Abstract

A few years ago a paper was presented at a SPIE meeting by DiVittorio and Walton[1].
The authors wanted to compare three instruments available commercially in view of measuring and monitoring coatings of telescope mirrors. For doing this, they prepared three sample mirrors coated with aluminum and over coated with a SiO2 layer of unknown thickness. The sample (B and C) were grinded to present a lowering of reflectivity and sample A was left untouched. They proceeded to have the three samples measured by an external laboratory using a Cary 5 spectrometer to have 'reference' reflectivity curves for the three samples.

They did not worry about the sharp spectral features appearing in the reference reflectivity measurements. These features, however, allow computing the thickness of the SiO2 layers. These turn out to be very thick (resp. 208, 313 and 303 nm for samples A, B and C), i.e. between a quarter wave and one wave of the spectral bands of interest, so that the sample 'mirrors' are in fact interference filters of poor quality. Moreover, second mistake, it is not at all representative to grind the over coating of a mirror in order to adjust its 'reflectivity' to a chosen values at 650 nm.

The consequence of the first error is that the reflectivity is extremely sensitive to the working angle of incidence of each instrument. The consequence of the second error is that the measurements are extremely sensitive to the acceptance angle of each instrument.

Hence, one can expect that measurements will differ (up and down), depending on the wave band, and that the instrument with the wider acceptance angle will show results even higher than the reference because it includes an amount of scattered light into the 'reflectivity' measurement.

Each of those effects can lead to differences up to +/- 5 % and can combine differently at various wavelengths, so that sometimes they compensate, other times they add up; the outcome is that one reads for all three tested instruments absolute discrepancies ranging from -9.7% to +4.4%. These discrepancies are accepted by the authors as instrumental errors without any criticism although the three instruments are certified to measure within 1 or 2% by their vendors.

Telescope mirrors will never show those two bad features present on the samples because either they are not over coated or they are over coated with a layer that does not show sharp absorption features in the middle of the useful range. Moreover in a large telescope, most over coated mirrors are small diagonal mirrors used at an angle of 45° and the result of the reflectometer working at that angle would be better than the reference measurement made at 7° incidence, even for a thick coating. When the telescope mirror becomes dusty or even dirty, its scattering properties are always at least an order of magnitude smaller than those of the grinded samples. It would have been much better to prepare samples with pure aluminum coatings of various thicknesses in order to obtain a

[1] Michael DiVittorio and Joshua Walton, Reflectometer shoot-out: comparing the performance and accuracy of hand held reflectometers, Proc. SPIE 5494, 434 (2004).

range of representative reflectivities. Pure Al coating is covered by a thin protective layer of Al2O3 as soon as it is exposed to air after coating. This is sufficient for insuring stability during the test operations.

One of the three tested instruments is specifically build in order to measure the albedo or the gloss of a non reflecting surface and in no case the reflectivity. It would measure the reflectivity only in the case of a mirror free of *any* scattering; it is thus a plain mistake to use it with samples prepared to scatter heavily (compared to any dirty telescope mirror).

The procedure of 'normalization' makes no sense when it is used to 'correct' for discrepancies other than the scale error of the instrument.

The standard deviation makes no sense to analyze the sensitivity of the measurements with temperature (or any other parameter). Its use should be restricted for analyzing measurements made in the same conditions and affected by accidental errors.

The sum of these elementary mistakes is such that they show a deep ignorance of the techniques of metrology. *'Ex aure asinus'* [one can recognize an ass when looking at its ears] : it would be nice that engineers in charge of telescope's mirrors maintenance look carefully at this 'paper' and at the ***arguments*** developed hereafter before showing their ears.

Adequacy of the sample preparation

Three samples were produced in order to cover a range of reflectivity and scattering representative of telescope mirrors.

