Instrumentation

PRECISION OF ATMOSPHERIC MEASUREMENTS

by W.E. Hoehne (1)
and R.J. Stone (2)
NOAA NWS Test & Evaluation Division (3)
Sterling, VA 22170

ABSTRACT

The accuracy of meteorological measurements cannot be determined because there is no demonstrable true value with which to compare the operational measurement. As an alternative the functional precision (repeatability) of measurements is determined to provide a means of evaluating the variability of measurements in a synoptic field. Mechanization and automation of the observation process enhance the requirement for determination of precision of the measuring process. Elimination of the human observer requires the use of processing algorithms to replace the human thought process. These algorithms require quantitative definition of variables previously reported subjectively. Some determinations of precision and comparability have been made but much more effort is needed.

1. INTRODUCTION

Traditionally, atmospheric measurements for meteorological purposes are made by observers using manually operated sensors following the guidelines and instructions provided by national and international meteorological organizations. (4) The quality of the measurements depends upon the training of the observers as well as the laboratory accuracy of the instrumentation that they use. Meteorological measurements are point samples from a time and space continuum and it is important that the measurements made under identical conditions provide identical results. Accuracy is desirable, but repeatability is essential to a realistic representation of the synoptic measurement field.

Before operational use of a new meteorological sensor system, answers are needed to a number of questions such as: "What is the accuracy of the new system?", "What change will there be in the data provided to the user when the new system is adopted?", and "What will be the variability of measurements in a network containing such systems?". The need to answer these questions in a systematic manner led to the development of the standardized functional testing program. The program addresses all but the question of accuracy, which is determined in a laboratory.

The introduction of electronics to meteorological instrumentation separated the sensors from the observers, and the use of automatic systems eliminates the human from the process of data collection. Increasingly, observers read dials or digital displays of measurements made by sensors whose output is converted by transducers into an electrical signal. Automatic observing systems process the electrical signals and transmit meteorological data directly to the user.

Many of the traditional measurements are a subjective evaluation by the observer. Cloud height, visibility, precipitation intensity and other measurements used in synoptic observations are a subjective output of the observer produced by following the training and guidance provided. Recently, more and more attempts have been made to produce such observations automatically using electro-mechanical and electro-optical sensors with automatic data processing as a substitute for the observers subjective evaluation.

To judge the quality of data obtained from a variety of sensors in a synoptic field it is necessary to compare the measurements made by such systems under real operational conditions. As new sensors and systems are developed it is necessary to know the comparability between the new sensors and those that are already in use. Comparability must be determined for chronological continuity and to assure the consistency of data from a mixture of sensors used over a wide geographical area. In addition the automation of subjective measurements requires the definition of those measurements in a quantitative numerical form.

Accuracy of an instrument can be determined in a laboratory by controlling the conditions to which the sensor is subjected, but the entire range of natural atmospheric conditions cannot be simulated in the laboratory. For that reason some means of determining variability must be provided. This can best be accomplished by determining the precision of repeatability of the system when it is operating in the external environment under actual or quasi-operational conditions. Following the suggestion of Ku (5) this determination is called "functional precision" (i.e., the precision of a system when it is performing its designed function). Accuracy is determined in a laboratory under controlled conditions where the measurement made by a system is compared with the known value. Functional precision is determined in the natural environment where simultaneous measurements of an unknown value made by identical instruments are compared.

After the establishment of the Test Evaluation Laboratory of the Weather Bureau in the early 1960's, newly developed sensors were sent for evaluation in environmental chambers and wind tunnels and in the natural environment. For operational evaluation newly developed systems were compared with systems that had been in use for a number of years. There was no standardized methodology for making such comparisons and when a difference between simultaneous measurements by an old and a new sensor was observed there was no information on the difference that could be expected in simultaneous measurements by two of the old systems. Hence, evaluation of the size of the difference between the old and the new sensor was almost impossible. For this reason an attempt was made to standardize functional (operational) testing as reported by Hoehne (6) to the Second Symposium on Meteorological Observations and Instrumentation. The rationale and methodology for this standardization is described in IEEE Transactions on Geoscience Electronics (7). Over the years the methods have been refined and developed and recently were submitted as a recommended standard practice, for balloting by the American Society for Testing and Materials.

