Development and Implementation of an ISO Quality System for the Time and Frequency Group, National Measurement Laboratory, Australia

Bruce Warrington and Peter Fisk

National Measurement Laboratory
CSIRO Telecommunications and Industrial Physics
PO Box 218, Lindfield NSW 2070
Sydney, Australia
Email: `bruce.warrington@tip.csiro.au`

Abstract

The Time and Frequency Group of the National Measurement Laboratory, CSIRO, in Sydney, Australia has developed a set of procedures, documentation and associated infrastructure for quality management of device calibrations. This has included the development of automated methods for collection and processing of calibration data. The quality manual, test methods, work instructions, and software form a tightly integrated package, and the quality system was granted third-party accreditation by an independent panel in October 1999 as being compliant with ISO Guide 25. The special character of time and frequency measurements presented some unusual challenges within the framework of the Guide. An overview of the development of the quality system is given, including test methods and the evaluation of uncertainties.

1. Introduction

The National Measurement Laboratory (NML) in Sydney, Australia, is completing a programme to develop quality management systems in all areas of device calibration. A quality system was developed within the Time and Frequency Group as part of this programme, and assessed in October 1999 against the requirements of ISO Guide 25. The assessment panel consisted of two international technical experts and an official from NATA, the National Association of Testing Authorities, who oversee the formal accreditation of testing and calibration laboratories within Australia.

There are three quality manuals which document the quality system, but only one was directly prepared by Time and Frequency staff. The top level of the quality system is the Quality Policy Manual for NML, which deals with quality policy throughout the Laboratory. The middle level is the Laboratory Operations Manual, which specifies procedures common to all calibration areas within NML (for example, central record keeping of the calibration history of an instrument, or the programme of internal audits). The third and lowest level is the Project Operations Manual, which deals with requirements specific to Time and Frequency; similar manuals are prepared for each 'Project' or calibration area.

The general requirements that we sought to address in preparing the Time and Frequency quality system were for standardisation and for record-keeping: for any device calibration, a standard procedure prepared to be appropriate for that device should be followed, and information recorded during the procedure should form a complete record of the calibration. This record should be able to demonstrate clearly how device data was recorded, how it was processed, and (importantly) how it was checked before the measurement report was formally issued. The approach we adopted was to develop each test method as an itemised check-list. A copy of the check-list is printed out for each calibration and followed in order; this standardises the testing procedure. At various points, spaces are provided on the check-list for the initials of the testing officer and another member of the Project staff; these provide written confirmation that particular steps have been carried out and checked, respectively. The completed check-list, the laboratory calibration notebook with associated measurements, and data files recorded electronically form the complete calibration record.

It is important to note that it is NML policy that the calibration result applies to the instrument *at the time of calibration only*. The uncertainty quoted in the measurement report is interpreted as the precision to which the instrument output could be measured at that time, not as specifying a confidence interval within which the instrument output will remain in the future. Under this policy, it is consequently the responsibility of the customer to account for additional uncertainties if the measurement report is used to specify such a confidence interval at some point in the future.

2. Technical Issues

The special nature of time and frequency measurement in comparison with other physical measurements requires special treatment within the quality system. The calculation of experimental uncertainties is significantly different from the usual statistical treatment, as is well known; in addition, however, the distinctive nature of atomic frequency standards also affects the requirements of the quality system.

2.1. Caesium Atomic Frequency Standards

Caesium frequency standards perform a direct, local realisation of the SI second within uncertainty limits specified by the manufacturer of the standard. Caesium frequency standards may therefore be regarded as primary standards, and as such do not necessarily need to be calibrated; in principle it is only necessary to confirm that they are operating correctly for them to be used as references for frequency measurements made with respect to the SI second. This is in sharp contrast to other calibration areas which must maintain a calibration chain for secondary standards back to primary standards of measurement.

At NML, correct operation of Cs standards is confirmed by continuous comparison with other NML standards, and comparison with International Atomic Time (TAI) using GPS common-view time transfer according to the protocol specified by BIPM. Participation in the BIPM GPS common-view intercomparison system provides traceability of NML's frequency measurements to BIPM, and links Australia's local timescale, UTC(AUS), to Coordinated Universal Time (UTC). The results of these intercomparisons are reported in BIPM *Circular T*. The local Cs standard can then be used as a reference for absolute epoch as well as for frequency, so that time-of-day can be provided as a calibration service. To support this service, and also a local requirement to maintain UTC(AUS) within a 1 µs tolerance of UTC, the quality system includes procedures for ensuring correct operation of the primary GPS time-transfer receiver, dealing for example with antenna coordinates, tracking schedules, internal receiver delays and control of cabling.

