Special Issue

METEOROLOGICAL EQUIPMENT TEST AND EVALUATION

IEEE Transactions on Geoscience Electronics Vol. GE-11 No. 2 April 1973

IEEE TRANSACTIONS ON

GEOSCIENCE ELECTRONICS

APRIL 1973 VOLUME GE-11 NUMBER 2

A PUBLICATION OF THE IEEE GEOSCIENCE ELECTRONICS GROUP

SPECIAL ISSUE—METEOROLOGICAL EQUIPMENT TEST AND EVALUATION

PAPERS

INTRODUCTORY REMARKS

One of the idiosyncrasies of the environmental sciences is that we cannot examine a slice of "nature", for example air, in the laboratory and expect that it will perform as it might in the free state. Another is that sensing devices evaluated in the laboratory do not perform in the same way as they do when exposed to the environment. Failure to heed the implied warnings of these two statements is the reason for most problems of environmental observing programs. Instrument accuracy statements, based on laboratory calibrations, commonly are used as data accuracy statements. Worse yet, data users too often state requirements for new sensors based upon the results of these same calibrations. This creates nearly impossible restrictions on the design and test of a field instrument, if indeed the prototypes are ever field tested.

It is my personal observation that engineers who design and build instruments and those who use the data understand too little of each other's problems. Let me give you four examples. While each is true, I will keep my remarks as general as possible so as to avoid personalities.

A paper appeared two years ago in which the author examined the Eckman Spiral in the lower atmosphere over ocean areas. In theory, the spiral is a turning of the horizontal wind vector with height (up to about 1 km) due to a decrease in the effect of surface friction with altitude. Oceans are a perfect place to study this phenomenon because terrain effects over land tend to distort the spiral. Based on long-term averages of wind data from several Ocean Station Vessels, the author concluded that the Eckman Spiral theory was valid. What he did not know was that the beam-width of the radar used on the ships to track the rising baloon was too broad to obtain accurate wind data in the lower levels. To compensate, observers on the ships assumed an Eckman Spiral in the first kilometer in order to provide continuity with upper layer wind information.

In recent years, numerous articles have been published in both the scientific and popular literature relating a decrease in incoming solar radiation to increased air pollution. Some have taken the argument even further. An article appeared in a Phoenix paper earlier this year which suggested that the decrease in the incidence of skin cancer was due to decreased solar radiation, which, in turn, arose from increased air pollution. To help support his theories, the author examined pyranometer data (incoming solar radiation intensity) for about a 20 year period for the Phoenix area. He used a regression curve fitted to the data which showed about a 20% decrease in the annual average of daily incoming radiation. Unfortunately, the absorbent material in the pyranometer had been changed and had deteriorated after long exposure to the bright desert sun. Also, the sensor had been relocated twice during the period of record. A new analysis of the data (to be published later this year) shows that the decreasing trend of solar radiation is in the "noise level" of the other uncertainties of the data.

In the early 1960's a requirement was stated for a sensor to measure atmospheric index - of - refraction gradients. It was believed that real-time information was needed to correct air-craft and space vehicle radar tracking data in order to realize the inherent accuracies of the radars. To meet the stated requirement, an expendable refractometer for use on a balloon was developed. However, it was later learned that the "4/3rds-earth" correction was not being applied to the radar data. This is a routine correction which compensates for radar beam "bending" assuming the conditions of a standard atmospheric profile. Additional corrections based on local climatology were also suggested. These corrections eliminated all but the most severe (and relatively rare) atmospheric effects. It was for these latter situations that the balloon-borne refractometer eventually proved to be useful. The requirement, and hence the instrument, might have been quite different had the situation been more fully known when development began.

The final example concerns the network of climatological stations managed by the National Weather Service. There are about 13,000 such stations across the United States that are operated by cooperative observers. These largely unpaid citizens take daily readings of maximum and minimum temperatures and precipitation accumulation. One observer was a lady in Massachusetts who had been recording observations for over 25 years without missing a day. One summer the people who compile the records, noticed that her data were not compatible with other data in the vicinity. After several weeks of investigation, we learned that she had left for a vacation at New York City and had taken her instruments with her. She did not want to break her record and had continued to take and send in her observations, but she still gave her Massachusetts address.

All four examples illustrate the hazards in using data provided by others without a thorough understanding of the user requirements, instruments, data processing and related "people-problems". In the example of the cooperative observer, her data may have been accurate, precise, and she certainly was reliable. The data just were not representative. The bright, new pyranometers in labortory calibration were precise. The long-term data accuracy however, was poor when compared to a standard.

This brings me to the purpose of the five papers assembled in this issue. Each of the above anecdotes illustrates a specific problem in translating data needs to suitable sensing programs. This translation is often confused and is frequently a failure. The following five papers are aimed at providing tools to answer one question -- "How good are the data?"

Gill and Hexter propose a standard set of terms for use by meteorologists and engineers. These two authors represent the dichotomy between engineer and data user. Gerry Gill is well known for his wind systems (Young, Co.), widely used in Micro-meteorology (a study of small scale meteorological features with horizontal dimensions of less than a kilometer or so).

Paul Hexter is a meteorologist with the National Weather Service, who is concerned with how the data are used. I was pleased to learn that the differences in opinion between engineer and user could be resolved so well. Their paper establishes a common terminology for discussing the question -- "How good are the data?"

Once we have agreed on the basic terminology, Dave Beaubien considers the need for a set of accepted and standard test procedures. These would be procedures to which manufactures could reference the performance of their products and to which users could reference their needs. He is well known in the field of meteorological instrumentation for his stand on this issue. His paper in this journal is meant to carry the message to the electrical and electronics engineers who are at the heart of all instrumentation development.

The next three papers take the argument one step further and propose means by which we can carry out test programs and interpret the test results. An underlying question is -- "How do data from a network of similar sensors compare with each other?" This leads to the concept of functional precision, as discussed by Walt Hoehne.

