



A11103 914923

NIST
PUBLICATIONS

NISTIR 4998

NIST Length Scale Interferometer Measurement Assurance

**John S. Beers
William B. Penzes**

U.S. DEPARTMENT OF COMMERCE
Technology Administration
National Institute of Standards
and Technology
Manufacturing Engineering Laboratory
Precision Engineering Division
Gaithersburg, MD 20899

QC
100
.U56
4998
1992

NIST

NIST Length Scale Interferometer Measurement Assurance

**John S. Beers
William B. Penzes**

U.S. DEPARTMENT OF COMMERCE
Technology Administration
National Institute of Standards
and Technology
Manufacturing Engineering Laboratory
Precision Engineering Division
Gaithersburg, MD 20899

December 1992



**U.S. DEPARTMENT OF COMMERCE
Barbara Hackman Franklin, Secretary**

**TECHNOLOGY ADMINISTRATION
Robert M. White, Under Secretary for Technology**

**NATIONAL INSTITUTE OF STANDARDS
AND TECHNOLOGY
John W. Lyons, Director**

NIST Length Scale Interferometer Measurement Assurance

CONTENTS

1. Introduction
2. The measurement assurance program
 - 2.1 Definition
 - 2.2 Control standards
3. Process changes and control charts
 - 3.1 M5727 control chart, 1965 through 1985
 - 3.2 Measurement process changes
 - 3.3 M5727 control data through 1985
 - 3.4 Control data, 1971 to 1991
 - 3.4.1 Meter bar M5727
 - 3.4.2 20 inch (508 mm) bar No.6495
4. Re-evaluation of measurement errors
 - 4.1 The need for re-evaluation
 - 4.2 Systematic and random errors
 - 4.3 Procedure
 - 4.4 Results
 - 4.4.1 Recalibration of the temperature, pressure and humidity measurement systems, July - August 1987
 - 4.4.2 The International Temperature Scale of 1990 (ITS-90)
 - 4.4.3 Interferometric zero shift corrections
 - 4.4.4 Laser wavelength
 - 4.4.5 Carbon dioxide content of laboratory air
 - 4.4.6 Interferometer and scale axis alignment
 - 4.4.7 BIPM meter No. 12924; international measurements
 - 4.4.8 The revised water vapor correction in the Edlén equation
 - 4.4.9 Correlation test: length vs atmospheric pressure
 - 4.5.0 Summary of data adjustments
5. Interpreting control charts
 - 5.1 Secular change, 1971 to 1991
 - 5.2 Control charts incorporating 1991-92 data
 - 5.3 Simulating stability in control standards
 - 5.4 The measurement process in an out-of-control state
 - 5.5 Control subintervals
 - 5.6 Long term measurement precision
6. Summary
7. References

Acknowledgements

NIST Length Scale Interferometer Measurement Assurance

By John S. Beers and William B. Penzes

ABSTRACT

This paper is an extension of NBSIR 87-3625, Length Scale Measurement Procedures at the National Bureau of Standards[1]. Results from the measurement assurance program (MAP) for graduated length scales over a twenty six year period are reviewed. Line scale interferometer modifications, measurement procedure changes, computational revisions, and a re-evaluation of measurement errors are described and their effects discussed. The question of whether the observed length changes in the MAP control standards are apparent or real is resolved. Improvements in precision and accuracy are demonstrated.

Key words: Length; graduated scales; interferometry; measurement assurance; uncertainty.

1. INTRODUCTION

The ability to measure accurately is vitally important to quality control in many industrial and scientific operations, and the measurement of length is fundamental. Since publishing NBSIR 87-3625, Length Scale Measurement Procedures at the National Bureau of Standards[1], considerable work has been done with the NIST line scale interferometer to improve and study its performance. This report on the results of that work has a direct bearing on improving high-accuracy length measurements.

2. THE MEASUREMENT ASSURANCE PROGRAM (MAP)

2.1 Definition

Measurement assurance can be defined as a systematic program employing redundant measurements of stable control standards to continually monitor a measurement process, and determine measurement variability and uncertainty. Treating regularly performed measurements as a process is the basis of the program while statistical analysis and control charts provide the means for monitoring. Any organized measurement procedure that generates statistically significant amounts of data can be treated as a process, and can be monitored with controls and analyzed to reveal its characteristics.

In a complex measuring system like the line scale interferometer it is essential to know that all its components are working properly. Substantial errors can enter the process if, for example, the

interferometer is misaligned, the barometer is out of calibration, or the distance between the microscope and the remote interferometer changes during a measurement. The MAP with its control standards, statistical analysis and control charts is an indispensable method for evaluating measurement repeatability and uncertainty, and for detecting process malfunctions. It is also useful for measuring the effects of planned process changes.

2.2 Control standards

A graduated invar meter bar, M5727, has been the principle control standard for measurement process monitoring since starting interferometric length scale measurements in 1966. A 20-inch (508 millimeter) steel scale, No. 6495, was added as a control in 1982. The controls are used not only to monitor the measurement process, but also to detect and quantify the effects of changes made to improve the process. A measurement process such as this one, which runs for decades, should have dimensionally stable control standards, but completely stable materials are rare and perhaps non-existent. This can make it difficult to distinguish real from apparent length changes in the controls.

3. PROCESS CHANGES AND CONTROL CHARTS

3.1 M5727 control chart 1965 through 1985

Figure 1 shows the measurement history of M5727 (0 to 1 meter interval) from 1965 through 1985. Individual measurement values are plotted. A change in the measurement process was made at each vertical line, dividing the plot into periods. The mean value for each period is shown as a horizontal solid line, and the limits shown with dotted lines are $\pm 3\sigma$ where σ is the standard deviation of a single value¹. Each change and the mean value for its period is designated by M1 through M4. The process change associated with each mean is described in table 1. Data taken and changes made after 1985, designated as M5 through M8, are also described in the table but will be dealt with later.

An individual measurement is defined here as the mean of two data sets of two passes each. One set is taken with the bar in the normal orientation (zero graduation at the left) and consists of one pass up the scale from the zero graduation to the terminal graduation (one meter in this case), and one pass down the scale, closing at the zero graduation. Data stops are made at each decimeter in both directions. The second set consists of two passes taken in a similar manner but with the bar in the reversed orientation (zero graduation at the right).

¹ These conventions will be followed throughout for control charts.

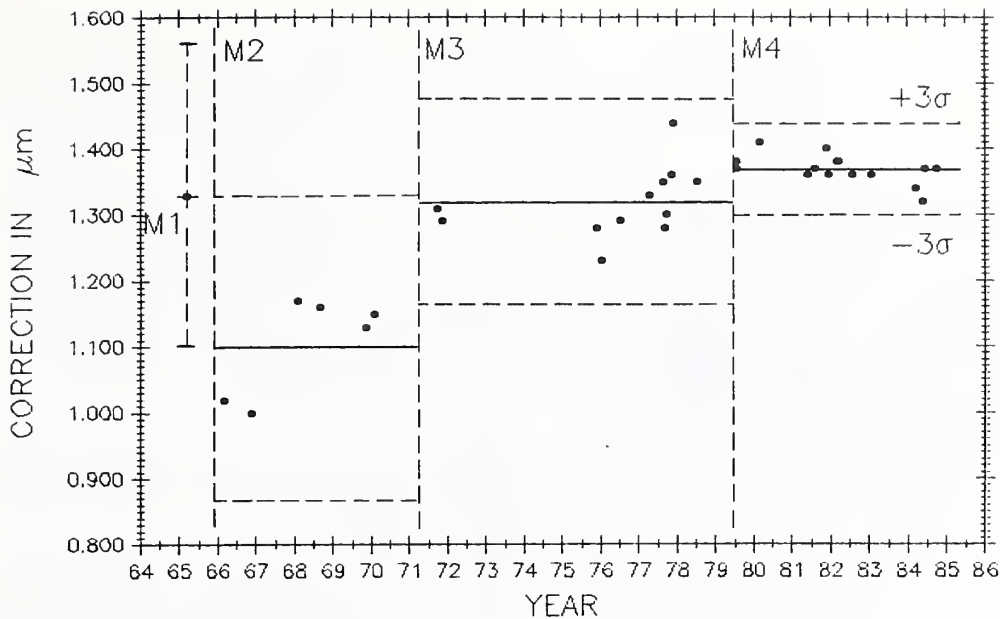


Figure 1. M5727 control chart, 0 to 1 meter, 1965 through 1985

3.2 Measurement process changes

Table 1 covers all changes up to the date of this publication, but the subject will be discussed in two groups, (1) M1 through M4, to show the situation at the time described in reference 1 (through 1985), and (2) M3 through M8, to evaluate the most reliable data.

Table 1. Sequence and Descriptions of Process Changes

Period Symbol	Process Change
M1	Mean of data from classical meter bar intercomparisons (individual points not shown).
M2	Data from the first interferometric measurements.
M3	Data taken after the microscope, beam splitter and reference mirror were rigidly coupled together forming an assembly that was kinematically mounted on the waybed. This was designed to reduce microscope movement cause by the heavy carriage as it moved along the waybed.
M4	Data taken after replacing the original Michelson type interferometer with a commercial interferometer system mounted at the opposite end of the waybed. No carriage pitch or yaw corrections were made after this change, but scales were mounted with their ruled axis coincident with the interferometer axis.
M5	Data taken after the interferometer retroreflector was

moved from its mount on the subcarriage to a mount on the scale support structure (see fig. 2). This reduced the possibility of change, during a scale measurement, in the critical distance between scale and retroreflector.

- M6 Data taken after recalibrating the barometer and temperature system, and replacing the hygrometer.
- M7 Data taken after the retroreflector was mounted on a new invar scale support structure to further reduce chances of a critical distance change. Zero shift (dead path) corrections were applied during this and the following period.
- M8 Data taken after incorporating the modified Edlén equation into the process and mounting a hygrometer inside the interferometer housing.

3.3 M5727 control data through 1985

Table 2 shows the mean value and precision (3σ) for each group shown graphically in figure 1.

Table 2.

Mean Value and Precision for Groups M1 through M4

Group	Mean Value (μm)	Precision 3σ (μm)
M1	1.33	0.24
M2	1.10	0.23
M3	1.32	0.16
M4	1.37	0.07

The M1 period measurements were made in 1965 by comparing M5727 with NBS laboratory standard meter bars in a comparator employing filar microscopes to measure the length differences between standard and unknown in the traditional way[2]. The mean value is plotted in Fig. 1 with its error bar.