The production of precision determinations in accordance with this methodology is a low-priority effort at the Test and Evaluation Division. Data accumulated for other test purposes is frequently used for precision determinations and specific determinations are made on a time available basis. As a result the total production to date has been relatively small. The resources committed to this program are insignificant.

2. METHODOLOGY

Functional precision of a meteorological measuring system is defined as the root mean square (rms) difference between a large number of simultaneous measurements of a particular atmospheric variable made by identical sensors in a specified atmospheric volume. Identical sensors are defined as being the same make and model. The atmospheric volume is restricted to a diameter of 10 meters and a depth of 1 meter or one-tenth the height of the sensors above the surface, whichever is smaller. Simultaneous measurements are defined as being separated by not more than one tenth the time constant of the sensor, but not more than 1 second. Individual pairs of measurements must be separated by at least four times the time constant of the sensor to insure that the statistical samples are independent. If the systems used to make the measurements are different, that is, a different manufacturer, or different model, or different design, then the root mean square difference obtained from the process is called comparability. Taken together, the two determinations (functional precision and comparability) can provide information on the variation of data obtained by a mixed network of sensors of two or more types. If the functional precision of a measuring system is known and the comparability between that measuring system and a different measuring system is determined, then the amount of variability introduced into a network as new sensors replace old sensors can be established.

To make the results of functional precision and comparability determinations consistent in time and repeatable from laboratory to laboratory, it is necessary to standardize the sampling procedure used and the format of the results. The Subcommittee on Meteorology of the American Society for Testing and Materials (8) has a standard method for determining comparability and functional precision of meteorological measurements under development. Action on this standard could be completed by mid 1984 and the standard should by available in the 1984 Book of Standards. In addition to the root mean square difference, the output format of the standard requires a report of statistical parameters such as mean difference, standard deviation, skewness and kurtosis. The method assumes that differences are randomly distributed about the mean difference and these parameters are used as a means of verifying that assumption. The standard also requires calculation of the correlations between the value of the measurements and the size of the difference between them. In that way an estimate of the change in variability with change in variable can be determined. Other parameters are also required by this standard practice including the range of measurements and a calculation to determine the reliability of the distribution.

An attempt was made to produce a sampling procedure that would be widely applicable to all sorts of meteorological variables. In some instances this presents special problems due to time and space discontinuities and the method by which a particular variable is measured. Care must be exercised to insure that determinations are made by comparing the same variable. For example, comparing a wind speed and direction measurement from a rotating-cup anemometer and vane with the measurement made by a laser system of the average wind speed and direction over an extended path would produce a special type of comparability. Obviously the point measurement made by the anemometer is not the same variable as the volume measurement made by a laser system. In some instances one of the measurements of position is actually the variable of interest (e.g., cloud height). Or the sensor system itself may exceed the size limit of the method. For example, the projector and detector of a transmissometer are more than 10 meters apart. In the first case the corresponding elements of the sensor system (i.e., the projector and detector of the cloud height measuring system) are located within the horizontal restraints and height is treated as a meteorological varible of interest. In the other case the corresponding components of the sensor (the projector and receiver) are located within the 10-meter by 1-meter volume and all other parts of the method proceed as specified.

As new technologies bring new instruments, the challenge to compare instruments or systems, as well as the danger of false comparison, increases. The radiosonde, for example, has for many years been the upper air measurement system from which vertical profiles of temerature, humidity, pressure-height, and wind information are obtained. At operational radiosonde stations, instruments are launched once every 12 hours and the profile is developed as the balloon ascends through the atmosphere for 90 minutes or so. The data is presented as a vertical profile even though the balloon might have drifted a hundred miles or more from the launch site. An atmospheric profiler system developed by NOAA's Wave Propagation Laboratory in Boulder (9) uses a passive six-channel microwave radiometer to provide continuous vertical profiles of temperature, humidity, and liquid water. UHF and VHF clear-air Doppler radars provide continuous vertical wind profiles through the tropopause with updates currently available every 20 minutes.