One of the more challenging developments in the field of environmental sciences is the trend toward large field experiments. Whole "chunks" of the atmosphere are being treated as a "laboratory." In the Barbados Oceanographic and Meteorological Experiment (BOMEX), the "chunk" was 500km on a side. Even reasonably controlled conditions are not possible in such experiments.

One way to estimate a "true" value or to intercompare data, particularly during large field experiments, is to use a variety of sensing systems from which data on the same parameters can be derived. Josh Holland and Don Acheson are two of those responsible for developing this technique for BOMEX data analysis. They discuss this procedure and the attendant problems in the fourth paper.

A second method of obtaining a "true" value is to devise your own field reference. Most laboratory standards do not operate in field environments -- most field instruments are not designed to be standards. Max Hinzpeter discusses the development of a reference radiosonde which can be used to calibrate upper air measurements. As an indication of the need for such a reference, upper air temperature data from the world-wide network show serious discontinuities at national boundaries where different sondes are used.

Given the increased complexity and sophistication of environmental sensing systems, and the need for more data of greater accuracy, the question of "how good are the data" will come back to haunt many an engineer and scientist. The message of these papers is that standard terminology, standard test procedures and standard test analysis are mandatory - and overdue - in the field of environmental instrumentation.

James Giraytys
Technical Editor, AIR
IEEE Transactions on
Geoscience Electronics

SOME INSTRUMENTATION DEFINITIONS FOR USE BY METEOROLOGISTS AND ENGINEERS

Gerald C. Gill
Professor of Meteorology
The University of Michigan
Ann Arbor, Michigan 48104

Paul L. Hexter
Data Acquisition Division
National Weather Service, NOAA
Silver Spring, Maryland 20910

Abstract

To bridge the growing communications gap between meteorologists and engineers, a provisional list of definitions of terms frequently used by both groups is presented. These definitions are divided into four categories: A) Basic terminology (definitions of terms such as sensor, transducer, instrument, and data acquisition system). B) Terms relating primarily to the sensor (definitions of terms such as time constant, distance constant, damping ratio, and hysteresis). C) Terms relating primarily to the instrument (definitions of terms such as sensitivity, resolution, error, accuracy, and linearity). D) Terms relating primarily to the measuring process (definitions and discussions of terms such as precision, reliability, and representativeness).

The authors hope this selected set of definitions will not only be of immediate use as a step towards a standard terminology but also will form the basis for a more comprehensive Glossary of Meteorological Instrumentation Terms.

Introduction

Meteorologists and engineers often have difficulty communicating clearly with one another concerning common problems of meteorological measurements. There is a need for common definitions of such terms as sensitivity, accuracy, reliability, etc. Some of the uses for such definitions are:

a) To enable the data user (meteorologist), the instrument designer, and the field engineer to understand each other.
b) To provide common terminology for use by equipment designers and manufacturers in specifying instrument characteristics, thus permitting direct comparisons of competing instrument systems.
c) To clarify statements on equipment requirements, standards, and deficiencies.

We have used, for the most part, terms previously defined by individuals and groups concerned with these problems and have drawn freely from these works (see references). Of course, these definitions often did not agree, and so we have selected what we considered to be the better definitions in such cases. In general, the basic terms are commonly found in the literature. We suggest that the meteorological community consider these definitions for general usage. The definitions have been reviewed and endorsed by most members of the Committee on Atmospheric Measurements (CAM) of the AMS. The authors feel that there is a very real and urgent need for uniformity of usage and are aware of current work on standardization of terms by such groups as the Working Group on Accuracy of Measurement which reports to the WMO Commission for Instruments and Methods of Observations (CIMO); Data Seminars of the WMO; Inter-Range Instrumentation Group-Meteorological Working Group; Instrument Society of America, and others. It is hoped that a glossary of instrument terms acceptable to the meteorological community will be prepared. As a preliminary step toward this goal, the authors have prepared the following provisional list of definitions. We hope that these definitions will be utilized by members of the meteorological community in scientific papers and in instrument specifications.

Definition of Terms

Basic Terminology

1. The term sensor refers to the specific sensing elements of an instrument which are designed to react to changes in the environment, e.g., the resistance wire in a resistance thermometer; the cupwheel of an anemometer; the thermistor, hygristor, and aneroid of a radiosonde.

2. The term transducer refers to a device for converting energy from one form to another form. In the case of meteorological instruments, it usually refers to the device used to convert the mechanical motion or mechanical position of the sensor to an electrical

Manuscript received Sept. 10, 1972

signal for remote indication, e.g., the tachometer generator, or photo chopper circuit of a cup anemometer; the potentiometer, or selsyn motor in a wind vane transmitter; or the linear differential transformer of a remote indicating aneroid barometer.

3. The term instrument is used to describe sensors; associated mechanical and/or electronic linkages (e.g., transducers); and the data readout or recording device. Illustrations would be the resistance thermometer, which includes the temperature sensitive resistance wire, the wire lead-ins, automatic bridge system, and panel meter or recorder; and the anemometer which includes the cupwheel or propeller, the mechanical to electrical transducer; and the readout indicator or recorder.

4. The term data processor refers to an accessory piece of equipment that accepts the output of several sensors (or transducers) and conditions the signals so that all may be recorded on the same recorder or readied for transmission for real time operational use. (The data are processed in real time, either as one or more channels of analog data and/or one or more channels of sequentially sampled digital or analog data from each of the several sensor inputs. Conditioning of the input signals may take the form of amplification, linearization, totalizing counts, off-setting the zero, etc.)

5. The data acquisition system consists of sensor(s), transducer(s), data processor(s), and a recording and/or data transmission device. Examples are the radiotheodolite, the weather radar, and the automatic meteorological observing station. The term is distinguished from instrument in that it refers to complex equipments, such as given in the examples, rather than simpler equipments such as the mercury-in-glass thermometer, aneroid barometer, or tipping bucket rainguage. A data acquisition system will include instruments, as for example the automatic meteorological observing station includes the tipping bucket rainguage and a three-cup anemometer; and the radiotheodolite system includes the radiosonde with its sensors, and the ground receiving and processing equipment.