The first interferometric measurement group, M2, was made from 1966 through 1970. The M2 mean agreed with M1 within $0.2 \mu\text{m}$. This was considered to be very good agreement for that time.

A major structural change in the interferometer was made in 1970. The photoelectric microscope was originally attached to the side of the waybed at the midpoint, and the beamsplitter assembly was attached to the waybed at one end (see figure 2). A steel beam (not shown in figure 2) extending from the microscope to the top of the beamsplitter assembly was installed to connect these two components together very early when it became apparent that the microscope angle changed when the carriage moved on the ways. The

beam reduced this mechanical distortion, but a more effective structure was subsequently designed and installed.

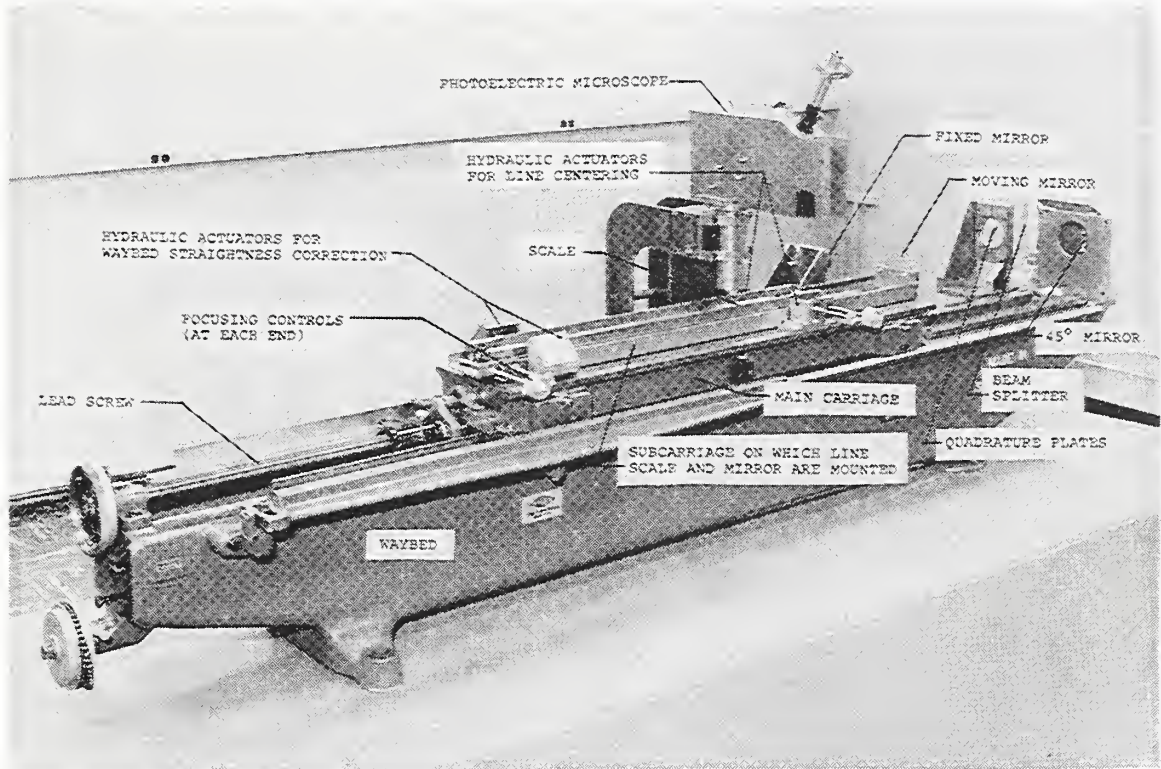


Figure 2. The line scale interferometer mechanical and optical components prior to 1970

In this new structure the microscope and beamsplitter were mounted at opposite ends of a rigid steel tube and the assembly kinematically mounted on the waybed. The beamsplitter end was fixed to the waybed and the microscope end was supported on two 25.4 mm ball bearings free to roll in a direction parallel to the ways (see figure 3). This decoupled the assembly from waybed distortions. By 1971 the change was completed and group M3 shows the resulting measurements. There was an improvement in precision and a change in apparent length.

By 1978 the original fringe counting and microscope electronics were failing from age. In 1979 the electronics were replaced with modern components and the original NBS-made interferometer was replaced with a commercial model. All this was done while preserving the principles of the original NBS design[3]. Commercial laser interferometers were by then quite reliable so a Hewlett-Packard² (HP) model was installed. Aligning the HP

² Trade names and company products are mentioned in the text to specify adequately the equipment and procedures used. Such identification does not imply recommendation or endorsement by the National Institute of Standards and Technology, nor does it imply that the products are necessarily the best available for the

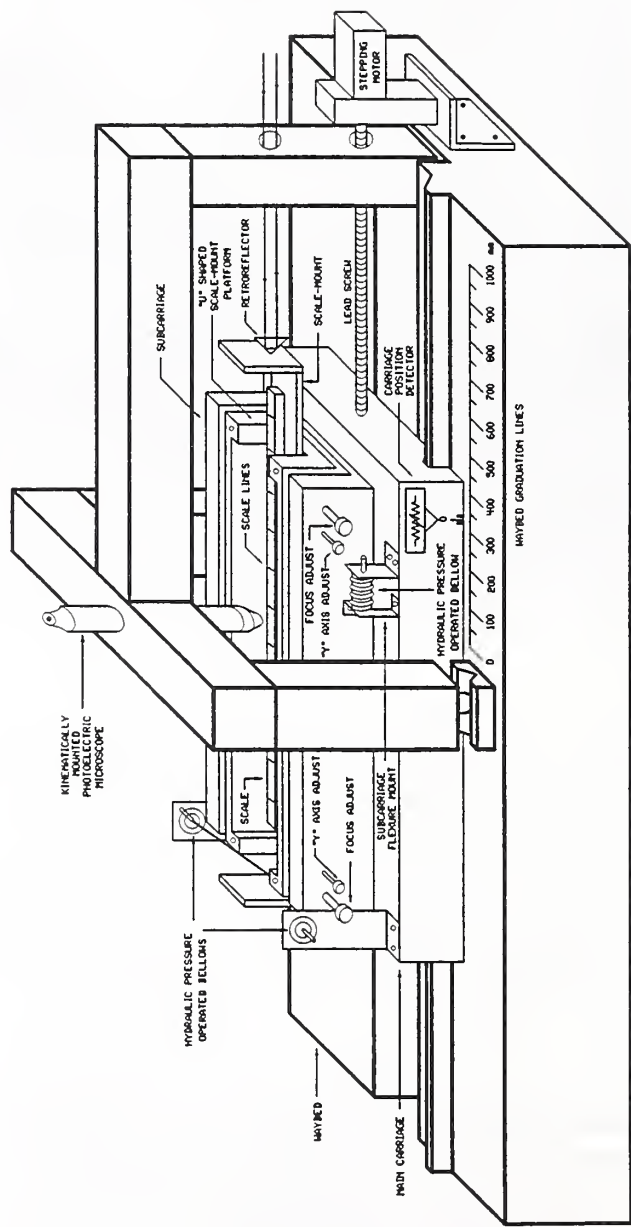
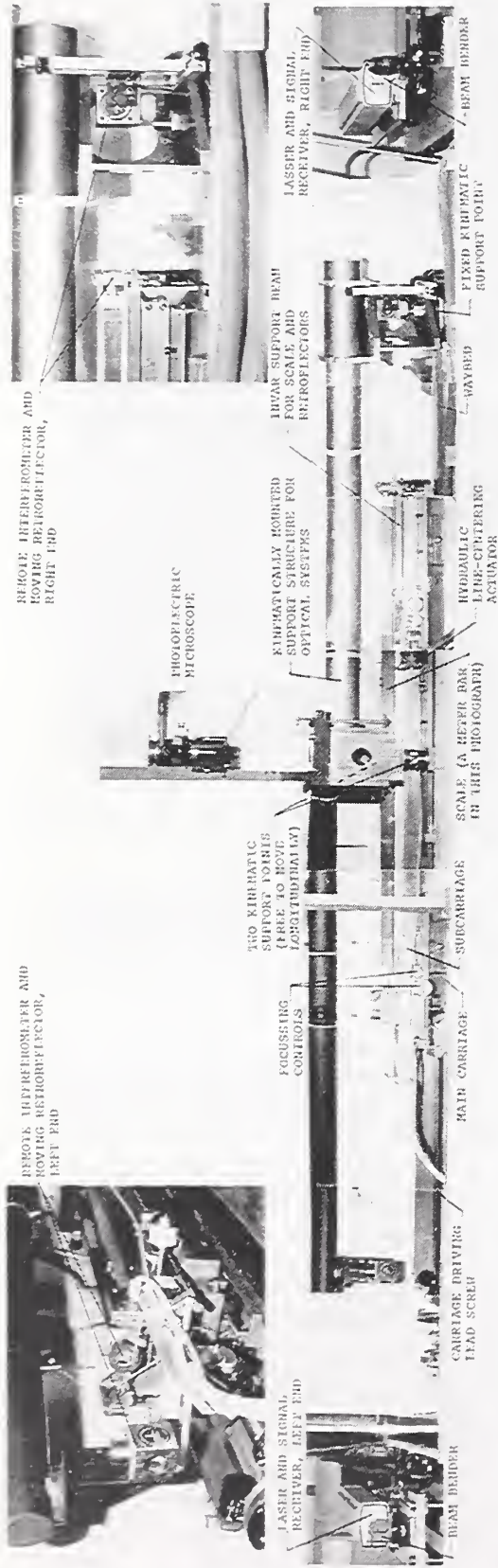


Figure 3. A photograph and a schematic diagram of the line scale interferometer in its present form

interferometer was easier and more precise than with the Michelson, and the resulting improved measurement precision can be seen in group M4.

In the 1986 evaluation[1] of M1 through M4 data the possibility of real secular lengthening of M5727 as evidenced by the slope of a fitted line was considered but rejected for several reasons. Most important was that the first interferometric measurement group, M2, was the least reliable, having been taken before the fixed optical components of the interferometer and the microscope were rigidly coupled and kinematically mounted to prevent mechanical distortion that could cause measurement errors. The mean value change from M2 to M3 was, therefore, discounted as evidence for dimensional change in M5727. Data on the 20 inch (508 mm) control bar was of little use because it had not been measured over a long enough time period. The change in No. 5727 from M3 to M4 was attributed, as justified by the available evidence, to measurement process change M4 (replacement of the Michelson type with an HP interferometer). For more information see section 5.4.2 of reference [1].