It is shown hereafter that the three samples do not represent anything similar to the degradation of telescope mirrors. They induce artifacts in the measurements which exceed by far the accuracy that one wish to appreciate. The main points that will be discussed are: the effect of poor coating of the samples; the effect of scattering; the effect of wavelength shift of the Cary 5 used to get the primary reflectivity data of the samples.

Coating of the samples and incidence angle

The samples were commercially available mirrors coated with Al and a SiO2 protecting layer. It is appreciated that the experiment demands stable samples and that the authors believe uncoated mirrors would have raised difficulties. Quality protected Al coating, however, are protected with a thin SiO2 coating that does not interfere with the reflectivity of the mirror in the range where it is used. For instance, Fig. 1 shows the measured and the computed reflectivity of the IRIS gauge for two incidence angles.

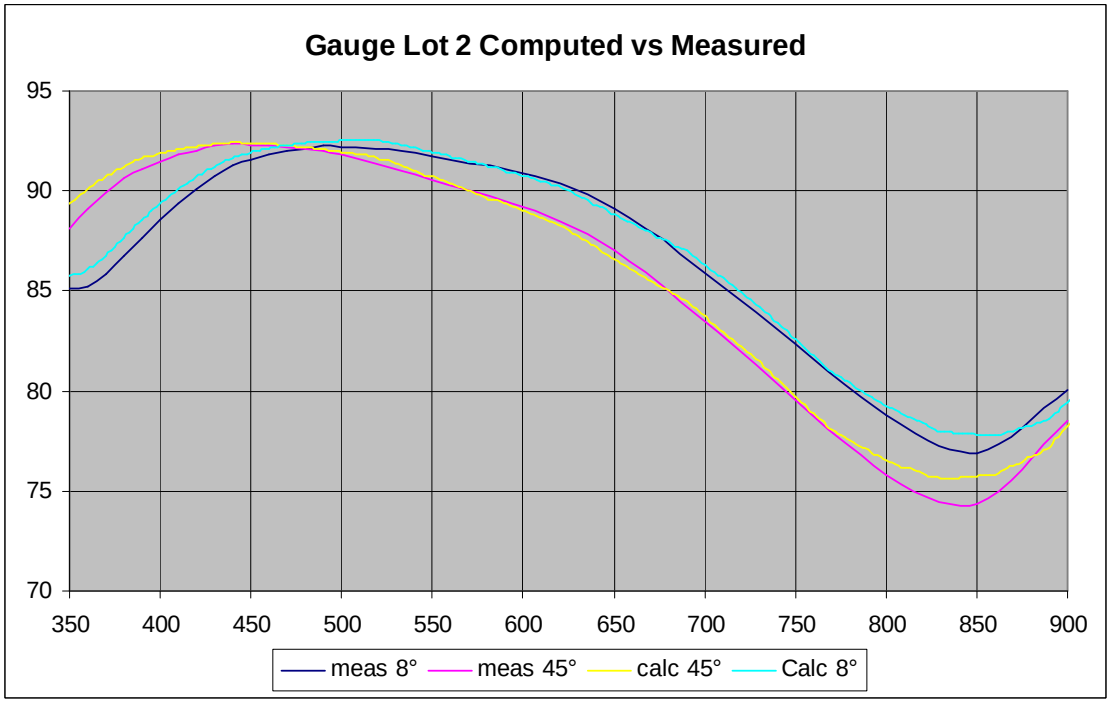


Fig.1 Quality Al over coated with 165 nm SiO2

The wide ditch around 835 nm is due to Al and is unavoidable when Al is used for coating the mirror; the first order ditch due to the SiO2 coating is at a wavelength lower than 350 nm; this is obtained by a 165 nm thick coating. The small differences between measured and computed reflectivity are due to differences in the complex refractive index used for the materials (Al and SiO2)[2]. The wavelengths utilized in IRIS are 460, 530, 650 and 860 nm all situated on smooth sections of the reflectivity spectrum. In the future, the IRIS' gauges will be protected by a SiO2 layer of 20 nm only. On the other hand, the overcoat's of the samples A, B and C used in the comparison work are much thicker so that the first order and the second order absorption features of the coating appears right in the middle of the useful spectral region as can be seen on Fig. 2. They act as interference filters.