6. A measuring process (method of measurement) is defined by specifying the apparatus and auxiliary equipment to be used, the operations to be performed, the sequence in which they are to be executed, and the conditions under which they are to be carried out.

Terms Relating Primarily to the Sensor

7. The term "time constant" is used to describe the response time of a temperature sensor whose rate of change of reading is directly proportional to the instantaneous temperature difference: see Fig. 1. The time constant is the period that is required for the temperature sensor to respond to 63.2 per cent (1 - 1/e) of the step-wise change in temperature. The time constant is usually expressed in seconds (or minutes). (The term "time constant" is equally applicable to sensors of humidity, pressure, wind speed, etc., as it is to temperature sensors. The time constant of a sensor is usually dependent on several factors of the environment in which the sensor is exposed, for instance: the fluid in which it is exposed, air or water; the density of the fluid, air at 1000 mb., or air at 100 mb.; the flow rate of the fluid, still air, or air at 20 knots; etc. Accordingly, when specifying the time constant of a sensor, one should specify the general conditions of the environment that directly affect the measurement.)

8. The distance constant of a sensor is the length of fluid flow past the sensor required to cause it to respond to 63.2 per cent of the step-function change in speed. It is measured in feet (or meters). For some sensors the term distance constant is more appropriate than the term "time constant." For instance, when a three-cup anemometer is suddenly transferred from quiet air into a wind of 10 ft/sec, the time constant might be 3.0 seconds, but if the same instrument were transferred from quiet air into a wind speed of 20 ft/sec, the time constant would be only 1.5 seconds. In each case, the same amount of air (3.0 sec x 10 ft/sec = 30 ft; 1.5 sec x 20 ft/sec = 30) will have passed for the sensor to respond to 63.2 per cent of the speed change. Thus, the term "distance constant" is more appropriate for such a sensor. This is likewise true for propeller anemometers and propeller-type flow meters.

(In determining the time constant of a cup anemometer by obtaining an acceleration curve similar to Fig. 1, θ_o corresponds to the cup wheel being held from turning; θ_e corresponds to the steady wind tunnel speed; and $t = 0$ corresponds to the instant of release of the cup wheel. Due to abnormal turbulence in the wake of a stalled cup wheel (or a propeller) it is best to use only the upper 50 to 60% of the acceleration curve for analysis purposes.)

9. In the case of temperature sensors and wind speed sensors, the sensor has a "first-order response" as shown in Fig. 1. In these cases, where there has been a sudden change of the measured variable from θ_o to θ_e, the sensor never overshoots the new steady state condition θ_e. This is not the case when a wind vane is subjected to a sudden shift in direction. The response of a typical wind vane to an instantaneous change in direction of 15° might be as shown in Fig. 2. The vane overshoots the final value, executing a damped simple harmonic motion from 2 to 10 oscillations before it reaches the steady state, depending on the "damping ratio" of the vane. The damping ratio of a sensor is the ratio

$$\frac{\text{actual damping coefficient}}{\text{critical damping coefficient}}$$

in the second order differential equation that specifies the response of the sensor.

The period of the damped oscillation as shown in Fig. 2 is called the damped period (equal to 2.0 sec in this case); and the corresponding amount of air that passes for one complete damped oscillation is called the damped natural wave length (equal to 60 ft. in this case; 2 sec x 30 ft/sec).

10. In order to use a characteristic distance for a wind vane which is analogous to a distance constant for a wind speed sensor, MacCready and Jex (1964) defined the delay distance as the length of air (in feet or meters) that passed a wind vane for the vane to respond to 50% of a sudden angular change in wind direction. From Fig. 2 the delay distance of the vane would be approximately 10 ft (0.33 sec times 30 ft/sec). (For a much fuller explanation of these terms and their use in fluctuating winds, the readers are referred to MacCready and Jex (1964), MacCready (1970), Gill (1967), and Moses (1968). (For routine use, wind vanes should have a damping ratio, $h > 0.2$; and for diffusion and turbulence studies where the standard deviation of azimuth angle is used, $h > 0.35$).

11. The term dead band (or back lash) refers to the range through which an input may be varied without initiating response (see Fig. 3b). Dead band is usually expressed in per cent of full-scale range.

12. Sensors may exhibit a phenomenon known as hysteresis. Hysteresis has been defined as the maximum difference in output for any given input value (within the specified range) when the value is approached first with increasing, and then with decreasing input signals (see Fig. 3a). By this usage both the dead band error and hysteretic error are combined into one error. (The hysteretic error is caused by energy absorption in the elements of the measuring instrument. This error is significant in some aneroid barometers and in radiosonde humidity elements.) Hysteresis is usually expressed in per cent of full-scale range.

Terms Relating Primarily to the Instrument and the Measuring Process

13. The resolution of an instrument may be defined as the smallest change in the environment that causes a detectable change in the indication of the instrument. (Example: For a recording resistance thermometer having a span of 100C with a 1000-turn slide wire, the resolution of a new and properly adjusted instrument might be ± 0.1C, corresponding to one turn of the slide wire. But the resolution might be as low as ±1.0C if the sliding contact were badly worn or the servo amplifier poorly adjusted.)

14. The error of an instrument is the algebraic difference between the indication and the true value of the measured signal. It is the quantity which algebraically subtracted from the indication gives the true value. (Error = indication - true value.)

15. The accuracy of an instrument (after application of its calibration curve) is the degree with which the instrument will measure the variable in terms of an accepted standard value or true value. (The term accuracy is usually measured in terms of inaccuracy but expressed as accuracy. Thus the accuracy of the resistance thermometer referenced above might be expressed as ±0.5C in range 0 to 50C and ±0.8C in range 50C to 100C.)