3.4 Control data from 1971 to 1991

3.4.1 Meter bar M5727

Figure 4 shows the M5727 measurement history to 1991 with M1 and M2 data deleted to eliminate the influence of the probably biased early data. Vertical lines correspond to process changes M5 through M7 in table 1 (M8 data will be shown and discussed later). Horizontal lines are period mean values. The additional data strengthens the case for a steady lengthening of this bar.

3.4.2 20 inch (508 mm) scale No.6495

Figure 5 is the history of steel control bar No. 6495 plotted on the same time scale as meter bar M5727 in figure 4. During the time common to both scales, the length of the former is increasing and the latter is decreasing.

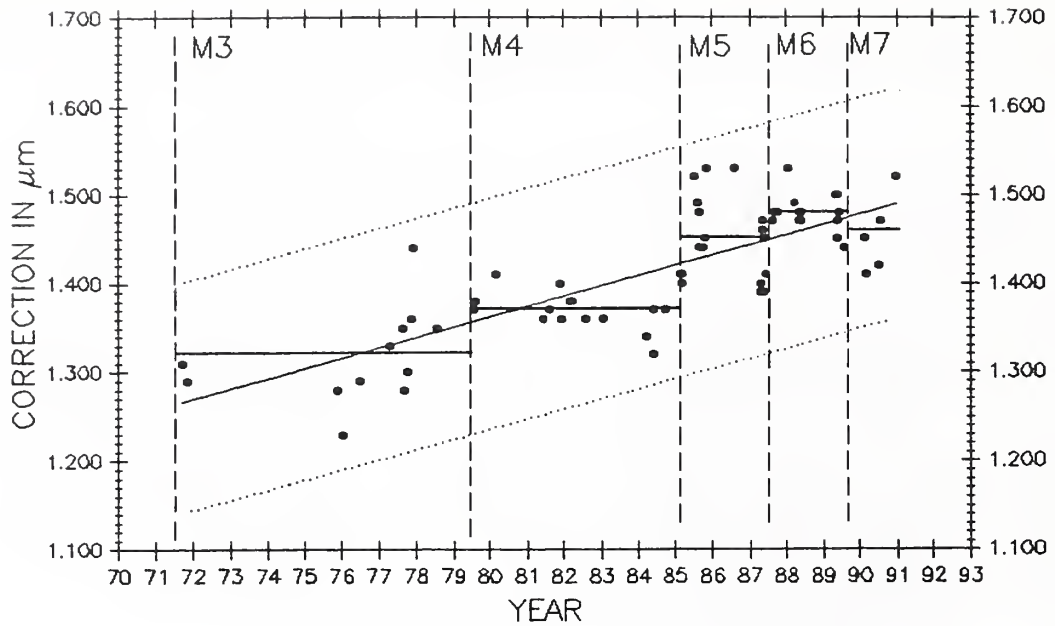
In 1987, a re-evaluation of measurement process errors was initiated to ensure that a valid case existed for real change in M5727. There was insufficient data on No. 6495 at that time to show a trend. More importantly, the error study was initiated to reduce measurement uncertainty.

4. RE-EVALUATION OF MEASUREMENT ERRORS

4.1 The need for re-evaluation

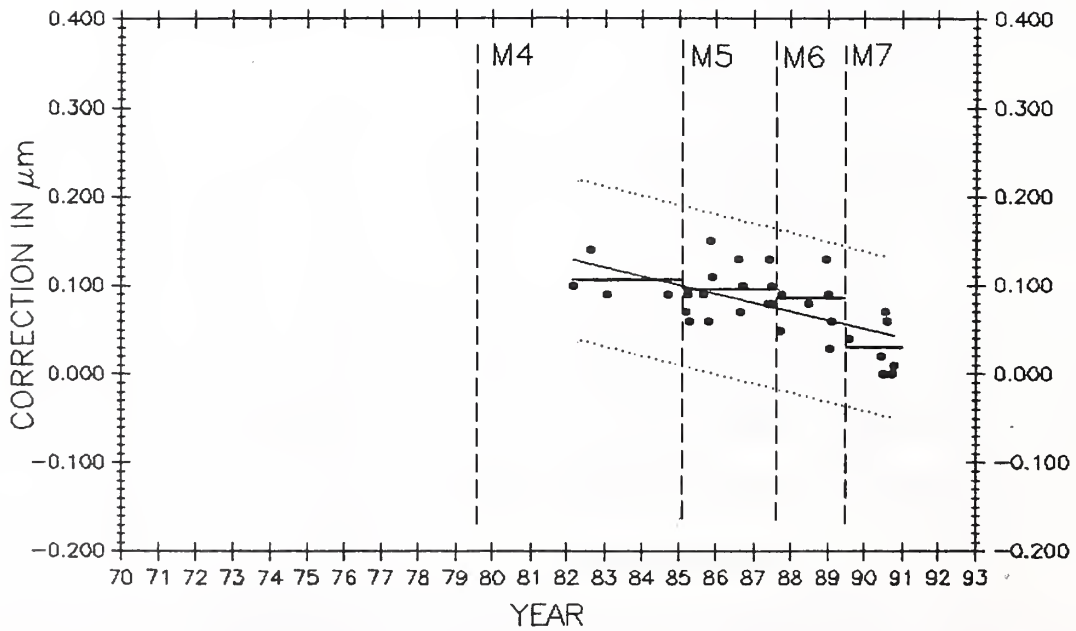
While control standards are indispensable tools for monitoring a measurement process they may have the disadvantage of undergoing slow secular length change, leaving some doubt about process

purpose.



CONTROL BAR M5727, 0 TO 1 METER

Figure 4. M5727 control chart, 0 to 1 meter, 1971 to 1991



CONTROL BAR NO. 6495, 0 TO 20 INCHES (.508 METER)

Figure 5. 20 inch bar No. 6495, 0 to 20 inches, 1982 to 1991, plotted on same time scale as figure 4

performance. Consequently, three goals were set for the error study: (1) reduce measurement errors, (2) determine control bar stability, and (3) establish a new value for measurement uncertainty.

4.2 Systematic and random errors

The following is a list of potential length scale measurement process systematic error sources:

- Wavelength
 1. Vacuum wavelength of the laser
 2. Refractive index of air determination
 - a. Refractive index equation
 - b. Air temperature measurement
 - c. Atmospheric pressure measurement
 - d. Humidity measurement
 - e. Air composition
 3. Zero shift (deadpath error)

- Interferometer
 1. Alignment of interferometer axis with scale graduation axis (minimum Abbé offset)
 2. Structural characteristics
 - a. Constancy of distance between reference mirror and microscope
 - b. Constancy of distance between beam splitter and microscope
 - c. Constancy of distance between measuring retro-reflector and scale

- Scale
 1. Temperature measurement
 2. Thermal expansion coefficient
 3. Graduation quality

All of the listed parameters are also sources of long term random error with the exception of the air refractivity equation, and scale thermal expansion coefficient. In a single measurement they all remain sources of systematic error, but here we are studying only the long term. Long term randomness is exemplified by interferometer structural characteristics where these critical distances can be changing in one direction during one measurement and in the opposite direction in another, or their magnitude can change from one measurement to the next.

With the exception of interferometer structural characteristics, scale graduation quality, and zero shift correction these parameters cause length-dependent errors. Their relative effects are illustrated in Table 3 where it is shown, for example, that a 0.1 μm length measurement error will result from an error of 0.009 $^{\circ}\text{C}$ in temperature measurement of a one meter steel scale.

Table 3
 Relationship Between Length-Dependent Errors and Errors
 in Measurement Parameters
 and
 Estimated Process Systematic Error (S.E.)

Parameter	Error in a One-Meter Length		Process S.E.	
	1 part in 10^7 or 0.1 $\mu\text{m}/\text{m}$	1 part in 10^8 or 0.01 $\mu\text{m}/\text{m}$	parts in 10^8 (1987) (1992)	
<u>Wavelength</u>				
Vacuum wavelength	1 part in 10^7	1 part in 10^8	5	2
Refract. index eq.	1 part in 10^7	1 part in 10^8	5	3
Air temp.	0.1 $^{\circ}\text{C}$	0.01 $^{\circ}\text{C}$	0	0
Pressure	0.3 mm	0.03 mm	4	2
Rel. humidity	12% rh	1.2% rh	3	1
CO ₂ content	-----	67 parts in 10^6	2	1
<u>Interferometer</u>				
Alignment	0.45 mm/m	0.14 mm/m	2	2
<u>Scale</u>				
Steel temp.	0.009 $^{\circ}\text{C}$	0.001 $^{\circ}\text{C}$	2	2
Glass temp.	0.012 $^{\circ}\text{C}$	0.001 $^{\circ}\text{C}$	---	---
Invar temp.	0.067 $^{\circ}\text{C}$	0.007 $^{\circ}\text{C}$	Sum = 23	13
Quartz temp.	0.250 $^{\circ}\text{C}$	0.025 $^{\circ}\text{C}$		

There are two ways of estimating the systematic component of process uncertainty. First, there is the error budget method in this table. Process S.E. is shown in the two right hand columns of this table for the period before (in 1987) and after (in 1992) this study. The sum of the individual S.E.s is the worst possible case where they all have the same sign, a highly unlikely event. An accepted method for calculating a realistic process S.E. is to add the individual values in quadrature, i.e., take the square root of the sum of their squares. These estimated process S.E. values are shown at the bottom of the two columns.

The second method is to compare measurements of the same scale performed by different but equally valid measurement methods. In section 4.4.7 results of international measurements of a meter bar will be analyzed to arrive at an estimate of systematic error.

4.3 Procedure

Actions taken for this study (but not necessarily in the order given) were to:

- a. Recalibrate the temperature measurement system
- b. Recalibrate the barometer
- c. Recalibrate the hygrometer
- d. Apply zero shift (deadpath) corrections to interferometric data.
- e. Redetermine the vacuum wavelength of the laser

- f. Measure the ambient carbon dioxide content of air in the length scale laboratory.
- g. Re-examine the interferometer and scale alignment procedures
- h. Carry out an international measurement interchange by obtaining, by loan from BIPM, steel meter bar No. 12924, together with its measurement history. Meter bar No. 12924 was used in an international intercomparison organized by BIPM. It was circulated among the major national measurement laboratories of the world from 1976 to the present and it is a valuable tool for evaluating measurement processes because of the quality of its measurement history. It was measured at NBS in 1977.
- i. Test the recently proposed revision of the water vapor correction factor in Edlén's air refractivity equation.
- j. Convert from the International Practical Temperature Scale of 1968 (IPTS-68) to the International Temperature Scale of 1990 (ITS-90).