2.2. Calculation of Uncertainties

A significant issue in the development of the quality system is the treatment of experimental uncertainties. It is well known that the high precision of time and frequency measurements can readily reveal drift in oscillators under calibration, and that this drift invalidates the usual statistical treatment outlined in the first edition of the *Guide to the Expression of Uncertainty in Measurement* published by the ISO (hereafter the GUM). The GUM itself contains the following clause:

> **4.2.7** If the random variations in the observations of an input quantity are correlated, for example, in time, the mean and experimental standard deviation… may be inappropriate estimators of the desired statistics. In such cases, the observations should be analysed by statistical methods specially designed to treat a series of correlated, randomly-varying measurements. [p11]

Such a mathematical treatment is based instead on the Allan deviation (see for example the papers in *Characterization of Clocks and Oscillators*, NIST Technical Note 1337), and the GUM goes on to mention the Allan variance in a note following §4.2.7. Thus the Allan deviation should be used to evaluate the stability of a device under calibration, and the use of the Allan deviation is fully in accordance with the GUM.

However, two practical issues immediately appear:

- Over what measurement interval t should the Allan deviation $s_y(t)$ be calculated?
- As t approaches the total measurement time, the uncertainty in $s_y(t)$ increases; what is then the best method to determine $s_y(t)$?

Associated with these, a number of subsidiary practical decisions must be made concerning the implementation of a calculation of $s_y(t)$. The particular approach used by NML to answer these questions is outlined in the following section.

3. The NML Approach

3.1. Verification of Primary Standards and of Calibration Equipment

At NML, calibrations are made with reference to a particular commercial Cs beam standard, designated both as the local realisation of UTC and as the National Frequency Standard (hereafter referred to respectively as UTC(AUS) and the NFS). This standard is continuously intercompared with others in the local ensemble. The comparison is made every minute using software developed in-house running on two separate monitoring systems for redundancy, and the data records are maintained indefinitely. At the conclusion of each calibration, an explicit check is made that the ensemble was functioning correctly throughout the calibration, and a separate record of the corresponding intercomparison data is placed in the calibration file.

Before every calibration, the calibrating frequency counter is used to measure a known frequency derived from the NFS. This confirms correct operation of the counter, in particular that the timebase is externally referred to the NFS, and also performs an important additional verification of the software used to automate the data acquisition. The data recorded are stored as part of the data set associated with that calibration.

Additional requirements for cabling—for example, to ensure correct identification of signals and correct cable delays—are specified in the quality system. Essentially these consist of methods of labelling cables which are designated as ‘critical’ (those which, if disconnected or connected incorrectly, could invalidate a device calibration) together with associated requirements for maintaining documentation on measurements of cable delays.

3.2. Recording of Calibration Data

Calibration data are recorded using commercial frequency and time-interval counters interfaced to Linux computer hosts running dedicated software developed in-house. Appropriate counter configuration settings for the device under calibration are established by the testing officer; these are stored in a file kept as part of the calibration data set, and used by the calibration software to configure the counter.

For low-stability oscillators (typically quartz crystals), the calibrating counter is used to make direct measurements of the oscillator frequency, with the gate time and the length of the data set selected appropriately by the testing officer. For high-stability oscillators (typically rubidium atomic frequency standards), the counter is used to perform a direct phase comparison with the caesium reference over a longer time interval, typically several days.

3.3. Calculation of Measured Frequency and Estimation of Uncertainty

We list below the principal steps by which a measured frequency and an uncertainty estimate are obtained from a set of data recorded over a total measurement time T_{obs}. These are quoted directly from the Project quality manual (in italics), together with some additional explanatory notes.

1. *The true value of the output frequency of the oscillator at the time of calibration shall be estimated by calculating the mean frequency over the period of observation* T_{obs}.

    The optimal estimator of the output frequency depends on the noise properties of the oscillator. For convenience, we use the mean of the sample frequencies; this is optimal where the frequency fluctuations have a white spectral composition, and we avoid the complication of having to establish the exact noise properties of the device.

2. *The uncertainty quoted in the measurement report shall be expanded from the best available estimate, denoted* u_A*, of the RMS fluctuations in the frequency which would be obtained if the measurement specified in point 1 were repeated continuously with no delay.*

    That is, if we performed another N–1 calibrations immediately after the one completed, the RMS scatter in the set of N mean frequencies would equal the uncertainty in the first. In practice, we only perform one calibration, so we must calculate the estimate u_A from the available data.

3. *The quantity* u_A *shall be given by* $\sqrt{2}$ *times the Allan deviation* $s_y(t)$ *of the frequency measurements over an averaging time* $t = T_{obs}$*, the period of observation.*

    Where the frequency fluctuations have a white spectral composition, it can be shown that $s_y(t)$ is equal to the conventional standard deviation; where the frequency fluctuations have a different spectral composition, $s_y(t)$ represents the best estimate of the magnitude of frequency fluctuation over a given averaging time. The extra factor of $\sqrt{2}$ is required to match the definition in point 2.

4. *The quantity u_A shall be estimated by extrapolating $s_y(T_{obs}/6)$, being the longest time interval over which it is possible to calculate $s_y(t)$ with reasonable confidence limits (using the overlapping data technique), out to $s_y(T_{obs})$. This extrapolation shall be carried out conservatively, by assuming that the oscillator under test exhibits random walk frequency fluctuations for averaging times $t>T_{obs}/6$:*

$$u_A = \sqrt{2}\,\sqrt{6}\; s_y\left(t = \frac{T_{obs}}{6}\right).$$

   As t increases, so does the uncertainty in $s_y(t)$ calculated from the available data, even using the overlapping data technique. We take the position that the uncertainty is too large to be useful for $t>T_{obs}/6$, and we must therefore extrapolate. We assume random-walk frequency fluctuations ($s_y(t) \propto t^{+1/2}$); this is conservative, but we have no information on the noise behaviour of the instrument for $t>T_{obs}/6$ as $s_y(t)$ is too uncertain here.