When defining the performance specifications of a device under active operating conditions, SAMA (Scientific Apparatus Manufacturers Association), 1970, uses the term reference accuracy.

They define *reference accuracy* as "a number or quantity which defines the limit that errors will not exceed when the device is used under 'reference operating conditions.' Reference accuracy includes the combined 'conformity, hysteresis and repeatability' errors" (see Fig. 4). Thus the various dynamic and static errors are combined into one reference accuracy for a particular set of operating conditions. Example: the 'reference accuracy' (of the previously referred to resistance thermometer) might be ±0.5C for range 0C to 50C and ± 0.8C for range 50C to 100C for input frequencies up to 0.001 Hz; ±2C over range 0C to 100C if the input frequency was 0.01 Hz with an amplitude of ±5C.

It is recommended that for performance specifications "accuracy" be assumed to mean "reference accuracy" unless otherwise stated.

16. The term *sensitivity* refers to the ratio of the full-scale output of the device to the full-scale input value. (It is thus quite different from the term resolution, and the two terms should not be used interchangeably.) (Examples: 1) The sensitivity of the recording thermometer might be 10 inches pen movement per 100 C temperature change; 2) The sensitivity of a digital indicating aneroid might be 2000 scale units per 200 mb pressure change, adjustable over range 600 to 1100 mb.)

17. The term *speed of response* of an instrument is variously applied. Often it indicates the time required for the indicator or recorder to follow 90% of a sudden full-scale change in the measured variable; sometimes 99% of full-scale. Sometimes the term indicates the time that elapses from the application of a sudden step-wise change until the recorder reading is steady. Thus, the term must always be defined, e.g., 90% response in 2 sec. We suggest that the use of this term be discouraged.

18. In the calibration of an instrument the indications of the instrument are usually plotted against known values of the measured variable for a number of points over the range of the instrument. Since these points generally do not yield exactly a straight line, instrument manufacturers usually draw a "best fit" straight line through the calibration points and specify *linearity* as the maximum deviation of any points from this straight line. The linearity, often expressed as a percentage, refers to the percentage of full-scale deflection rather than percentage of the indication. (Example: For the recording resistance thermometer, the linearity might be expressed as ±0.5%. This would indicate a deviation of ±0.5C from true value over the complete range 0C to 100C.)

19. The *reporting increment* is the smallest unit of measurement to be used in each reading of the instrument. The reporting increment should always be greater than the resolution of the instrument. To return to the example of the resistance wire thermometer, the meteorologist may say he desires read-out to the nearest whole degree Celcius. The reporting increment is then 1C.

20. The *repeatability* of an instrument is the closeness of agreement among a number of consecutive measurements of the output for the same value of the input under the same operating conditions, approaching from the same direction. (It is usually measured as "nonrepeatability" but expressed as "repeatability," in per cent of span. It does not include hysteresis.)

21. The *reliability* (Norton 1969) of an instrument is a measure of the probability that the instrument will continue to perform within specified limits of error for a specified length of time under specified conditions. (For a discussion of some of the practical problems relating to data reliability, readers are referred to IRIG Document #110-70). (Meteorological Data Accuracies Committee, MWG, IRIG, 1970)

Summary

The authors have submitted definitions of selected terms for use by meteorologists and engineers, to clarify the thinking of both groups and to bridge a growing communications gap that has been developing as instruments and instrument needs have become more sophisticated. We are cognizant of the incompleteness of these definitions, but hope the list will be added to as clear definitions become available of other important meteorological instrument properties. It would be most helpful if a *Glossary of Meteorological Instrument Terms were developed, somewhat along the lines* of *Glossary of Meteorology*, but probably *not over 50 pages in extent*.

The authors are most grateful for the constructive replies they received from members of the Committee on Atmospheric

Measurements of the AMS and associates of each of us. The final selection of definitions is our own, and we accept the responsibility for the selection and the definitions. We hope both meteorologists and engineers will use these definitions.

References

Gill, G. C., 1964: Data validation. Symposium on Environmental Measurements: Valid Data and Logical Interpretation. PHS Publication #999-AP-15, pp. 85-100.

Instrument Society of American, 1963: Tentative Recommended Practice No. ISA-RP37.1 "Nomenclature and Specification Terminology for Aerospace Test Transducers with Electrical Output."

Instrument Society of America, 1963: *ISA Transducer Compendium*, Plenum Press.

MacCready, P. B., Jr., 1970: Theoretical considerations in instrument design, *Meteor. Monogr.*, Vol. II, 202-210.

MacCready, P. B., Jr., and H. R. Jex, 1964: Response characteristics and meteorological utilization of propeller and vane wind sensors. *J. Appl. Meteor.*, 3, 183-193.

Meteorological Data Accuracies Committee, MWG, IRIG, 1970: Reliability of meteorological data. IRIG Document 110-70.

Morrissey, J. F., and F. J. Brousaides, 1970: Temperature induced errors in the ML-476 humidity data. *J. Appl. Meteor.* 9, 805-808

Moses, H., 1968: *Meteorology and Atomic Energy*, Ch. 6, 257-300.

Norton, Harry N., 1969: *Handbook of Transducers for Electronic Measuring Systems*, Prentice-Hall Inc.

Ruskin, R. E. (moderator), 1970: Panel discussion 3: Can instruments be designed to be both accurate and representative? *Meteor. Monogr.*, Vol. II, No. 33.

Scientific Apparatus Makers Association, 1970: Process measurement and control terminology. SAMA Standard PMC 20-2-1970.

Von Alven, W. H., (editor) 1964: *Reliability Engineering*, ARINC Research Corporation, Prentice Hall, Inc. p.6.