4.4 Results

4.4.1 Recalibration of the temperature, pressure and humidity measurement systems in July and August 1987

The temperature measurement system was found to remain within its uncertainty of 0.002 °C. The barometer calibration changed 16 Pa since 1985 resulting in a length change of 4 parts in 10⁸. The hygrometer calibration changed 5% r.h. since November 1985 resulting in a length change of 4 parts in 10⁸. The algebraic signs of these changes tended to make the length measurement errors they created cancel (i.e., both the barometer and hygrometer were reading low before recalibration), thus reducing the net effect. Undoubtedly the contribution of these errors added to random variability and measurement uncertainty by amounts varying from zero to several parts in 10⁸.

No corrections to existing data have been made, but barometers and hygrometers with better claimed stability were obtained. More frequent checks on these instruments have been made since March 1990. Some instabilities are still being found, but humidity measurement uncertainty is now 1.2% r.h., and barometric pressure measurement uncertainty is 8 Pa (0.06 mm of Hg).

4.4.2 The International Temperature Scale of 1990 (ITS-90)

As a result of the adoption of a new international temperature scale[4] (ITS-90) to replace the International Practical Temperature Scale of 1968 (IPTS-68) the temperature measurement system was recalibrated again in September 1991. The difference between the two scales at 20 °C is

$$(ITS-90) - (IPTS-68) = -0.005 \text{ } ^\circ\text{C}$$

During the period between the official adoption of ITS-90 and this

recalibration, line scale measurements were corrected for the difference without converting to the new temperature scale.

Measurements of steel control bar No. 6495 were made before and after the change in temperature scales as a verification. Since the linear thermal expansion coefficient of this bar is $11.5 \times 10^{-6}/^{\circ}\text{C}$ the computed length change from 20°C IPTS-68 to 20°C ITS-90 is $0.005 \times 11.5 \times 0.508 = +0.03$ micrometer. The measured change was $+0.04$ micrometer with an uncertainty of 0.01 micrometer. The 0.01 micrometer difference between the measured and computed length change is equivalent to 0.002°C .

4.4.3 Interferometric zero shift corrections

Beginning in 1989, zero shift corrections[5] (often called deadpath corrections) were applied to compensate for barometric pressure changes occurring during length measurements. These pressure changes can expand or contract the standing wave train between the remote interferometer and the retroreflector and thus affect the interferometric count.

The 43 measurements of M5727 in the 1991 revised Edlén equation experiment provide data for evaluating the zero shift correction. Table 4 shows mean values and statistics for the one meter interval on M5727 with and without correction. Although the mean length changes by only 2 nanometers, precision is improved by more than a factor of two by correcting. Algebraic signs of the corrections change with direction of pressure change and with direction of interferometric counting, so over the long term, zero shift corrections tend to cancel and have minimal effect on mean length. However, making the corrections improves both short and long term precision.

Table 4.
Effect of making zero shift corrections

Mode	Deviation from nominal length (μm)	Precision 3σ (μm)
Without correction	1.420	0.079
With correction	1.418	0.037

4.4.4 Laser wavelength

A measurement of the laser vacuum wavelength revealed a shortening of 7 parts in 10^8 . For lack of a better model this change was prorated over the data on M5727 and No. 6495 for the period from 1979 to 1989. Laser wavelength measurements are now more readily available than they were in the past.

4.4.5 Carbon dioxide content of laboratory air

Carbon dioxide levels were measured in the NIST length scale laboratory in October of 1990. The ambient level with no one in the room averaged 350 ppm. With one person in the room the average increased to 375 ppm, and with two people it increased to 400 ppm. During length scale measurements there is occasionally more than one person in the laboratory so 380 ppm (1.2 people) is a reasonable average value for control bar measurements. In the Edlén equation[6] for the refractive index of air a value of 300 ppm is assumed. Figure 6 shows the relationship between CO₂ content of air and the 0.6328 μ m laser wavelength applying Jones's[7] analysis of the effect. Changing the CO₂ value from 300 to 380 ppm will change measurements of a one meter length by -0.012 μ m (-1.2 parts in 10⁸).

Figure 7 shows the trend in ambient atmospheric CO₂ levels according to historical records[8]. From 1958 to 1988 the CO₂ level increased by approximately 40 ppm. While values may change with the season and from one geographic location to another the chart is representative of a worldwide trend. The length change is small but all values on the control standards have been adjusted by assuming an average laboratory level 30 ppm above average ambient and pro-rating the change over the 26 years of interferometric measurements.

4.4.6 Interferometer and scale axis alignment

A more precise and less tedious method for aligning the scale graduation axis with the interferometer optical axis was devised in August 1987.

A short 3 mm diameter peg in the back of the retroreflector case, located on the axis of the retroreflector, was lengthened a few millimeters so that it could extend under the microscope objective lens when the carriage was moved all the way to the right. Milling a flat bottomed notch halfway through the extended peg created a surface coincident with the retroreflector apex and scribing an axial line at the center of this surface created a coincident reference line.

To make use of the reference line, the scale to be measured is first focused and aligned in the center of the microscope field. Then the carriage is moved to bring the reference mark under the microscope and the retroreflector is adjusted vertically and laterally to bring the reference mark into focus in the center of the microscope field. The interferometer axis is then coincident with the scale axis by virtue of both being in the focal plane of the microscope and aligned laterally in the microscope. This procedure ensures the accuracy of this critical adjustment.

4.4.7 BIPM meter bar No. 12924; international measurements

The second method for estimating process systematic error mentioned in section 4.1 is to compare measurements of a scale as performed

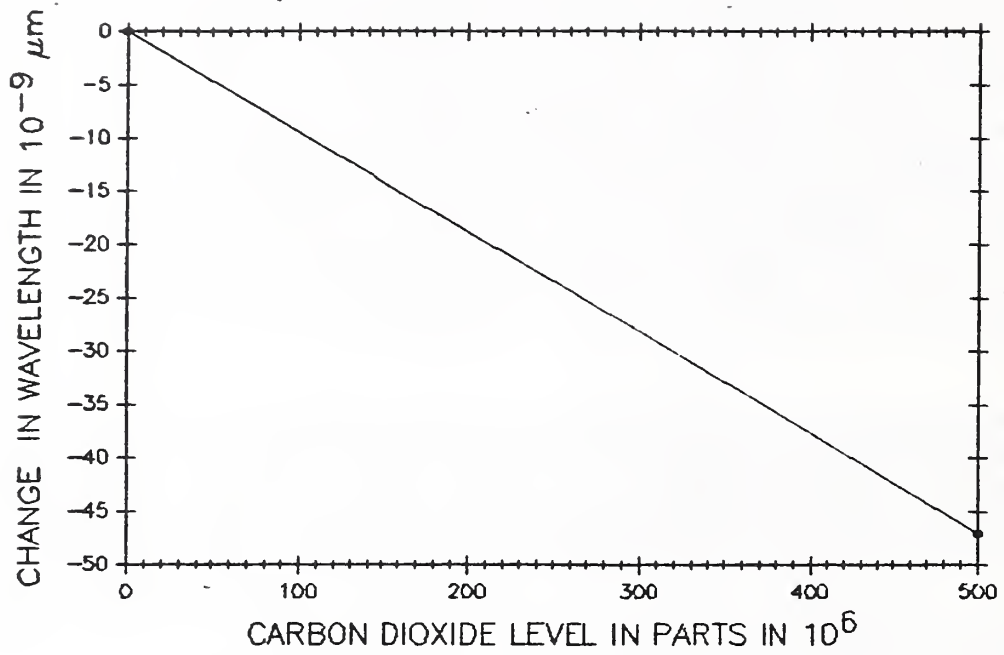


Figure 6. Effect of CO_2 content of air on the laser wavelength

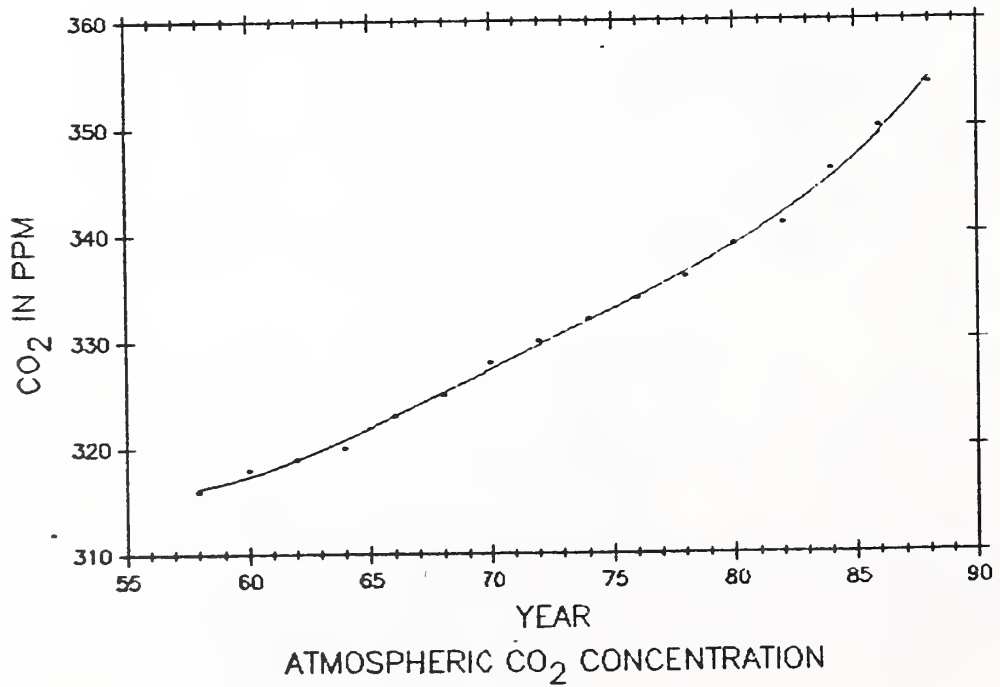


Figure 7. Atmospheric CO_2 levels, 1958-88

by several independent and valid methods. This was done through the good offices and efforts of the Bureau International des Poids et Mesures (BIPM). Starting in 1976 steel meter bar No. 12924 was sent to the national standards laboratories of most of the world's industrialized nations in a very successful and useful international standardization effort.