The data from both systems are presented to the user as vertical profiles of atmospheric varibles and comparability must be determined if data from a mix of profilers and radiosonde observations is to be used in a synoptic network. Other applications of those comparability determinations must be done cautiously with full knowledge of the real differences in the measurement process.

As information from new technologies are brought into the meteorological data base, the users must be made aware of the possible changes. The NWS is (for example) in the midst of a program to replace its substation temperature measuring equipment, (the "cotton region shelter" with Townsend Support, mercury-in-glass maximum temperature and spirit-in-glass minimum temperature thermometers) with a small plastic shelter containing a thermistor. The new system may have the same laboratory accuracy as the old, but has shorter response time when exposed to actual operational conditions because of its smaller volume. The wooden shelter dries more slowly after being wet by precipitation, fog, or dew, causing evaporative cooling to alter the measured air temperature. The cumulative effect of these differences may be a subtle change in the reported temperatures that might be erroneously interpreted as a real climatic change. Although the standard recommended practice for determining comparability of meteorological measurements is structured to be as broadly applicable as possible, there remains an area where common sense and good scientific practice must prevail.

3. PROSPECTS FOR THE FUTURE

Increased mechanization of the meteorological data gathering process will require the determination of precision and comparability for a wide range of meteorological measurements if the user of those measurements is to know what variability can be expected. Automation of atmospheric measurements has been increasing and can be expected to increase continuously in the near future. Booz Allen and Hamilton (10) recommend and increase in the efforts toward automation. Within the National Weather Service the general trend away from manual observations produced by trained observers, to automatic observations processed and transmitted by machines, can be expected to accelerate within the near future.

Insuring continuity in climatological data and synoptic reports made from a mixture of automatic and human observations will require an increased effort in determining the precision and comparability of such observations. As old equipment becomes obsolete and is being replaced with new sensors which use new technologies to measure the "same" atmospheric elements, and as the pace of automation increases, the importance of precision determinations increases. In addition, more effort must be directed toward definition of the variables now reported subjectively by the observer. Such definitions must be quantitative and numerical so tha the quality of automatic measurements can be established.

The National Weather Service has begun to automate subjective portions of the surface observation such as clouds and visibility. Measurements from a number of sensors were compared to human observations reported to the FAA (11). Generally, there is good agreement. The main point to be made is that automated observations standardize the values reported for visibility and clouds but those values will be somewhat different from what the user gets today.

Meteorologists have never strongly supported determination of meteorological data quality. The academic community is not interested in the production of precision and comparability measurements because there are no theses or research papers to be written from such efforts. Most meteorologists are at least partially research scientists who have depended on the skill and training of observers and have by and large assumed that this quality is maintained even though measurements are less and less frequently made by a human observer and more and more frequently by automated systems. Repeatedly, gross errors have been made by meteorologists who have accepted meteorological measurements at face value assuming either zero or very small variance. Some horror stories concerning this problem were related by Giraytys (12).

About 10 years ago the American Meteorological Society recognized the need for information on the variability of atmospheric measurements and the demand generated by laws written to protect the natural environment. Beckman (13), Lamb & Pharo (14), Beaubien (15), Hoehne (16) and others have cited the need for information on the operational accuracy of meteorological data and procedures for determining this accuracy. In addition, non-meteorological organizations such as The American Society for Testing and Materials, The American Nuclear Society and several other standard-setting organizations have instituted efforts to prepare procedures by which the operational accuracy of meteorological measurements can be determined. At least one non-profit institution (Meteorological Standards Institute, Fox Island, Washington) has been established to help fill the need for standardized tests and procedures.

Office of Management and Budget and Department of Commerce directives encourage federal agencies to use voluntary standard procedures and participate in their development, but recent and expected future cuts in the federal government will reduce the production of precision and comparability determinations made by government supported laboratories. If present trends continue the very small effort presently supported by the National Weather Service can be expected to disappear within the next few years. The purchaser of meteorological instrumentation and the user of meteorological data must depend more and more on the the determination of precision and accuracy provided by non-government organizations.