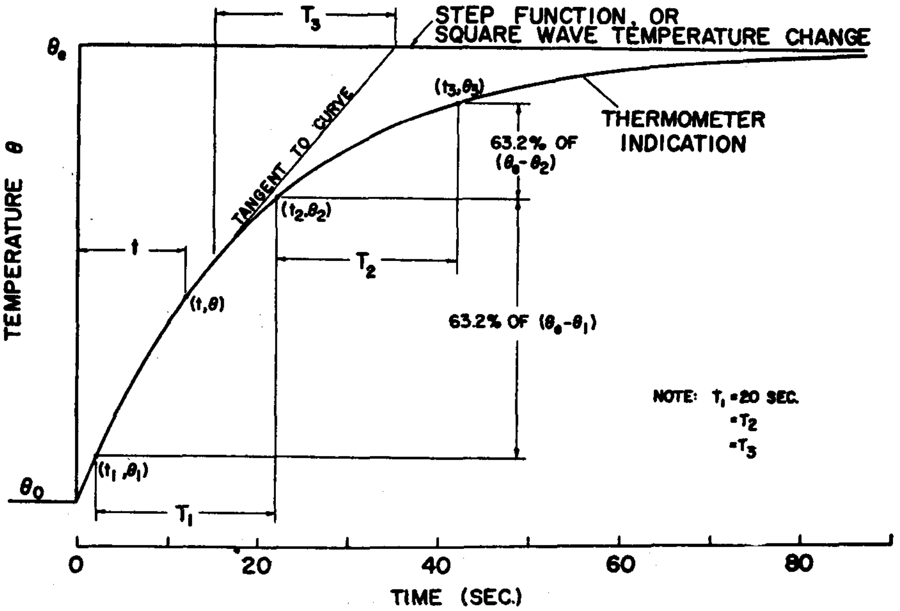

FIG. 1. Response to a thermometer at temperature θ_0 and time constant T to a sudden change in the environment (step function) to a new temperature θ_e.

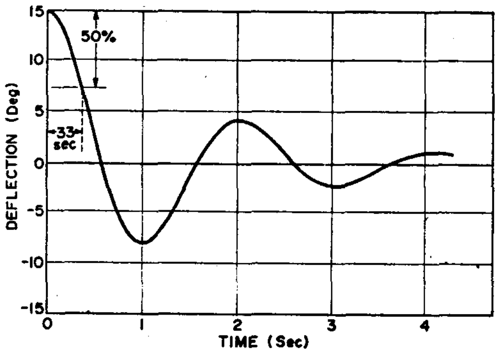

FIG. 2. Oscillation of a typical wind vane exposed in wind tunnel when held 15° from axis of tunnel and then released. (Tunnel speed, 30 ft/sec; vane damping ratio, h = 0.2)

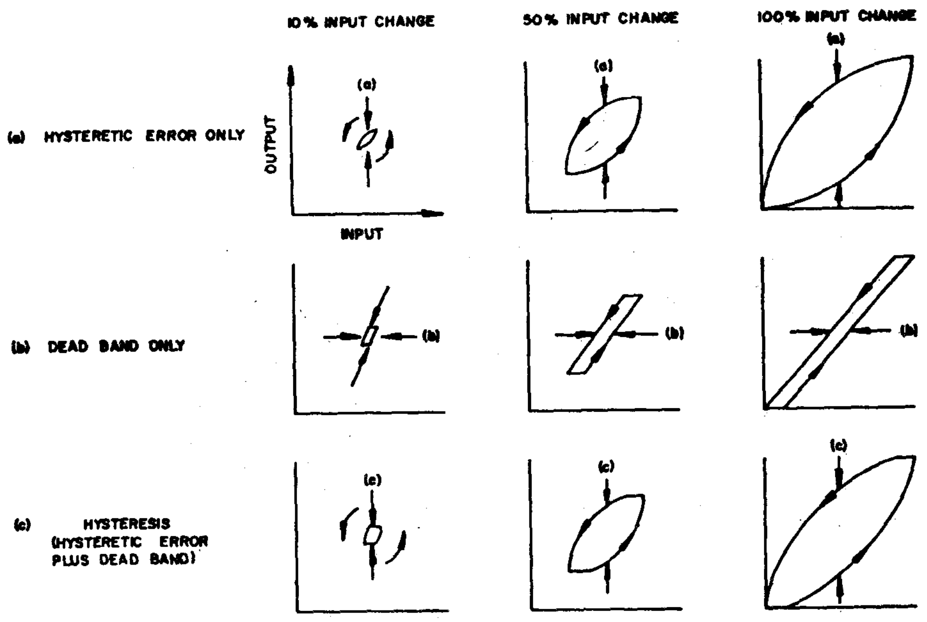

FIG. 3. Hysterectic error, dead band and hysteresis, (exaggerated for easy understanding). (From SAMA Standard PMC 20-2 1970.)

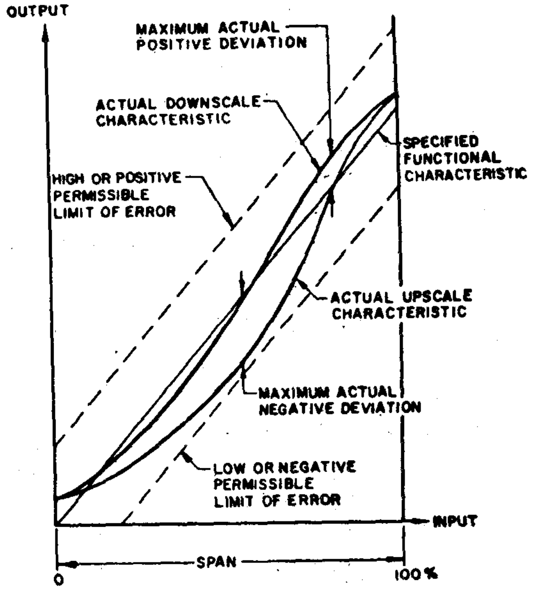

FIG. 4. Reference accuracy (exaggerated for easy understanding). (From SAMA Standard PMC 20-2-1970.)