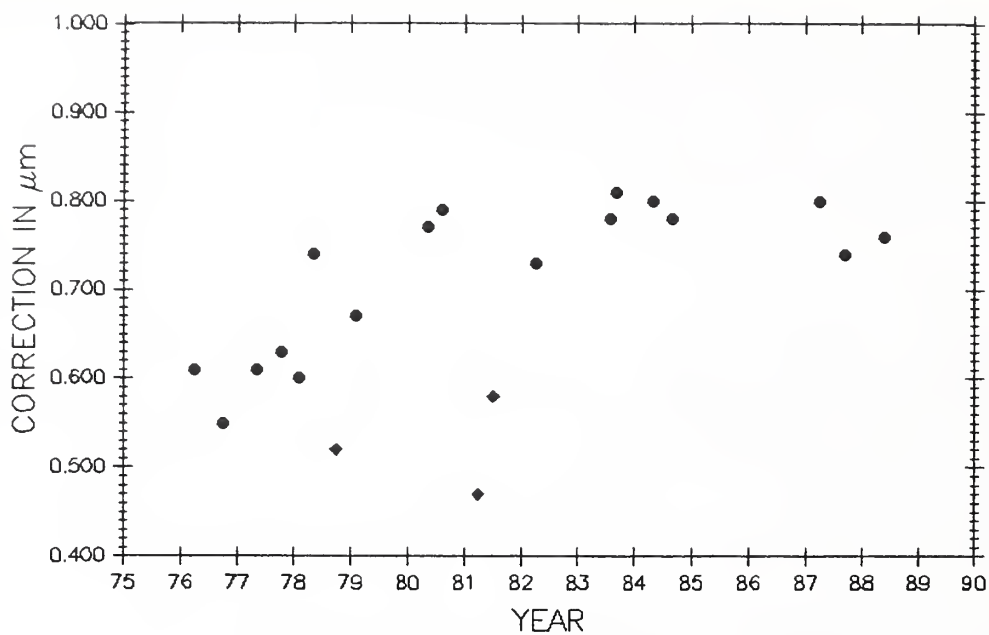
The following laboratories measured this bar:

1. Bureau International des Poids et Mesures (BIPM), Sevres, France
2. National Measurement Laboratory (NML), later CSIRO, Lindfield, Australia
3. National Research Council (NRC), Ottawa, Canada
4. National Bureau of Standards (NBS), later NIST, Washington, DC, USA
5. National Research Laboratory of Metrology (NRLM), Ibaraki, Japan
6. National Physical Laboratory (NPL), Teddington, UK
7. Amt fur Standardisierung, Messwesen und Warenprufung (ASMW), Berlin, East Germany
8. D.I. Mendeleef Institute of Metrology (MIM), Leningrad, USSR
9. Physikalisch-Technische Bundesanstalt (PTB), Braunschweig, West Germany
10. Instituto di Metrologia G. Colonnetti (IMGC), Turino, Italy
11. National Institute of Metrology (NIM), Beijing, PRC
12. Federal Office of Metrology (OFMET), Wabern, Switzerland

BIPM has measured the bar 6 times, NBS/NIST 3 times, CSIRO and MIM twice each, and the remaining laboratories once each for a total of 21 measurements over a period of 12 years. The values for the 1 meter length are shown graphically in figure 8.

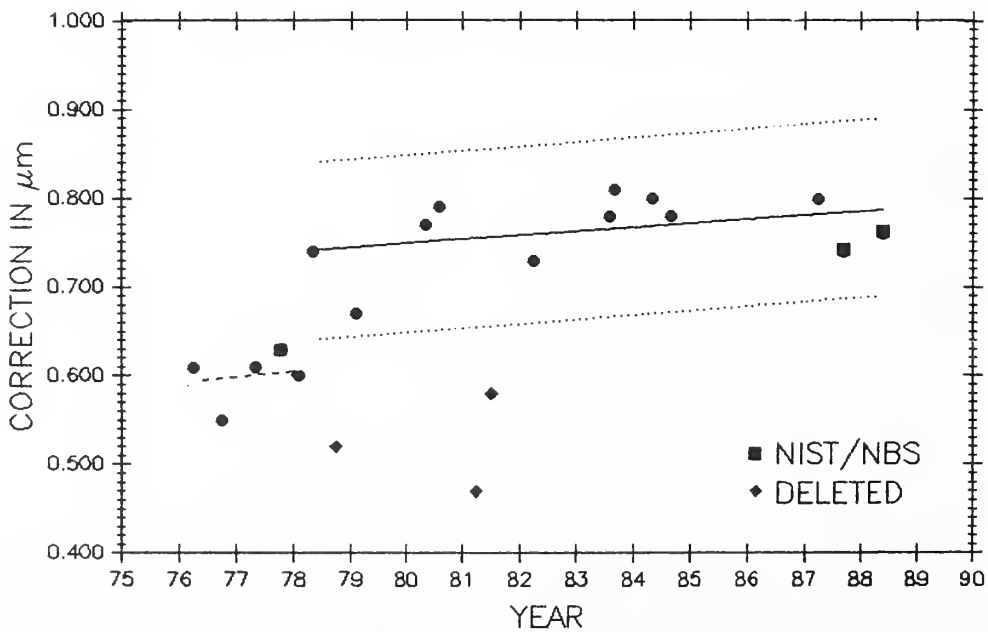
In evaluating these data, BIPM[9] concluded from its own and other measurements that the bar experienced a sudden lengthening of over 0.1 μm early in 1978. Such a change could have been caused by a severe mechanical shock, probably during shipment. Based on this conclusion the data were divided into a group before the change and a group after the change. BIPM further concluded that the bar has a long term linear growth trend. This is shown in figure 9 where linear fits are made to the two groups of points. The first group lacks sufficient data points and time span to establish an accurate slope value so it has been given the same slope as the second group. Three points, indicated on the graph by diamonds, were deleted from the analysis. When they were included in the data the 3 sigma limit becomes much larger and they could not be rejected by the accepted statistical test of exceeding the 3 sigma limit. In this case they were rejected because they did not fit the pattern convincingly established by the other 12 points in the second group.

NBS/NIST measurements are indicated in figure 9 by squares and the



BIPM METER NO. 12924, 0 TO 1 METER

Figure 8. International measurement data on BIPM Meter No. 12924, 0 to 1 meter



BIPM METER NO. 12924, 0 TO 1 METER

Figure 9. Lines fitted to the data on meter No.12924

1987 and 1988 points agree with the international average, as represented by the line fitted to all the accepted points, within $0.04 \mu\text{m}$ and $0.03 \mu\text{m}$ (4 and 3 parts in 10^8) respectively. Control bar M5727 was measured a number of times during the periods No. 12924 was being measured, so the validity of these M5727 measurements is greatly enhanced, the stability of the NIST measurement process is verified, and the growth of M5727 is proven. Table 5 shows these values.

Table 5. Measurements of control bar M5727 made during same period as measurements of meter No. 12924

Decimalized Date (in years from 1900)	M5727 Correction to 1 M Interval (in μm at 20°C.)
77.73	1.34
87.67	1.47
88.30	1.48

4.4.8 The revised water vapor correction in the Edlén equation

In 1988, Birch and Downs[10] published new values of the refractive indices of dry and moist standard air for the $0.633 \mu\text{m}$ wavelength of He-Ne lasers. The result was a revised water vapor correction in the Edlén equation for the refractive index of air. Indirect testing of the revised correction was done at NIST by measuring control standard M5727 with the line scale interferometer over a range of air moisture. These measurements were made as part of this study of systematic errors but the results were published as a separate paper[11].

Results of testing the revised water vapor correction are shown in figure 10. Measured length values of M5727, computed two different ways, are plotted against partial pressure of water vapor. The upper plot (A) is the data computed with the 1966 Edlén equation and the lower plot (B) is the same data computed with the revised Edlén equation. The reduction in the correlation between length and water vapor content in the lower plot proves that the revised correction provides a much better estimate of the refractive index of moist air.

In addition to verifying the revised water vapor correction two things are demonstrated and discussed in reference 11: (1) accurate and reliable hygrometers are essential to making high accuracy length measurements, and (2) air moisture content must be measured inside the interferometer housing. Using the revised equation improved measurement precision (3σ) from $0.15 \mu\text{m}$ for the period from 1971 to the end of 1991 to $0.04 \mu\text{m}$ for the seven month period shown in figure 11.