[2] Computations use the data given in the Handbook of optical constant of Solids.

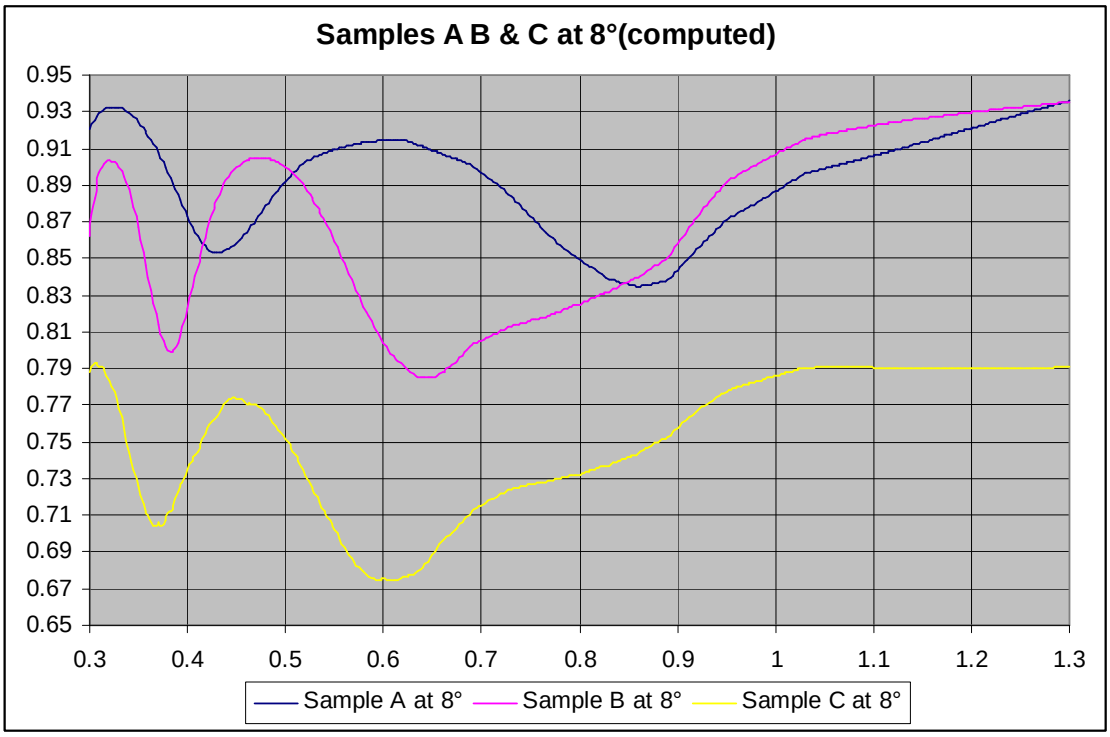


Fig. 2 Computed coating reflectivity of samples A, B and C used in the test.

These computed curves compare very well with those given in the paper by DiVittorio and Walton (hereafter referenced as “the paper”). The spectral feature directly seen in the curves presented in the paper offer a very accurate way of measuring the thickness of the coatings.

The protective SiO2 layer thickness is the driving factor of the reflectivity variations that are observed in the three samples. Sample A has a thickness of 208 nm, sample B 313 nm and sample C 303 nm. All three samples are of very poor quality for use in optical metrology; sample A has the first order SiO2 absorption feature at 432 nm quite close to the useful wavelength range, sample B and C have this first order absorption feature shifted to respectively 642 and 611 nm joining with the Al feature, while the second order absorption feature appears at 382 and 371 nm. In between the two spectral features, the reflectivity varies wildly by 10% up and down exactly in the useful range of wavelengths. Of course, this kind of situation is *never* encountered when measuring telescope mirrors. When such mirrors are coated, the coating has to yield a smooth high reflectivity in the useful wavelength band, so that a practical reflecto-meter always measures within a smooth spectral region. The consequences of this situation for the comparative measurements is best understood when plotting the same reflectivity curves for an angle of 45° as shown in Fig.3.