ON THE NEED FOR DEVELOPING METEOROLOGICAL TEST STANDARDS[1, 2]

DAVID J. BEAUBIEN
EG&G, INC.
Bedford, Massachusetts 01730

Abstract- This paper outlines an American Meteorological Society-sponsored approach for developing a system of standard test procedures for the basic meteorological instruments. The approach involves establishing a number of working groups, under the direction of a suitable committee, composed of members from both industry and the profession. The approach is parallel to similar approaches successfully taken by other professional societies faced with similar problems.

I. INTRODUCTION

The need for standards for meteorological instrument manufacturers has been a frequent topic during the past decade, as the quantity and sophistication of meteorological instrumentation have increased. Users and manufacturers alike have asked for a policing function by a non-partisan organization to minimize the occasional but increasing number of disputes over claims of instrument performance, and to act as an interpretive aid to the purchaser of environmental monitoring equipment. Whether the problem has reached a sufficient magnitude to warrant the development of a solution at this time might be argued, for it can be shown that the necessary solution is a major undertaking for the profession. Developing and implementing a successful standards program for a profession even as small as our own requires the expenditure of many man-hours of effort over a long period of time. However, the need for generally-accepted equipment standards is likely to increase, at least within the United States, as new environmental pollution legislation is enacted. In many instances, the new statutes will require measurement of meteorological variables by the industrial polluter, using unspecified instrumentation. It would seem that we as meteorologists have the responsibility for structuring a method by which some of the present problems regarding a uniform instrumentation language can be minimized.

The concern over the lack of "standards" first began to surface several years ago. Poncelot (1959)[1] suggested the "standard equipment" approach to rainguage testing. This approach entails defining a standard instrument to which all other instruments are compared. The standard instrument concept is a valid solution to many of the basic problems of metrology. However, it is not a general solution for applied instruments, and its adoption limits the effective testing of improved instruments as they evolve. Lamb and Pharo (1966)[2] proposed that the concept of "standard tests" be adopted by the profession. Again in 1967, Beckman,[3] then active on the AMS Committee on Atmospheric Measurements, published a survey paper on the problem in the Bulletin of the American Meteorological Society, recommending that responsibility for the effort be given to the Committee on Atmospheric Measurements. In the past several years, as more non-meteorologists have found it necessary to measure the basic meteorological variables as part of the overall environmental monitoring problem, the need for standards has increased, as evidenced by increased requests to the American Meteorological Society for assistance. Unfortunately, the need for more standardization only surfaces when a problem develops, and this is too late for corrective action to be taken. Several times a year procurement officers find themselves in dispute with vendors over performance of purchased equipment. More often than not, the performance dispute is a result of inadequate specification of the performance criteria, both by the purchaser and the vendor.

II. TYPES OF SOLUTIONS AVAILABLE TO US

What is a standard? Unfortunately, the problem is viewed differently by the various user groups within the profession. Webster says simply: "a standard is that which is set up and established by authority, custom or general consent, as a model or example; criterion; test; in general, a definite level, degree, material, character, quality, or the like, viewed as that which is proper and adequate for a given purpose." Twenty-three other definitions then follow! The variability of the concept of a standard within the profession can be seen from the use of the word

---

[1] Paper presented at the AMS Second Symposium on Meteorological Observations and Instrumentation, San Diego, Calif., 27-30 March 1972.

[2] Editor's Note: Because of its interest to developers as well as users, this paper will be published in the November 1972 issue of the Bulletin of the American Meteorological Society.

Manuscript received Aug. 11, 1972

in the subjects covered in the session on standards at the Second Symposium on Meteorological Observations and Instrumentation in San Diego, Calif., March 1972.

The suggestions made in the past for a standardization program for meteorological instrumentation have, for the most part, stemmed directly from the procurement problems. Equipment users have asked for an "AMS-Seal of Approval" program for the meteorological equipment manufacturers, in the hopes that this would guarantee the purchaser that the equipment was "as advertised." The American Meteorological Society can not perform this function, nor would such a function serve to correct the basic problem without acceptable criteria being first developed. From time to time it has been suggested that the National Bureau of Standards assume the responsibility for establishing meteorological standards for manufacturers. Again, this is not within the charter of the NBS and even if it were, budgetary restrictions at NBS would probably preclude their undertaking such an enormous task, particularly when one considers that professions would likely come to the NBS for similar assistance, once a precedent was established.

The problem we are facing is not a new one; it has been faced by other professions with similar responsibilities to its members and society. We can, therefore, look to them for guidelines in developing a solution. A variety of avenues are available to us, including turning over the entire standards development problem for meteorological instrumentation to a sister organization specializing in instrumentation, such as the Instrument Society of America or the Scientific Apparatus Makers Association. Another approach would be to engage the services of the American National Standards Institute (ANSI), an organization which specializes in the establishment of standards. Lastly, we can undertake the task within our own society. Regardless of the approach taken for developing a solution to the problem, the single most important criterion for success in generating a working standard is structuring the approach taken in such a manner that the resulting standard will be fully acceptable to the using group when it is completed. Many organizations have spent countless man-hours in committee meetings developing standards to find that, upon completion of the task, usually after a period of several years, the result is not generally adopted within the profession.

A. What is needed - The concept of a standard test procedure

Ideally what is needed is a set of rules by which both the manufacturer and the purchaser agree to live by, which results in the equipment performing the way the purchaser thinks it ought to, to solve his problem! One cannot hope to achieve such a euphoric solution. The best solution one can hope for is the establishment of Standard Test Procedures to which manufacturers and users alike reference the performance characteristics of the equipment under consideration. The basic purpose of the Standard Test Procedure is to specify uniform methods of measuring performance characteristics of equipment. Numerous models for such standards can be found in many of the other professions. Standards for testing electronic components and systems such as microphones, amplifiers, etc., have been developed and maintained by the Institute for Electrical and Electronic Engineers. These standards are widely used and universally accepted within these engineering professions. However, no comparable universally accepted Standard Test Procedures are employed within the meteorological community for performance testing even of the simplest instruments such as thermometers, hygrometers, anemometers, barometers, etc. This is not to imply that such test procedures do not exist! In fact, each military agency and most of the civil agencies concerned with meteorological instruments have their own test procedures for the particular equipment types they purchase. However, there are no universally accepted Standard Test Procedures available to and endorsed by the overall profession for the majority of basic instruments we use.