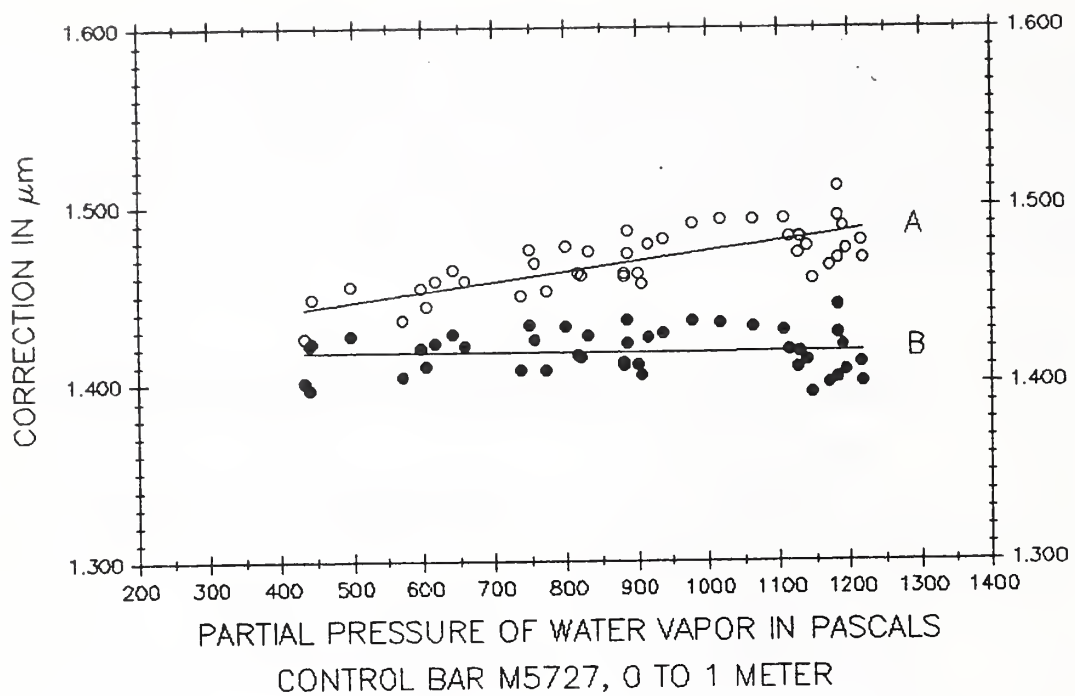


Figure 10. M5727 length vs water vapor content of air

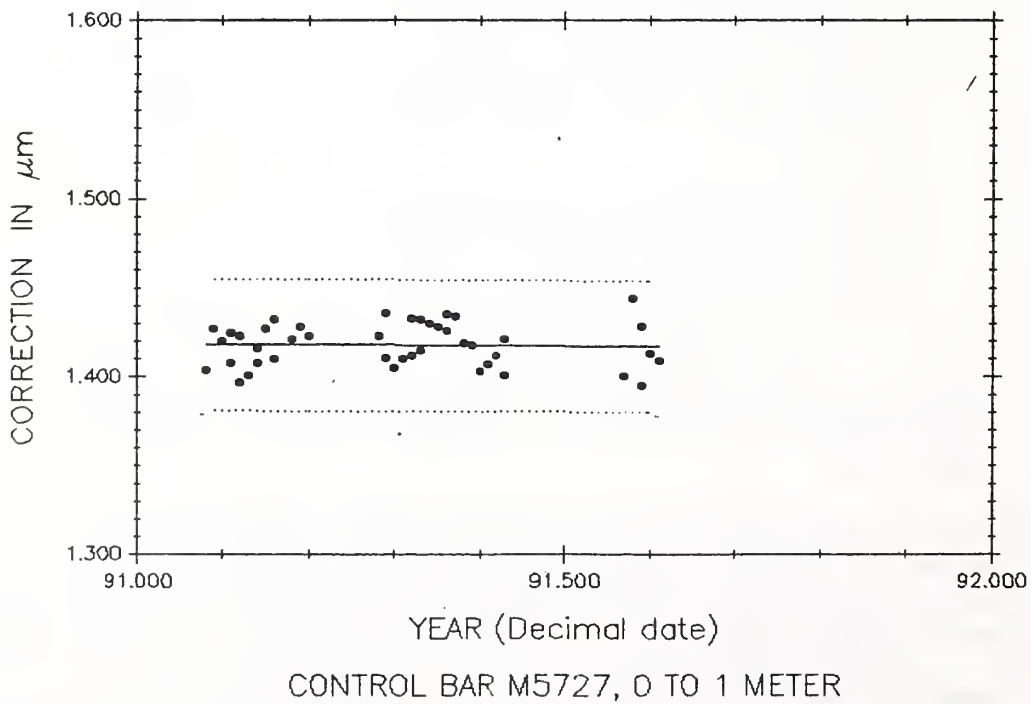


Figure 11. M5727 length vs measurement date in the water vapor experiment

All control data has been retroactively adjusted for the change in the water vapor correction, but errors still exist through period M7 in measuring water vapor because the hygrometer was outside the interferometer housing and, in some cases, the hygrometer was out of calibration. Precision was improved from $0.15 \mu\text{m}$ to the $0.12 \mu\text{m}$ value shown in figure 4 by this adjustment.

4.4.9 Correlation test: length vs barometric pressure

A correlation between control standard length and atmospheric pressure could indicate further problems with the wavelength correction equation. Figure 12 is the data from the air moisture experiment (plot B, figure 10) with corrections to the one meter length of M5727 plotted against observed barometric pressure. No statistically significant correlation is indicated.

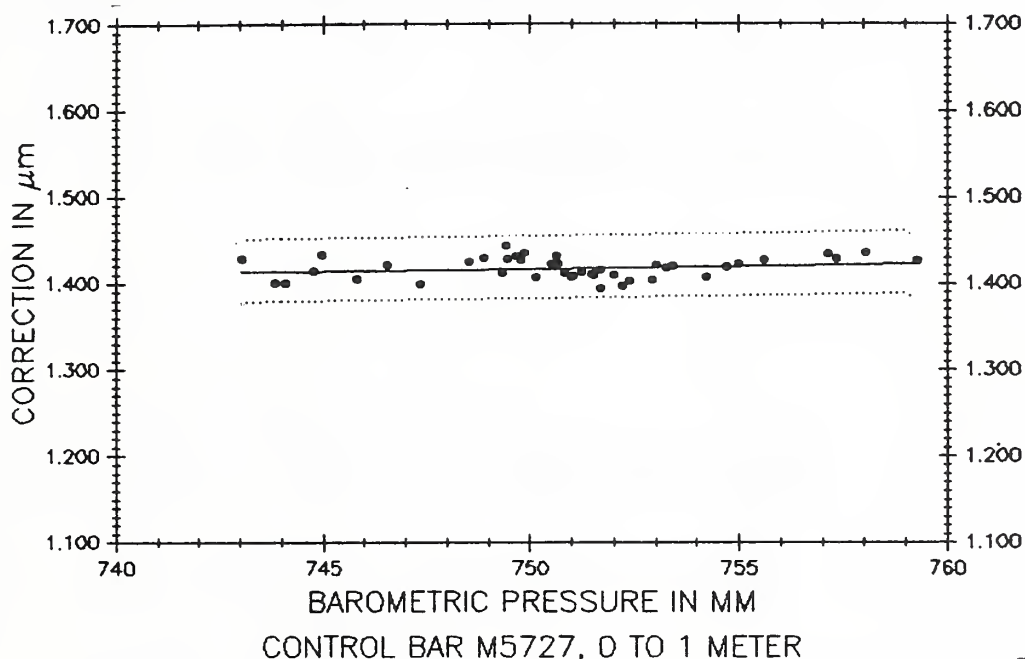


Figure 12. M5727 length vs barometric pressure correlation test

4.5.0 Summary of data adjustments

All measurement data on both controls have been adjusted retroactively for the following changes:

1. The change in international temperature scales from IPTS-68 to ITS-90 (see 4.4.2).
2. The laser vacuum wavelength change between 1979 and 1989 (see 4.4.4)
3. The carbon dioxide content of laboratory air from 1966 to date (see 4.4.5).
4. The revised water vapor correction to the laser wavelength (see 4.4.8).

Measurement data for the controls were adjusted for interferometric zero shift starting in 1989 (see 4.4.3), a procedure that is now

standard practice.

5.0 INTERPRETING CONTROL CHARTS

5.1 Secular change, 1971 to 1991

Long term length changes have occurred in both control standards. If all 68 data points on M5727 from 1971 to 1991 are used in a linear regression the following equation results:

$$Y = 0.0115X + 0.438 \quad (1)$$

where Y is the correction axis, X is the time axis and the Y intercept is at the year 1900.

Using only the nine data points taken concurrently with measurements of BIPM meter No. 12924 the equation is:

$$Y = 0.0133X + 0.311 \quad (2)$$

The equation from data taken concurrently with meter No. 12924, and all data taken between them (total = 49 points) is:

$$Y = 0.0127X + 0.353 \quad (3)$$

The mean of equations (1), (2), and (3) is:

$$Y = 0.0125X + 0.367 \quad (4)$$

Table 6 shows a comparison of the four equations. Over the 20 year period the first three equations, each based on a different data subset, differ by only 0.037 μm .

Table 6. Comparison of the linear regression equations for M5727

Equation	Slope ($\mu\text{m}/\text{m}/\text{yr}$)	Computed Y values	
		1971 (μm)	1991 (μm)
(1)	0.0115	1.254	1.484
(2)	0.0133	1.255	1.521
(3)	0.0127	1.255	1.509
(4) Mean	0.0125	1.255	1.505

Equation (4), an unweighted mean, is a compromise. The equations could be weighted according to the number of observations, but that would have denied the importance of equation (2) where the observations are few but internationally verified. In this case any weighting scheme is difficult to justify, and in any event will make little real difference in the result.

The 20 inch control No. 6495 (see figure 5) shows negative growth and a linear fit to the data is

$$Y = -0.00978X + 0.932 \quad (5)$$

It is interesting that one control is growing and the other is shrinking with time. The processes that cause secular volumetric changes in steel are obviously different from those in Invar.

5.2 Control charts incorporating 1991-92 data

Figures 13 and 14 are the revised control charts on M5727 and No. 6495 with 1991 (M8) data included. The mean values for periods M3 through M8 are shown as horizontal lines in addition to the line fitted to all the plotted data.

Of the 111 data points available on M5727, 43 are in M8 (nearly 40%). Plotting all 43 points in M8 will give excessive weight to this group. Since 1976 there has been an average of 5 data points per year. On that basis 5 points in the M8 period would be a reasonable weighting of the data. The 5 points selected give the same mean value and approximately the same spread as the 43 points.

A linear fit to all the plotted M5727 data is

$$Y = 0.01024X + 0.544 \quad (6)$$

No. 6495 has 8 points in M8 and needs no weighting. The fit to the No. 6495 data is

$$Y = -0.00598X + 0.605 \quad (7)$$

Control bar M5727 seems to show the reversal from positive to negative growth starting in period M7. Control bar No. 6495, on the other hand, seems to show a reversal from negative to positive growth in period M8. It is premature to conclude that a reversal has actually occurred. More data and more time is needed to determine the changes in trends, if any, and equations (6) and (7) are considered as interim.

In table 7 measurement precision (3σ) for both controls is shown for each period.