4. RESULTS

4.1 Surface Measurements: A large amount of the meteorological information produced each day can be classified as surface data. These are measurements made within 10 to 20 meters above the surface of the Earth, and for the most part, point measurements of meteorological variables. Stone (17) reported on a program by which over the past 14 years the Test and Evaluation Division of the National Weather Service has been able to determine the functional precision and/or comparability of measurements made by a few of the wide variety of sensor systems available. The variables for which such determinations have been made include air pressure; temperature (current, maximum and minimum); relative humidity; dew point; wind speed; wind direction, accumulation of precipitation; cloud height and solar radiation amount. Table I lists the determinations that have been made.

4.2 Upper Air: Measurements made by balloon-borne sensors, aircraft or indirect means from about 20 meters above the surface to near the top of the atmosphere are referred to as upper air variables. These include pressure, temperature, dew point depression, and/or relative humidity, pressure height, wind speed, and wind vector. Hoehne (18) has determined the precision of the National Weather Service system for measuring these variables as given in Table II. The difference between simultaneous measurements was in general independant of the size of the measurement. There was a weak correlation between the rms difference in height and the measured pressure and rms difference in pressure and the calculated height. The functional relationship is given along with the overall precision.

5. CONCLUSIONS

The increase in the use of electromechanical measuring systems for meteorological variables increases the need for determinations of precision and accuracy of the measurements. The quality of forecasts, both those made manually and those made by machine, is largely dependant on the quality of the meteorological data used in forecast preparation. To attempt to improve the quality of this data it is first necessary to determine its present quality, then measure the improvement. The small efforts of the Federal Government to produce such information can be expected to decrease continually with time and it will be more and more the job of the private industry to produce such determinations. If operational meteorologists are interested in the quality of the data they are provided and hope to see an increase in quality as well as quantity, they should support the production of standard methods for determining meteorological data quality, and encourage instrument and system producers to include determinations of functional precision and comparability in the advertised specifications of their instrument systems.

FOOTNOTES AND REFERENCES

1. Walter E. Hoehne is Chief of the Functional Experimentation and Test Branch of the National Weather Service Test and Evaluation Division. His interests include the evaluation of new methods for obtaining, processing, transmitting and displaying meteorological data with special emphasis on the impact on the user of data produced by the new methods. He has supported the standardization of meteorological measurement procedures and the tests to determine measurement quality.

2. Richard J. Stone is a Functional Test Program Manager in the Functional Experimentation and Test Branch of the National Weather Service Test & Evaluation Division. His areas of interest include laser remote sensing of the atmosphere, automation of meteorological observations, instrument development and testing in general. He was active in meteorological measurements related to nuclear power plant siting and was on an ANS working group which developed standards for those measurements.

3. The Test and Evaluation Division of the National Weather Service evaluates systems and instruments used and/or proposed for use by the National Weather Service for gathering, processing and displaying meteorological data.

4. Leftowitz, Matthew and J. T. Bradley, 1978: Relationship of Automated Weather Observations to Subjective Elements. Air Quality Meteorology and Atmospheric Ozone, STP 653, American Society for Testing and Materials, Philadelphia, PA, pp. 35-52.

5. Ku, Harry H., 1967: Statistical Concepts in Meteorology. The Handbook of Industrial Meteorology, American Society of Tool and Manufacturing Engineers, Prentice Hall, Inc., New York, NY, pp. 20-50.

6. Hoehne, Walter E., 1972: Standardizing Functional Tests. Preprints: Second Symposium on Meteorological Observations and Instrumentation, San Diego, CA, Amer. Meteor.Soc.Boston, MA, pp. 161-165.

7. Hoehne, Walter E., 1973: Standardizing Functional Tests. IEEE Transactions on Geoscience Electronics, Vol. GE-11 No. 2, The Institute of Electrical and Electronic Engineers, Inc., New York, NY, pp. 95-100.