The methods for developing Standard Test Procedures vary considerably depending on the nature of the problem, the complexity of equipment, the structure of the society responsible for the task, and many other variables. However, the following general guidelines for developing such standards usually prevail:

a) The manufacturers and the user groups within a particular geographical area, in this case the United States, must mutually agree that such standard test procedures are needed. This agreement is usually reached by polling the members of societies that use the equipment.

b) Assuming that there is agreement on the need for a particular standard, users and manufacturers must agree on which society will have the responsibility for developing Standard Test Procedures.

c) The selected society then examines its committee structure and selects or creates a committee to have the overall responsibility for developing the procedures. The rules governing the function of this committee must be such that a continuity of membership prevails with sufficient overlap in active terms so as to provide continuity of interest and effective performance over a long period, sometimes as long as ten years. An initial task for the general committee is to establish a glossary of terminology for use by all subcommittees and to establish a format for the Standard Test Procedures. (Work towards a uniform glossary has been in process for some time, the most recent contribution being that of Gill and Hexter (1972).[4]

d) Once the general committee is established, subcommittees are created to study, evaluate and recommend a common test procedure for equipment types for use within the profession. Usually, one subcommittee consists of from four to ten people focusing on only one or two equipment types. Seldom must new test procedures be developed: the usual course the subcommittee follows is to select and adapt some existing standard to the new format.

e) The subcommittees meet as often as possible - usually four to six times a year - over a period of several years, before a suitable Standard Test Procedure evolves. The work is split up among the members as a function of the capabilities and talents represented on the committee.

f) Lastly, as the individual Standard Test Procedures are completed, they are made available to all of the manufacturers and user groups for implementation, the most critical step in the process.

B. Existing activities on the standards problem

At the present time, both the American Meteorological Society and the World Meteorological Organization sponsor committees which are actively engaged in matters affecting meteorological instrumentation. These committees are the Committee on Atmospheric Measurements (CAM) and the Commission for Instruments and Methods of Observation (CIMO), respectively. Joint membership on these committees by a few individuals ensures cross fertilization of ideas and policies. These committees were created as sounding boards for their respective societies on matters having to do with meteorological instrumentation. The scope of activities of these committees is substantial. The activities of CIMO are particularly wide ranging and include working groups in the fields of meteorological radar, automatic weather stations, meteorological rocketry, upper air sounding systems, satellite measurements, etc. The nature of the World Meteorological Organization is such that the efforts of the working groups are primarily directed towards solving instrumentation problems that are international in scope. These working groups are burdened with the development of international standards of measurement and establishing priorities for development of new measurement systems.

The American Meteorological Society's Committee on Atmospheric Measurements tends to focus its activities on atmospheric measurement problems within the United States, with close cooperation with Canadian interests. Although both committees fully recognize the need for additional instrumentation standards within the profession, progress towards development of "working standards" such as standard test procedures has been negligible. A basic problem both committees face is the infrequency of meetings and the lack of any full time activity between meetings. Funds for supporting such activity are simply not available from organizations' coffers. However, if the problem of instrumentation standards is to be solved, time and money will be required. The necessary investment would appear to be minor in light of the savings to be gained by the profession. In fact, if patterns of standard test procedure implementation in other societies are representative, the sale of test procedure documents themselves by the sponsoring society is a source of significant income to the sponsor.

III. DEVELOPING UNIVERSALLY ACCEPTED STANDARD TEST PROCEDURES

It is worthwhile to review the process for developing standard test procedures in some detail following guidelines described previously.

A. The polling process

In the United States, the most widely read meteorological periodical is ~~the~~ Bulletin of the American Meteorological Society. This medium provides a means for soliciting the attitude and recommendations of the profession regarding the necessity for the standards program. A simple questionnaire inserted with the regular monthly mailing of the Bulletin would serve to ascertain whether our profession is ready for a full-fledged standards program. Alternately, a letter questionnaire directly to the membership can be designed so as to provide both an indication of the need and an assessment of the ingredients of the standards. Questions identifying which equipment areas need the earliest attention, the extent and type of test required, etc., should have early exposure to the profession in order to provide feedback to guide the committee in its work. The polling process must be considered an essential and on-going process of the standards effort.

B. Selecting the responsible society

We have a natural tendency to selfishly assume that our own American Meteorological Society is best qualified to undertake the task of developing a standards program. This, in fact, may not be the correct course for us to pursue, in view of the fact that we have never engaged in such activity in the past, and that only a small percentage of our membership is engaged in instrumentation. Several of our sister societies, on the other hand, have elaborate committee structures and have successfully demonstrated their ability to develop and implement such programs. Therefore, an additional objective of the polling process should be the solicitation of opinions as to which society should be given the responsibility.

C. Committee structuring

The committee charged with developing standards must be a working committee. It is not unusual for such committees, if they are successful, to meet six times a year in strenuous 12- to 14-hour sessions, taking home another week's work at the conclusion. Unfortunately, few qualified people can afford to donate this amount of effort to the cause. A second problem in committee structuring relates to balanced participation. Committees must be composed of representatives of each sector of our profession. Committee membership representing both industrial and scientific users, and from manufacturers themselves, must be carefully balanced, lest the resulting standard be weighed in favor of one or the other's interest. A method for stimulating such a committee would seem to fall under the responsibility of the Committee on Atmospheric Measurements.