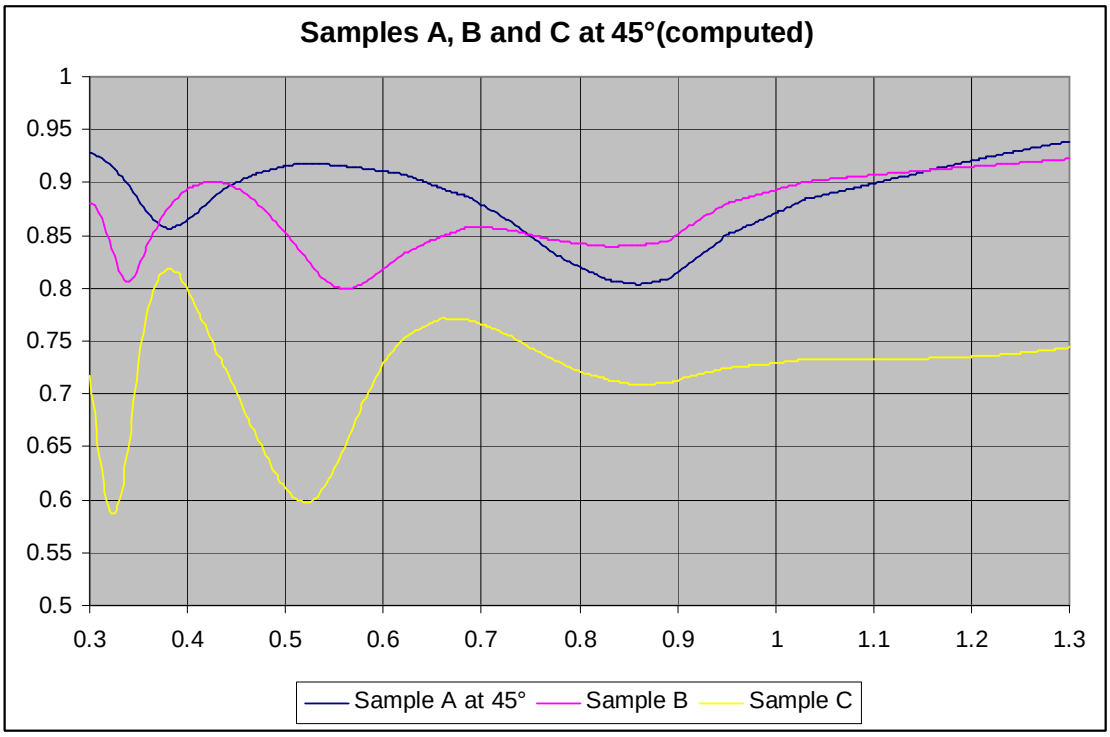


Fig. 3 Computed reflectivity of the three samples at 45° incidence.

The wavelength shifts of the sharp spectral features will induce a large difference in the measurements made at 45°, and the error will be very different for sample A and for Sample B or C, because the spectral features are not at the same place in the three samples. Let us insist that these errors are pure artifacts induced by the use of improper samples and that these errors will *never appear* when measuring actual telescope mirrors (because the latter show no spectral features in their reflectivity).