8. Hoehne, Walter E., 1982: Development of Meteorological Standards. Bull. Amer. Meteor. Soc. 63, pp. 304-305.

9. Hogg, D. C., M. T. Decker, F. O. Fuirand, K. B. Earnshaw, D.A. Merrit, K. P. Moran, W. B. Sweezy, R. G. Strauch, E. R. Westwater and C. G. Little, 1983: An Automatic Profiler of the Temperature, Wind and Humidity in the Troposphere. Journal of Applied Meteorology, 22, pp. 807-831.

10. Booz Allen & Hamilton, 1983: National Weather Service: A Strategy and Organization Concept For the Future. Results of a Study for the National Oceanic and Atmospheric Administration, U.S. Department of Commerce, Booz Allen & Hamilton, Inc., Bethesda, MD, 158 pp.

11. Federal Aviation Administration, 1979: Aviation Automated Weather Observation System (AV-AWOS), FAA AD/A-072 119, 55 pp.

12. Giraytys, James, 1973: Introductory Remarks. IEEE Transactions on Geoscience Electronics, Vol. GE-11 No. 2, The Institute of Electrical and Electronic Engineers, Inc., New York, NY, pp. 81-82.

13. Beckman, J. C., 1967: On the need for standardization of instruments calibration and related terminology. Bull. Amer. Meteor. Soc., 48, pp. 704-705.

14. Lamb, R. C., and J. A. Pharo, 1966: Standardizing meteorological instrument testing. Paper presented at the Seventh Conference on Agriculture Meteorology, New Brunswick, NJ, 30 Aug. - Sept., Amer. Meteor. Soc., Boston, MA.

15. Beaubien, D. J., 1972: On the need for developing meteorological test standards. Bull. American Meteor. Soc., 53, 1103-1106.

16. Hoehne, Walter E., 1978: Standards for Meteorology. Bull. Amer. Meteor. Soc. 59 1156-1158.

17. Stone, R. J., 1978: National Weather Service Automated Observational Networks and the Test & Evaluation Division Functional Testing Program. Preprints: Fourth Symposium on Meteorological Observations and Instrumentations, Denver, CO Amer. Meteor. Soc., Boston, MA 131-135.

18. Hoehne, Walter E., 1980: Precision of National Weather Service Upper Air Measurements. NOAA Tech. Memo T&ED - 16 U.S. Department of Commerce, Springfield, VA 23 pp.

TABLE I

PRECISION AND COMPARABILITY OF METEOROLOGICAL DATA

SURFACE SENSORS

VARIABLE	SENSOR	FUNCTIONAL PRECISION
Altimeter Setting (Pressure)	Kollsman Aneroid Barometer	±0.000" Hg.
Temperature	HO-61	$\pm 0.8^{o}$ F
Temperature	AMOS III-73	$+0.4^{o}$ F
Dew Point	HO-61	$+0.9^{o}$ F
24-Hour Max. Temp.	Mercury-in-Glass Thermometer Cotton Region Shelter	$\pm 0.2^{o}$ F
24-Hour Min. Temp.	Mercury-in-Glass Thermometer Cotton Region Shelter	$\pm 0.2^{o}$ F
Wind Speed	F 420	±1.0 kt
Wind Speed	DARDC	±0.8 kt
Wind Speed	AMOS III-73	±0.7 kt
Peak Wind Speed	DARDC	±4 kt
Wind Direction	F 420	$\pm 5^{o}$
Wind Direction	DARDC	$\pm 22.5^{o}$
Wind Direction	AMOS III-73	$\pm 6.3^{o}$
24-Hour Precipitation	Weighing Rain Gage	$\pm 0.02^{o}$
24-Hour Precipitation	U. S. 8-Inch Gage	±0.3 mm
24-Hour Precipitation	Snowden Pit Gage	±0.7 mm
24-Hour Precipitation	AMOS III-73	±0.4 mm
Solar Radiation	Eppley Pyranometer + Monitor Labs Integrator	±0.35 langley/hr
Visibility	Videograph	±0.2 mi.
Cloud Height	Rotating Beam Ceilometer	$\frac{\pm 0.01309(B^2 + h^2)}{(B + 0.01309h)}$ B = Baseline h = Measured Height