D. Implementing standard test procedures

Upon completion of the standard test procedures by the subcommittees, the critical task of implementation is faced. If the tests are not successfully implemented the entire effort has been in vain. If general acceptance of the standards is to be achieved, implementation must start with the manufacturer. By involving the manufacturer as well as the user in the committee process, the manufacturer should find the standard in which he himself participated, to be acceptable. The manufacturer must then take the initiative in utilizing the standard test criteria in defining the performance of his products. One role that the AMS can play at this juncture is to encourage all Bulletin advertisers to include performance criteria in their advertising, in accordance with the language of the applicable standard. The manufacturer's advertising literature must contain specific reference to the standard test procedure rather than broad generalizations as to performance. Phraseology such as "Model XYZ when tested in accordance with AMS Standard Test Procedure performed as follows:" must become the accepted method of presenting specifications on data sheets. Once this stage of standard development is reached, the user will find it to his advantage to specify his equipment performance requirements in terms of or with reference to the Standard Test Procedure. The implementation of the procedure effectively adds considerable meaningful technical detail to the specification, without requiring the user to become an expert on the particular equipment type. Its use aids the manufacturer in defining his product, while at the same time insuring to the purchaser that the equipment is being described by a language which has been universally accepted by other users.

IV. CONCLUDING REMARKS

The need for establishing "standards" within the meteorological profession is increasing with the timely enactment of environmental legislation. Standard Test Procedures for use both by manufacturers and users are employed successfully by many other industries as an effective means of presenting equipment performance criteria and to minimize fraud and simple misunderstanding in the procurement process. The task of developing and successfully implementing such standards is known to be time consuming, typically taking from five to ten years depending on the complexity of the equipment involved. On the assumption that the arguments presented in this paper are supported by the general opinion of the profession, it is recommended that increased attention be given by the existing committees within the meteorological societies to developing a structure by which Standard Test Procedures can be brought into existence.

REFERENCES

1. Poncelot, L., 1959: Comparison of rain gauges. WMO Bulletin, VIII, 186-190
2. Lamb, R. C. and J. A. Pharo, 1966: Standardizing meteorological instrument testing. Presented at Seventh National Conference on Agricultural Meteorology of the Amer. Meteor. Soc.
3. Beckman, J. C., 1967: On the need for standardization of instruments calibration and related terminology. Bull. Amer. Meteor. Soc., 48, 704-705.
4. Gill, G. C., and P. L. Hexter, 1972: Some instrumentation definitions for use by meteorologists and engineers. Bull. Amer. Meteor. Soc., 53, 846-851.

STANDARDIZING FUNCTIONAL TESTS

Walter E. Hoehne
Functional Experimentation & Test Branch
National Oceanic and Atmospheric Administration
National Weather Service
Sterling, Virginia

Abstract-A standardized test for evaluation of meteorological measuring systems in the natural environment is described. The test is designed to provide a quantitative statistical value that will indicate the reliability of a particular system output. This quantity called functional precision provides a quantitative estimate of the difference in readings that can be expected from systems of identical design and construction when exposed to the same environmental conditions. The mathematical definition of this parameter is described and methods of application are discussed. Two specific examples are presented. A new surface sensor for measuring wind gust is compared with present observational practice and the functional precision determination for an upper air sounding system is summarized.

Introduction

The adequacy of proposed new equipment is determined for the National Weather Service by the Systems Development Office, Test and Evaluation Laboratory. In addition to questions of accuracy and general utility, one question that must be answered is: "What change will there be in data provided to the user when a new system is adopted?" To answer this question, comparison is made between the output of the new system and the output of a system already in use. The Functional Experimentation and Test Branch has developed a program to standardize the evaluation of differences in output from meteorological instruments. Functional is used here to indicate tests made with the equipment being operated in the natural environment and not under controlled laboratory conditions.

The value of a particular measurement for meteorological purposes has in most cases not been objectively defined. Some efforts are being made to make such definitions. For example, the WMO Commission for Instruments and Methods of Observation (CIMO) has appointed a committee to define temperature for meteorological purposes. Physical measurements may be defined in terms of physical phenomena (e.g., the phase changes of water were chosen as two points on a temperature scale). Meteorological measurements are not so clearly defined. A measurement for meteorological purposes is associated with physical conditions, with a volume larger than the volume immediately in contact with the sensor and an arbitrary time period. Perlat and Petit[1], and Bragenskaia and Kagan[2,3] have investigated the problems of associating instantaneous point measurement with a time and/or space domain. Many investigators have addressed themselves to the accuracy of particular instruments and laboratory methods for determining accuracy. Recently Beckman[4] proposed a means of setting up standards for instrument parameter definitions and for test procedures. Lamb and Pharo[5] also proposed standardizing meteorological instrument testing. In both cases, the proposed tests are those to be conducted in a laboratory with a controlled environment that simulates the natural environment. The variability in reading due to the natural variability of the atmosphere can be considered only to the extent that such variability can be simulated.

The Meteorological Working Group (MWG) Inter-Range Instrumentation Group (IRIG) compiled a set of accuracies for meteorological equipment used on the National Missile Ranges[6]. The MWG has revised that document[7] and in it a new concept is employed expressing "reliability" of data rather than accuracy. "Reliability is defined as the best available quantitative estimate of the quality of the data for operational use at the test ranges. Where possible, and as noted, the term reliability includes errors resulting from human and instrumental sources. Where standards have been established, reliability is a statement of accuracy. In general, however, the values of reliability are statements of data precision to be expected from well maintained equipment, operated by competent individuals according to a well defined procedure." The program described here is an attempt to standardize one test for reliability of meteorological data.

In the past when a new sensor system was developed, its data reliability was evaluated by comparing it with an existing system. Output differences between the two systems were tabulated and analyzed statistically to produce mean difference,

Manuscript received Aug. 11, 1972