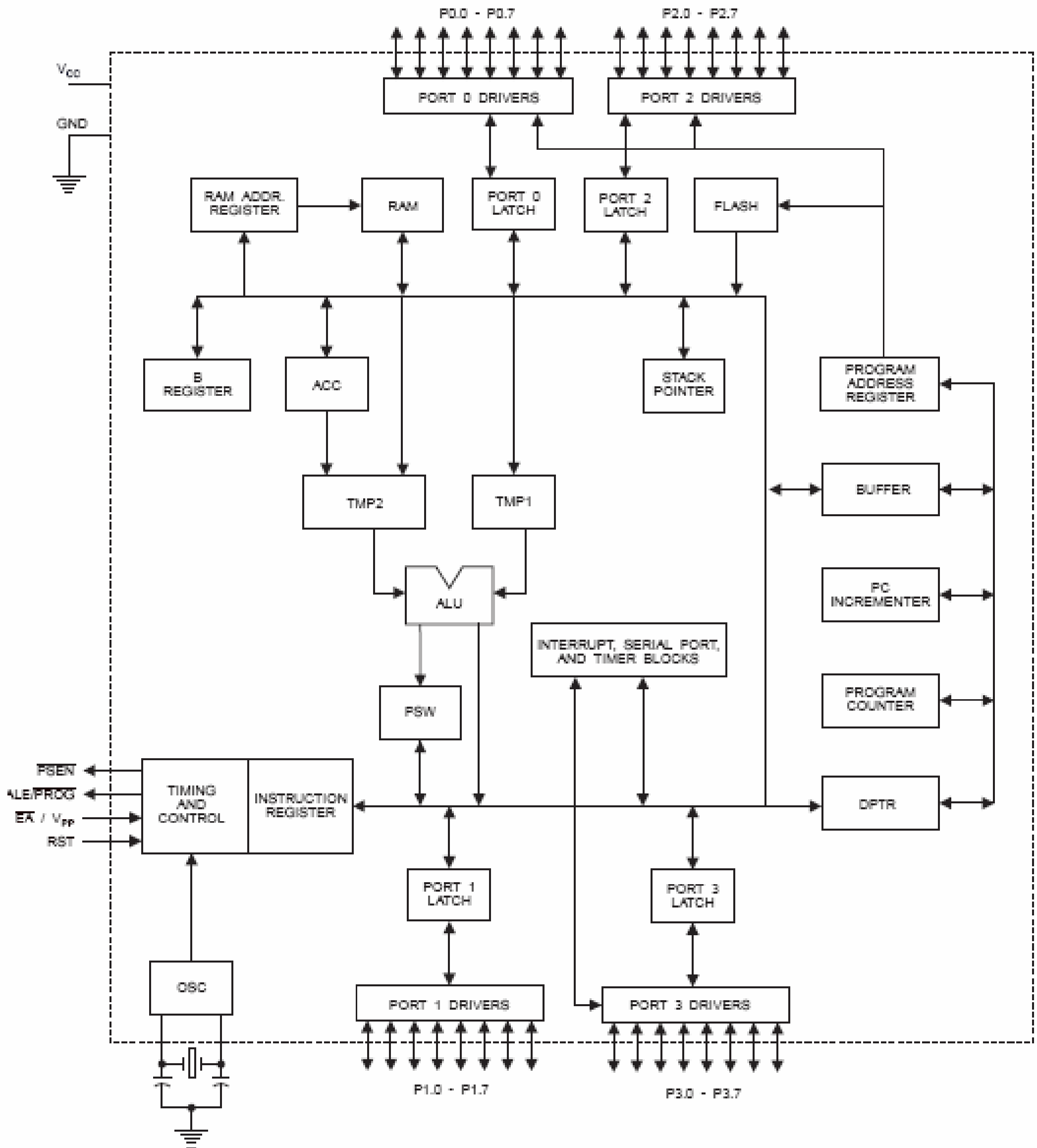
Micro Controller 8951

Pin Configurations



Port Pin	Alternate Functions
P3.0	RXD (serial input port)
P3.1	TXD (serial output port)
P3.2	$\overline{\text{INT0}}$ (external interrupt 0)
P3.3	$\overline{\text{INT1}}$ (external interrupt 1)
P3.4	T0 (timer 0 external input)
P3.5	T1 (timer 1 external input)
P3.6	$\overline{\text{WR}}$ (external data memory write strobe)
P3.7	$\overline{\text{RD}}$ (external data memory read strobe)