Table 7. Measurement precisions for each time period

Period Symbol	M5727 0 to 1 m (μm)	No. 6495 0 to 0.508 m (μm)
M1	0.24	---
M2	0.23	---
M3	0.16	---
M4	0.07	0.07
M5	0.14	0.08
M6	0.06	0.10
M7	0.10	0.09
M8	0.04	0.04

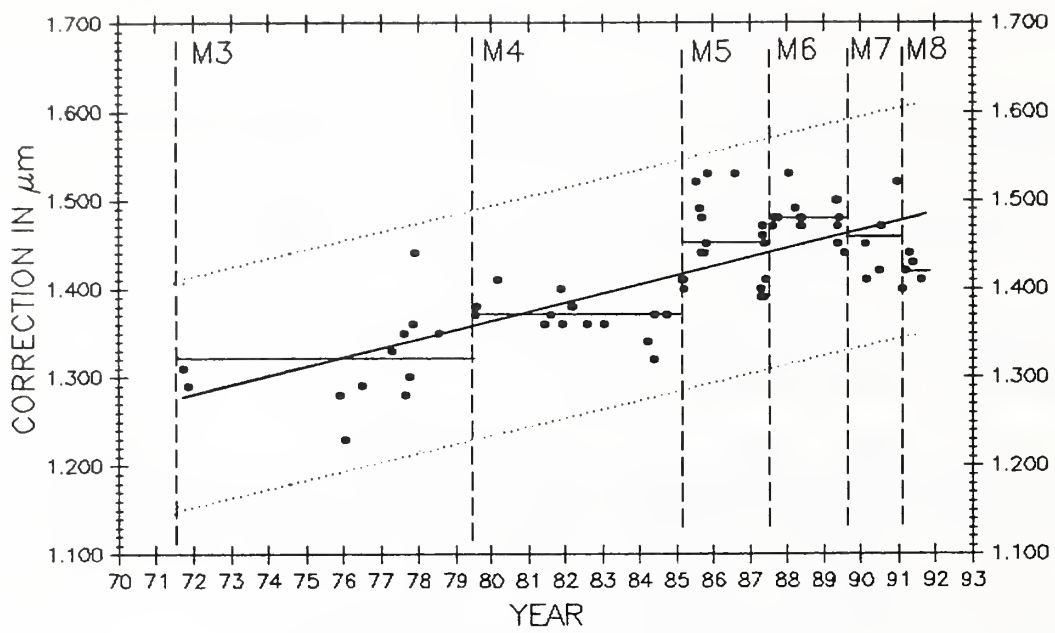


Figure 13. M5727 control chart, 1971-92

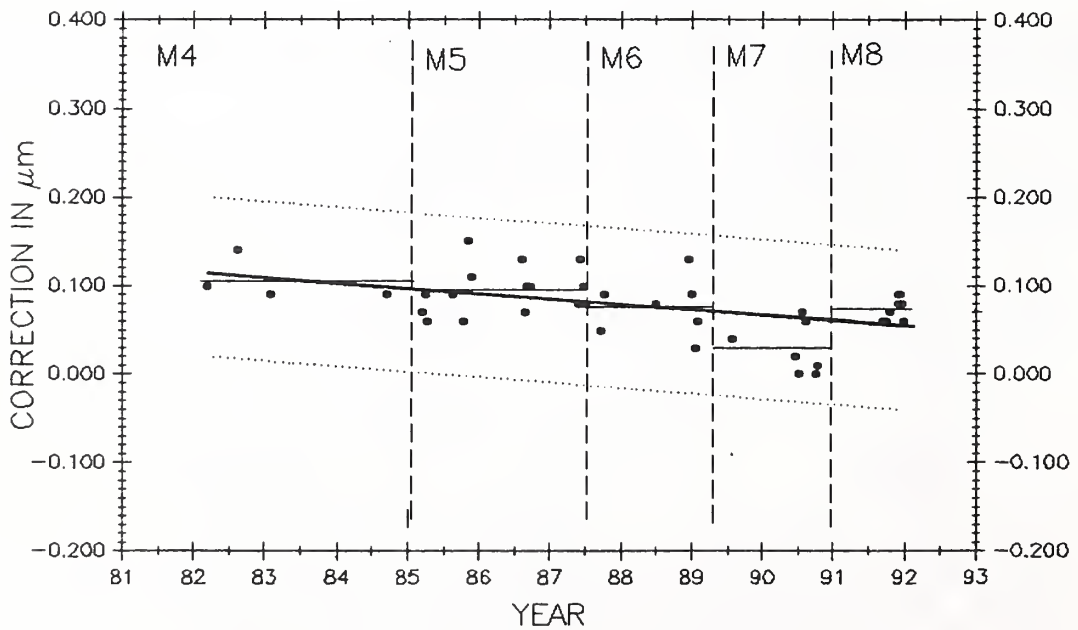


Figure 14. Bar No. 6495 control chart, 1982-92

After the major process changes were made (kinematic mounting of the optical components at M3, and changing to a commercial interferometer at M4) the variations in precision in M5, M6 and M7 were probably caused not only by the planned process changes but by factors such as undetected changes in the calibration of barometers and hygrometers, variations in interferometer alignment, and errors in air moisture measurements caused by the hygrometer being outside the interferometer housing.

5.3 Simulating stability in control standards

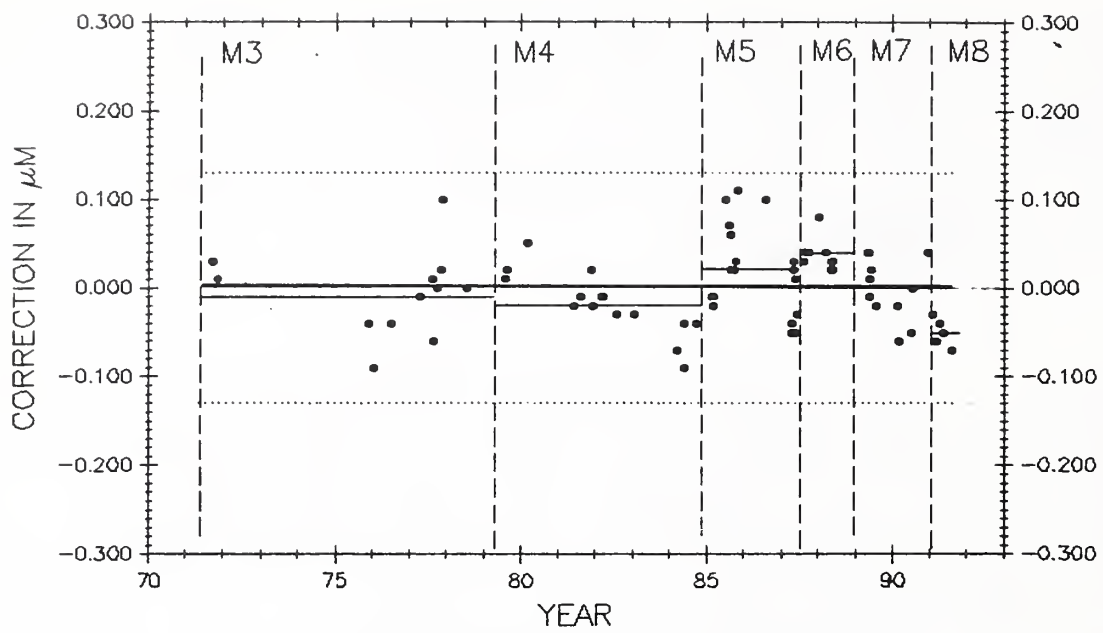
Figure 15 is the 1971 through 1991 data on M5727 corrected for secular length change and showing the mean for each period (individual horizontal solid lines) in addition to the horizontal line fitted to all the data. This chart is produced by plotting the residuals, i.e. observed minus computed values, based on the equation for the fitted line in figure 13. If the control standard had remained constant in length throughout this time period its measurement history would look like figure 15. Simulated stability makes evaluating process changes easier by releasing group mean values from the data slope influence (compare with figure 13). In this evaluation none of the process changes make a statistically significant change in the mean length. This demonstrates that the measurement process is consistent throughout the 20 years since 1971 despite changes made in the process.

Figure 16 is similar to 15 except that all the data are considered to be a single group. The basic assumption in most statistical techniques is that the data are a random sample from a stable probability distribution and, in most cases a normal frequency distribution is formed. Stability in this case is simulated by the secular change adjustment but it is none the less valid.

The histogram in figure 17 is a test of how closely the data approximates a normal distribution. It peaks near the zero value as it should, but its lack of symmetry probably indicates insufficient data (there are 120 points) rather than a serious bias.

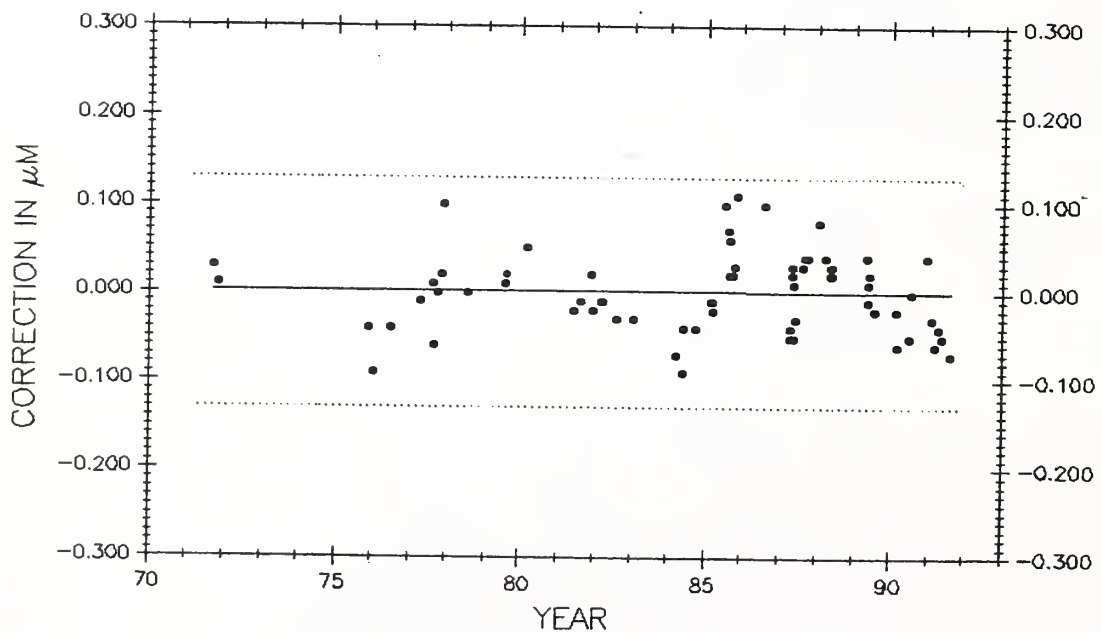
5.4 The measurement process in an out-of-control state