5. *Further, for reasons of numerical convenience, the Allan deviation is only calculated for an averaging time which is 2^n times the averaging time t for a single reading, with n=0,1,2… Denote by T the largest such multiple less than $T_{obs}/6$. Then*

$$u_A = \sqrt{2}\,\sqrt{\frac{T_{obs}}{T}}\; s_y(t = T).$$

6. *The uncertainty quoted in the test report may need to be further expanded to account for the finite resolution of the counter/timer.*

   This is done following the prescription of the GUM, assuming that a counter reading x with minimum resolution of δ was drawn from a uniform distribution over the interval $[x-\delta/2, x+\delta/2]$.

7. *Calculated uncertainty contributions shall be combined following the procedures laid out in the ISO GUM. The uncertainty quoted in the measurement report shall be the combined standard uncertainty multiplied by a coverage factor k to give an expanded uncertainty representing an estimated confidence level.*

   It is not possible to calculate analytically the coverage factor k required for a specified confidence level, as the distribution associated with the uncertainty component due to the frequency fluctuations is not well known. However, following the general guidelines laid out in Annex G of the GUM, we adopt $k=2$ and assume that an expanded uncertainty equal to twice the combined uncertainty defines an interval having a level of confidence of approximately 95%.

This gives the complete prescription for obtaining the measured frequency and an uncertainty estimate from the set of calibration measurements. In practice, these calculations are carried out by in-house software, to ensure that all data is treated similarly. The calculation software is designed to operate on the output data files recorded by the data acquisition software.

3.3. Additional Issues in Uncertainty Evaluation

The uncertainty of the primary standard itself—that is, the uncertainty in the realisation of the SI second specified by the manufacturer of the caesium beam frequency standard—is in practice a negligible contribution to the combined uncertainty. Uncertainties associated with temperature-induced delay variations of distribution amplifiers and connecting cables are similarly negligible.

It can also be shown that dead time between measurements is not a significant source of error. If each reading is averaged over an interval t_a (this is the gate time set for the counter) and there is dead time t_d between readings, we can either set the measurement time $t = t_a + t_d$ or we can set $t = t_a$ and make a small correction, following for example Barnes and Allan 1990. If we take $t = t_a + t_d$, we overestimate $s_y(t)$ slightly; the overestimate is typically small with $t_d << t_a$. It is obviously especially important to determine $t_a + t_d$ empirically if phase measurements are being used to determine the instrument frequency f, so that $f_i=\Delta f_i/(t_a+t_d)$. Each reading is time-stamped by the data acquisition software for this purpose.

3.4. Software Verification

Correct operation of software for data acquisition and for calculation is clearly critical. All software is extensively verified before use. For example, data acquisition software is verified at the beginning of each calibration by recording a known frequency (see §3.1), and software which calculates the Allan deviation is verified by cross-checking against an independent calculation carried out by other means. Software must be re-verified on any modification. Software verification records are maintained with other quality system documentation.

3.3. Measurement Reports

Measurement reports are prepared using fixed document templates. A facility exists for merging information extracted from the client database into a measurement report using the template, to minimise transcription errors. The measurement report is checked by one or two staff members other than the testing officer before being passed to the head of the Group for signature.

4. Conclusions

The quality system has operated successfully within the Time and Frequency Project for more than a year.

An ongoing issue which is likely to become more important in future is that of software verification. At present we are exploring the possible use of CVS, a public-domain package for UNIX systems which provides a central software repository, maintains records of modifications and can be used to enforce software version control.

The National Measurement Laboratory has extended the development of its quality systems to require assessment against the newer ISO 17025. The most recent assessments of other calibration areas at NML have been carried out against this standard, as will be the next re-assessment made of the quality system within the Time and Frequency Project. The extensions required largely fall within the Laboratory Operations Manual common to all calibration areas, so that substantial additional work should not be necessary.

Acknowledgments

We were able to draw on the experience of other calibration areas at NML with their own Project quality systems already developed and assessed; in particular, Time and Frequency staff benefited greatly from the prior experience of staff in the Mass Project. Malcolm Lawn, John Thorn and Steve Quigg assisted with the implementation of the quality system and with additional documentation. We would also like to thank the assessment panel for their thorough and careful assessment of the quality system.

References

Guide to the Expression of Uncertainty in Measurement 1993, first edition, ISO, Geneva.

Characterization of Clocks and Oscillators 1990, ed. D. B. Sullivan, D. W. Allan, D. A. Howe and F. L. Walls, NIST Technical Note 1337.

Barnes J. A. and Allan D. W. 1990, "Variances Based on Data with Dead Time Between the Measurements", NIST Technical Note 1318 (reprinted in *Characterization of Clocks and Oscillators*)