Block Diagram



Status of External Pins During Idle and Power-down Modes

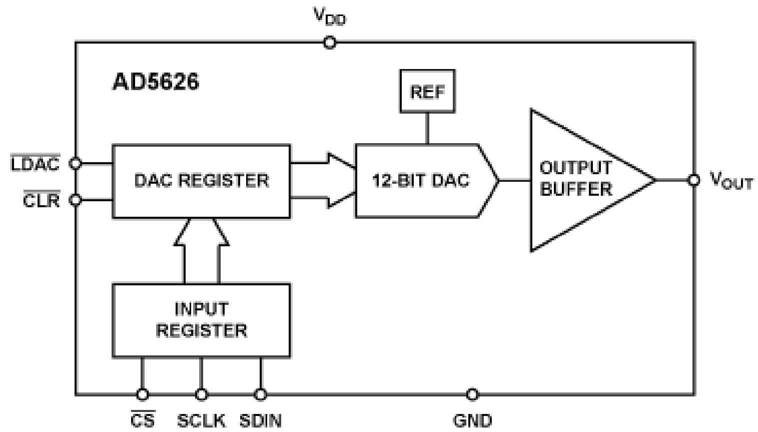
Mode	Program Memory	ALE	$\overline{\text{PSEN}}$	PORT0	PORT1	PORT2	PORT3
Idle	Internal	1	1	Data	Data	Data	Data
Idle	External	1	1	Float	Data	Address	Data
Power-down	Internal	0	0	Data	Data	Data	Data
Power-down	External	0	0	Float	Data	Data	Data

12 bit Serial DAC AD5626

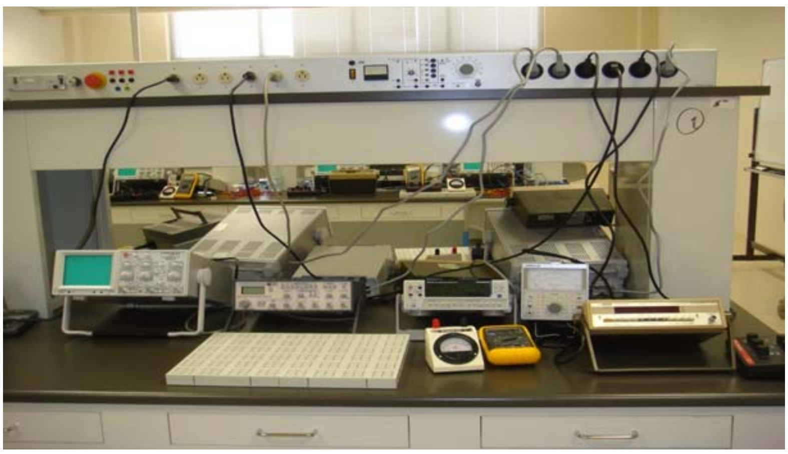
FEATURES AND BENEFITS

- 8-lead MSOP and 8-lead LFCSP packages
- Complete, voltage output with internal reference
- 1 mV/bit with 4.095 V full scale
- 5 V single-supply operation
- No external components required
- 3-wire serial interface, 20 MHz data loading rate
- Low power: 2.5 mW

FUNCTIONAL BLOCK DIAGRAM FOR AD5626



Photographs of Experiments on Human Related Measurement Errors



Appendix - II

Survey to analyze perception on Human Errors on Professionals

Name : ………………………………………….….Designation : …………………………

(Please tick [✓] the correct answer / fill in the blanks)