The 3σ limits are a predictor of future measurement performance based on past performance. There is a 99.7% chance that the next measurement value will fall within the 3σ bounds if the process is in statistical control. Values that fall outside the bounds indicate that the measurement process is out of statistical control. If the cause of the out-of-control measurement is found, or if the next measurement is back in control, the errant measurement value can be deleted. If out-of-control measurements persist, the cause must be found and corrected before the errant measurements can be deleted. During periods when the process is out of control the uncertainty of measurements made with the process cannot be stated at all or must be increased. Statistical bounds are periodically recomputed, as data accumulates, to reflect changes in measurement process precision. Frequent measurement of



CONTROL BAR M5727, 0 TO 1 METER

Figure 15. Simulated length stability in M5727



CONTROL BAR M5727, 0 TO 1 METER

Figure 16. Simulated length stability chart with data treated as one group

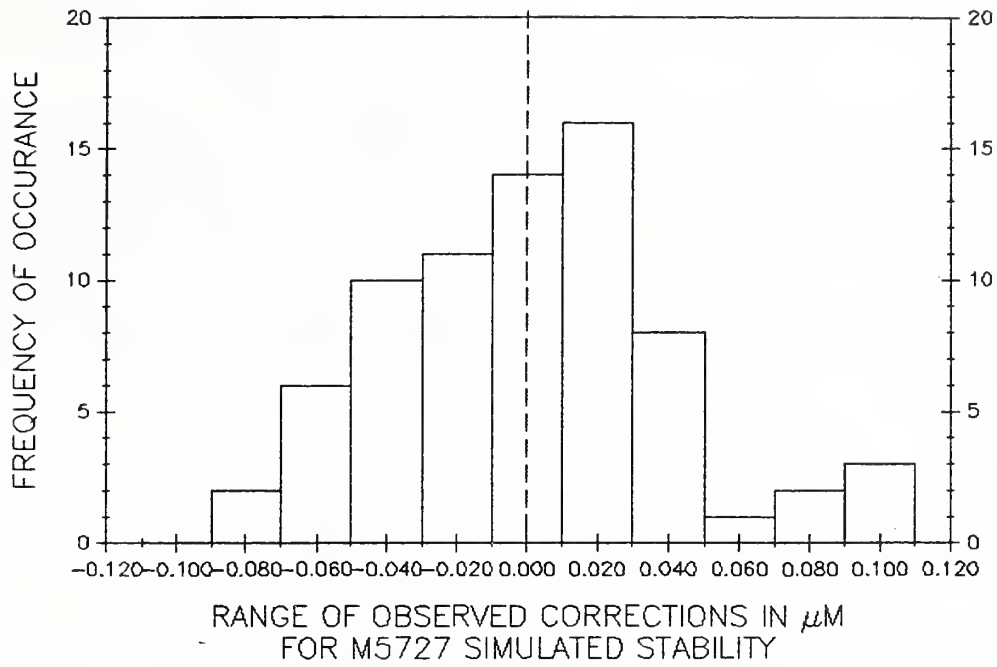


Figure 17. Histogram of simulated stability data

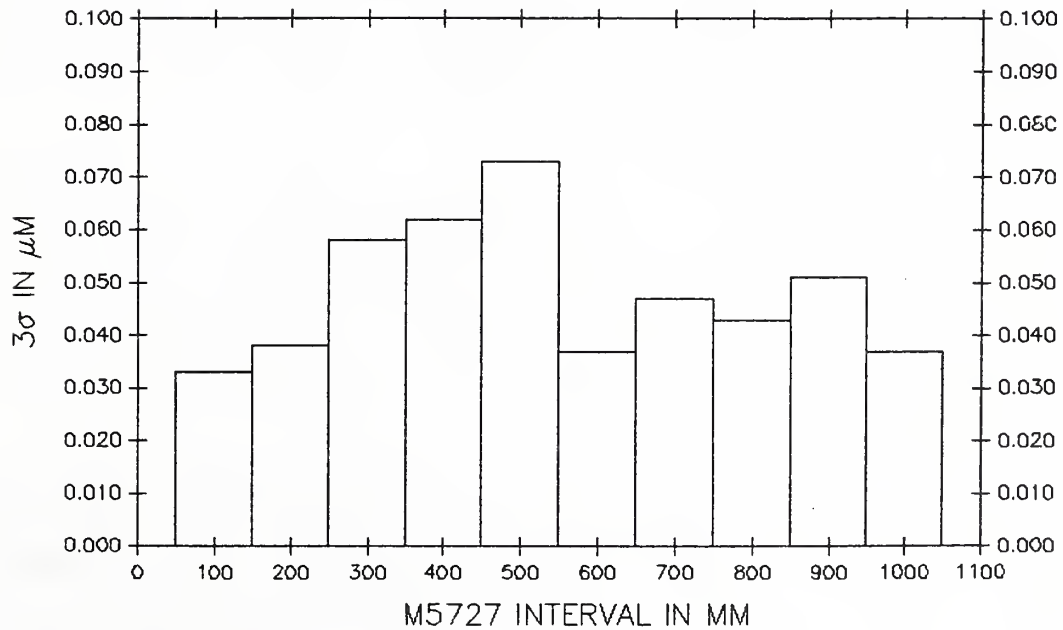


Figure 18. M5727 measurement precision for decimeter subintervals

the control standards will detect problems when they arise and ensure that the process is performing as predicted.

For a period in 1986, the process did go out of control. A series of erratic values on M5727, many of which fell outside the existing bounds, had a 3σ value in excess of 0.4 μm . A search was made for the cause and each time a remedy was tried more measurements of M5727 were made. After a number of attempts, it was found that the retroreflector on the carriage had loosened and was shifting slightly during measurements. Once the mount was tightened the process returned to a state of control and the out-of-control data were deleted.

5.5 Control subintervals

Figure 18 shows the variation in measurement precision among the M5727 decimeter subintervals in the water vapor compensation experiment (plot B in figure 10).

In the setup procedure for any scale measurement the terminal graduations are brought into sharp focus, so there may be degradation of focus on subinterval graduations if the scale surface is not flat. Because of production difficulties, long scales are more likely than short ones to be out-of-flat. Any scale deemed likely to be distorted by gravity is supported at its Airy points to minimize bending and to provide a reproducible support method. A number of effects are operating to create the variations in subinterval precision seen in figure 19. In addition to gravitational sag there may be slight variations in graduation symmetry and blemishes on or near the graduations that can influence the centering operation of the photoelectric microscope.

5.6 Long term measurement precision

The somewhat structured pattern in measurement results from 1971 to 1991 is cause for concern and is exemplified in figure 16 by data trends such as the downward slope of points from 1979 to the end of 1984, the upward swing of points from the beginning of 1985 to the end of 1986, or the double cluster of points in mid 1987. Causes for these patterns are difficult to pinpoint but not difficult to generalize.

Secular change was eliminated as a factor in this chart and the water vapor correction effect was partially reduced. What remains as the cause of these patterns? Most likely it is variations, often not random, in instrument calibrations (i.e. barometer and hygrometer), misalignment of the interferometer that may persist for several measurements, gradual or sudden change in laser vacuum wavelength, or a temporary mechanical distortion in the apparatus. These problems seem to be mitigated in the 1991 data but may return unless precautions are taken and improved instruments are acquired. Meticulous care is needed to make high precision measurements of any kind.

6.0 SUMMARY

This effort to reduce errors has improved measurement precision, both short and long term, and it has reduced measurement uncertainty. Improvements were significant or noteworthy in the following cases:

- a. The revised water vapor correction to air refractivity produced the most dramatic increases in precision and accuracy (see 4.4.8).
- b. Improved hygrometric measurements also contributed to the improvements but problems still exist in instrument reliability (see 4.4.8).
- c. Barometric pressure measurements are now more reliable but an uncertainty of 8 Pa (0.06 mm of Hg) remains. This can probably be improved with better instrumentation and calibration (see 4.4.1).
- d. Correcting for interferometric zero shift (deadpath correction) has little effect on long term values but improves short term and long term precision (see 4.4.3).
- e. The improved interferometer and scale axis alignment method, although not yet proven, should reduce errors from this source (see 4.4.6).
- f. Because laser vacuum wavelength measurements are now more readily available, errors from this source should be reduced (see 4.4.4).

Finally, the international interchanges with the BIPM meter bar verifies that the NIST line scale measurement process is consistent and well within the bounds of the world's best national laboratories (see 4.4.7). The simulation of control bar stability described in section 5.3 further demonstrates process stability. Process systematic error is now estimated to be 5 parts in 10^8 by the error budget method (see 4.1) and 3.5 parts in 10^8 by the independent measurements method (see 4.4.7). This is a significant improvement over the 10 parts in 10^8 value in 1987.

7. REFERENCES

1. Beers J.S. Length scale measurement procedures at the National Bureau of Standards. NBSIR 87-3625, 1987
2. Page B.L. Calibration of meter line standards of length at the National Bureau of Standards. J. of Res. of the NBS, January 1955, 54(1)
3. Cook H.D. and Marzetta L.A. An automatic fringe counting interferometer for use in the calibration of line scales. J. of Res. of the NBS, April - June 1961, 65C(2)
4. International Temperature Scale of 1990, Metrologia 26, 1989
5. Stone J.A., Phillips S.D., Mandolfo G.A. Corrections for Wavelength Variations in Precision Interferometry Displacement Measurements, in process
6. Edlén B. Metrologia, 1966, 2(2), 71
7. Jones F.E. The refractivity of air. J. of Res. National

- Bureau of Standards, Vol. 86, 1:27-30; 1980
8. (U.S.) Intergovernmental Panel (IPCC) Report on Climate Change, 1990
 9. BIPM Report on Meter Bar Round Robin
 10. Birch K.P. and Downs M.J. The results of comparisons between calculated and measured values of the refractive index of air. J. Phys. E: Sci. Instrum. 21, 694 (1988)
 11. Beers J.S. and Doiron T. Change in Edlén's air refractivity equation verified, *Metologia* 1992, 29, 315-316

Acknowledgements: The authors wish to thank Ralph Veale for his encouragement in this project and his patience in waiting for "just one more experiment" needed to complete the work. We are also indebted to Ted Doiron for many useful discussions and contributions, and especially for co-authoring the paper on testing the revised air refractivity index; to Jack Stone and Steven Phillips for their valuable work on the zero shift correction; and to Churchill Eisenhart for helping us through some knotty statistical problems.