VARIABLE	SENSOR	COMPARABILITY
Temperature	HO-61/AMOS III-70	$+0.7^{o}$F
Temperature	HO-61/AMOS III-73	$\pm 1.0^{o}$F
Dew Point	HO-61/AMOS III-70	$\pm 0.8^{o}$F
Dew Point	HO-61/AMOS III-73	$\pm 2.0^{o}$F
Wind Speed	F420/AMOS III-70	±1.0 kt
Wind Speed	AMOS III-70/AMOS III-73	±0.8 kt
Wind Direction	F420/AMOS III-70	$\pm 7^{o}$
Wind Direction	DARDC/AMOS III-70	$\pm 18^{o}$
Wind Direction	AMOS III-70 AMOS III-73	$\pm 11^{o}$
24-Hour Precip.	U. S. 8-Inch Gage/ Belfort Weighing Gage	±0.03"
24-Hour Precip.	U. S. 8-Inch Gage/ Snowden Pit Gage	0.08T ± 0.4 mm T=24-Hour Precip. in mm

TABLE II

PRECISION OF UPPER AIR DATA

VARIABLE	FUNCTIONAL PRECISION	BIAS
	At same time of flight	
Pressure	± 1.9 mb	
As a function of height H	$\pm (12.1-6x10^{-5}H)$	
Temperature	$\pm 0.67^{o}$ C*	-0.14^{o} C
Dew-point depression	$\pm 3.67^{o}$ C*	0.35 C
Height	± 92.9 m*	-7.6 m
	At Same Height	
Pressure	± 0.7 mb	-0.1 mb
As a Function of Height H	$\pm(1.79-2x10^{-5}H)$	
Temperature	$\pm 0.84^{o}$ C*	-0.19^{o} C
Dew-point depression	$\pm 3.42^{o}$ C*	0.38^{o} C
Wind Vector Difference	6.0 kts (3.1 mps)	
Wind Speed	± 6.0 kts (± 3.1 mps)	
	At Same Pressure	
Height	± 23.7 m*	-4.0 m
As a Function of Pressure P	± (38-.038 P)	
Temperature	$\pm 0.61^{o}$ C*	-0.13^{o} C
Dew-point depression	$\pm 3.26^{o}$ C*	0.35 C

* Precision taken from standard deviation because of bias introduced by heat and humidity of upper sonde in balloon train.

TABLE III

PROCESS FOR DETERMINING OPERATIONAL COMPARABILITY

1. Obtain two or more systems for measuring a particular atmospheric variable.

2. Expose the sensors for those systems in a natural atmospheric volume as defined in Paragraph 1 of Section 2.

3. Record pairs of measurements (one from each of two systems measuring the variable) at time intervals described in Paragraph 1 of Section 2.

4. Sum the differences of these pairs and divide by the number of pairs to provide the mean or average difference.

5. Multiply the differences by themselves; sum the products, divide the result by the number of pairs and extract the square root of the quotient to obtain the root mean square (rms) of the differences. This rms is the operational comparability if the systms are of different make, model or design. The rms is the functional precision if the two systems are of the same make, model, and design.

6. Multiply the rms by itself and multiply the mean difference by itself. Subtract the product of the mean from the product of the rms and extract the square root of the difference to obtain standard deviation.

7. Multiply the standard deviation by three and divide by one increment of resolution (the smallest change in the variable measurable by the system). The number of pairs used should equal or exceed the square of the quotient.

8. Subtract the mean difference from the difference of each pair. Raise these new differences to the third power. Sum these third powers, divide by the number of pairs, and extract the third root of this sum to obtain the skewness of the distribution of the differences. A normal distribution has a skewness equal to zero. Skewness is a measure of the asymetry of the distribution about the mean.

9. Subtract the mean difference from the difference of each pair and raise these new differences to the fourth power. Sum the fourth powers, divide by the number of pairs and extract the fourth root of the quotient to obtain the kurtosis. A normal distribution has a kurtosis of three. Values greater than three indicate relatively more values near or at the mean difference. Kurtosis is a measure of the peakedness of the frequency distribution.