1. The ascending order of measurement error produced among the following display devices is:
a. Analog multimeter [ ] *b. Digital multimeter* [ ] *c. DSO***[ ]**
Write 1,2,3in the appropriate boxes to indicate the answer.
2. The ascending order of measurement error produced among the following mechanical devices is:
a. Analog Vernier Caliper [ ] *b. Digital Caliper [ ]*
Write 1,2,3in the appropriate boxes to indicate the answer.
3. The ascending order of measurement error produced depending upon the time of day is :
a. FN [ ], b. AN [ ], c. Night hours [ ]
Write 1,2,3in the appropriate boxes to indicate the answer.
4. The error produced in the measurement of ……. sized material is greater than …….. sized.
a. *smaller, larger* *b. larger, smaller*
5. Experience causes an error reduction in Analog measurements of about
a. 0 –5% *b. 6 – 10%* *c. 11 – 15%*
6. Experience causes an error reduction in Digital Measurements of about
a. *0 – 0.49%* *b. 0.5 – 0.99%* *c. 1 – 1.5%*
7. Compare the Analog measurement error created by subjects with above average intelligent quotient [IQ1] and average intelligent quotient [IQ2]
a. IQ1 > IQ2 b. IQ2 > IQ1 if so the range of difference in error is …………..
a. 0 – 2% *b. 2 - 4%* *c. 4 – 6%*
8. Compare the Digital measurement error created by subjects with above average intelligent quotient [IQ1] and average intelligent quotient [IQ2]
a. IQ1 > IQ2 b. IQ2 > IQ1 if so the range of difference in error is …………..
a. *0 – .99%* *b.1 – 1.99%* *c.2 – 2.99%*
9. Compare the Analog measurement error created by subjects with above average intelligent quotient [IQ1] and below average intelligent quotient [IQ3]
a. IQ1 > IQ3 b. IQ3 > IQ1 if so the range of difference in error is …………..
a. 0 – 2.99% *b. 3 – 5.99%* *c. 6 – 9%*
10. Compare the Digital measurement error created by subjects with above average intelligent quotient [IQ1] and average intelligent quotient [IQ2]
a. IQ1 > IQ2 b. IQ2 > IQ1 if so the range of difference in error is …………..
a. *0 – 1%* *b.1.1 – 2%* *c.2.1 – 3%*
11. Whether the measurement error created by experienced technicians in the age group of 41 – 50 [ET2] is greater than the age group of 31 – 40 [ET1]
a. Yes b. No
12. The range of measurement error difference between ET1 and ET2
a. 0 – 2% *b. 2.1 -4%* *c. 4.1 – 6%*
13. The experienced technicians may make an Analog measurement error more in the afternoon session than forenoon session in the range of
a. *0 – 2%* *b. 2.1 -4%* *c. 4.1 – 6%*

14. The experienced technicians may make a digital measurement error more in the afternoon session than forenoon session in the range of

a. *0 – 2%* *b. 2.1 -4%* *c. 4.1 – 6%*

15. Whether the measurement error produced by the inexperienced subjects in the afternoon compared to forenoon is more than the same produced by the experienced technicians (Y/N) ……….., If so the percentage of error difference is ………….

a. 0 – 3.99% *b. 4 -7.99%* *c. 8 – 12%*

16. The inexperienced subjects make an analog measurement error more in the afternoon than forenoon in the range of

a. 0 – 6.99% *b. 7-13.99%* *c. 14 – 21%*

17. The inexperienced subjects make a digital measurement error more in the afternoon than forenoon in the range of

a. 0 – 3.99% *b. 4 -7.99%* *c. 8 – 12%*

18. The mean percentage error produced by experienced technicians in analog measurement is more compared to digital measurement by

a. 0 – 3.99% *b. 4 -7.99%* *c. 8 – 12%*

19. The mean percentage error produced by inexperienced subjects in analog measurement is more compared to digital measurement by

a. 0 – 6.99% *b. 7-13.99%* *c. 14 – 21%*

20. Training shows an improvement in analog measurement error of about

a. 0 – 6.99% *b. 7-13.99%* *c. 14 – 21%*

21. Training shows an improvement in digital measurement error of about

a. 0 – 3.99% *b. 4 -7.99%* *c. 8 – 12%*

22. The performance of experienced technicians with below average IQ is comparable to inexperienced with IQ level

a. above average *b. average* *c. below average* *d. none of these*

23. The performance of engineering students with Kerala State Entrance Rank 1 to 5000 in measurements is comparable to inexperienced youngsters with IQ level

a. above average *b. average* *c. below average* *d. none of these*

24. The performance of engineering students with Kerala State Entrance Rank 5001 to 15000 in measurements is comparable to inexperienced youngsters with IQ level

a. above average *b. average* *c. below average* *d. none of these*

25. The performance of engineering students with Kerala State Entrance Rank above 15000 in measurements is comparable to inexperienced youngsters with IQ level

a. above average *b. average* *c. below average* *d. none of these*

26. Whether the measurement error produced by 1st year B. Tech. students is more than final year B.Tech. students (Y/N)……., if so what is the percentage ?

a. 0 – 10% *b. 11 – 20%* *c. 21 – 30%*

27. Which among the following make more measurement errors

a. male *b. female* *c. no significant difference between male and female*

28. If the subjects were not put under time pressure may cause an error reduction of about ……..

a. 0 -10% *b. 11-20%* *c. 21 – 30%*

29. Better environment causes an error reduction by

a. 0 -7% *b. 8-15%* *c. 16 – 23%*

30.A. Whether task difference will have any effect on measurement error ?

a. YES b. NO

b. If so, which type of task will create less error a. Simple b. Difficult