This point is illustrated on Fig. 4., where one has plotted the reflectivity of the same sample C (the worse case) as seen by the reference instrument (8° incidence), the SMS instrument (25° incidence) and the DMO instrument (45° incidence). For the SMS instrument, we have plotted the curves for S- and P-polarization, since this instrument uses a laser diode as a source; one should consider the curve corresponding to the polarization plane of the source (if known). DMO instrument uses strictly unpolarized sources. The Minolta instrument is not a reflectometer; it uses an integrating sphere and measures the integrated reflectivity for a solid angle approaching 2π (what astronomers call the albedo) that is exactly what one does not want to measure for monitoring the quality of a telescope mirror.
Considering the curve, one sees that the error induced by the spectral features would be for the DMO instrument: -10% at 460 and 530, +7% at 650 and -4% at 860 nm. These are very approximate because the DMO sources have a rather large spectral width (30 to 60 nm) and one should integrate over the band pass. The induced error on the SMS instrument would be approximately between +2 and +4% at 670 nm and -2% at 1300 nm. (Note that this is only the contribution of the incidence angle).

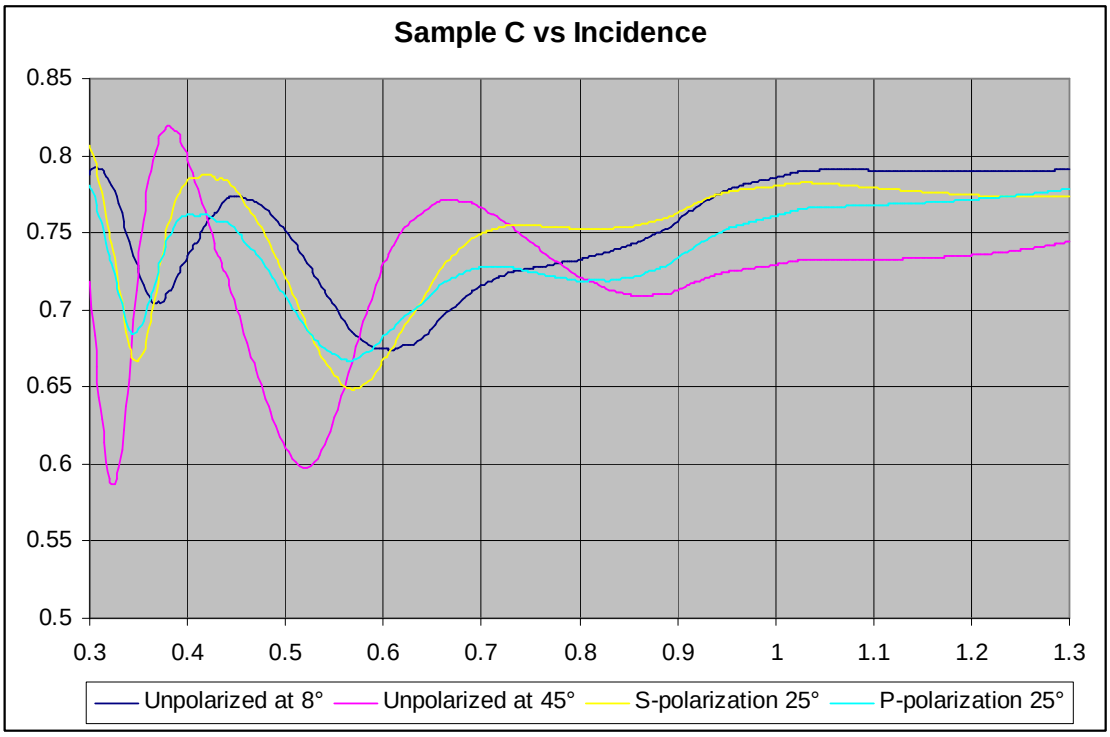


Fig.4. Computed reflectivity of sample C for three incidence angles.

As already mentioned, this is an artifact of the test due to the steep slope of the spectral features of the samples. A telescope mirror, either has no coating (majority of primary mirrors), or has a coating specially designed *not* to show any steep spectral feature in the useful range. Moreover, coated mirrors are mostly used on small diagonal mirrors that are used at 45° incidence and are best measured with an instrument working at this angle.

The DMO instrument has been specially designed for monitoring astronomical mirrors; it works naturally on pure Al or pure gold mirrors and on coated diagonals. It has a build in correction for the case where one needs to know the normal incidence reflectivity of a coated mirror used at normal incidence[3] with an absolute accuracy better than 1%. It is normally delivered with the Fresnel correction for dielectric material, but the correction for any coating can be inscribed in the calibration if the user gives the coating specifications[4].

When the instrument is used to monitor the degradation of the mirrors with time; it can be used as such without care of any correction, knowing that when measuring a coated secondary, the absolute reflectivity value must be corrected by a fixed amount: -1% at 460; +1.5% at 530; +2.5% at 650 and +2.5% at 860 nm; less if the coating is thinner.

[3] The only practical case is coated secondary mirrors; In the VLT Nashmit focus for instance, there is one uncoated mirror at normal incidence (primary), one normal incidence mirror that could be coated (secondary), and 6 diagonal mirrors at various angles (3 working at 45°).

[4] When The DMO instrument was designed, coated mirrors were quite unusual; the new model has the possibility of recording several correction curves for a number of coatings. This is not the model that was used during the tests.

Effect of scattering

The mirrors exposed to weather are degrading because of chemical etching and because of dust deposition and abrasion. The relative contribution of these effects depends on the coating (silver decays rapidly by chemical pollution if any trace of Sulfur is present in the atmosphere, over-coated mirrors are much less sensitive to chemical decay). Chemical decay produces mainly a loss of pure reflectivity and is very damageable for the light collection; dust induces mainly scattering which has little effect both on reflectivity and on resolving power, unless the pollution is massive. If the reflectivity of a mirror would go down by more than 10% due to scattering, one would not "see" a mirror anymore: it would appear like a ground glass. A telescope mirror will never loose 20% of its reflectivity by scattering as it is the case for Sample C of the test. Anyway, a reflectometer is build to measure specular reflectivity of an imaging mirror and is completely lost when forced to measure a diffusing mirror.

This is due to the angle of acceptance of the instrument: a good specular reflectometer has a very small acceptance angle; that is it measures only the light reflected in a small solid angle around the specular reflected beam. The international standards use two conventional values for this angle: namely 2° (narrow beam) and 10 ° (wide beam); this is already too much for measuring correctly the specular reflectivity of a diffraction limited mirror.

Now, let us examine quantitatively what is the effect of the acceptance angle on the measurement of the reflectivity *when the loss of reflectivity is mainly due to scattering.*

Fig. 5 shows the computed scattering of a point source in the image plane, assuming a total integrated reflectivity of 89.65% and a scattering fraction of 28% (and a specular fraction of 72%) representative of sample B. Fig. 6 shows the same for a scattering fraction of 55% representative of sample C.

An astronomer is interested in knowing the specular reflectivity because this is exactly the proportion of incoming power which he will find in the star image. This specular reflectivity is 64.71% (~72% of 89.65%) for sample B and 40.64% for sample C. One could measure these values with an *hypothetic* instrument with a minute angle of acceptance; The measured "reflectivity" with three normally available instruments (1.5°, 2° and 2.5° acceptance) is shown in the figures; they vary from 74.84% up to 83.717% (60.53% to 78.01% for sample C) for acceptance angles varying from 1.5° to 2.5°. Note that all three measurements are far from the truth, they are 10 to 19% too high for sample B; the instrument with the smallest acceptance is the closest to the correct value.

This is easy to understand: unless you have an instrument that has an extremely narrow angle of acceptance, it is not possible to measure the specular reflectivity of a mirror in the presence of any significant amount of scattering. What you measure with an instrument having a few degrees of acceptance is closer to the albedo. If you insist in knowing the absolute specular reflectivity, you should measure separately the total

scattering, using a scatterometer, and subtract the amount of scattered power from the albedo[5].

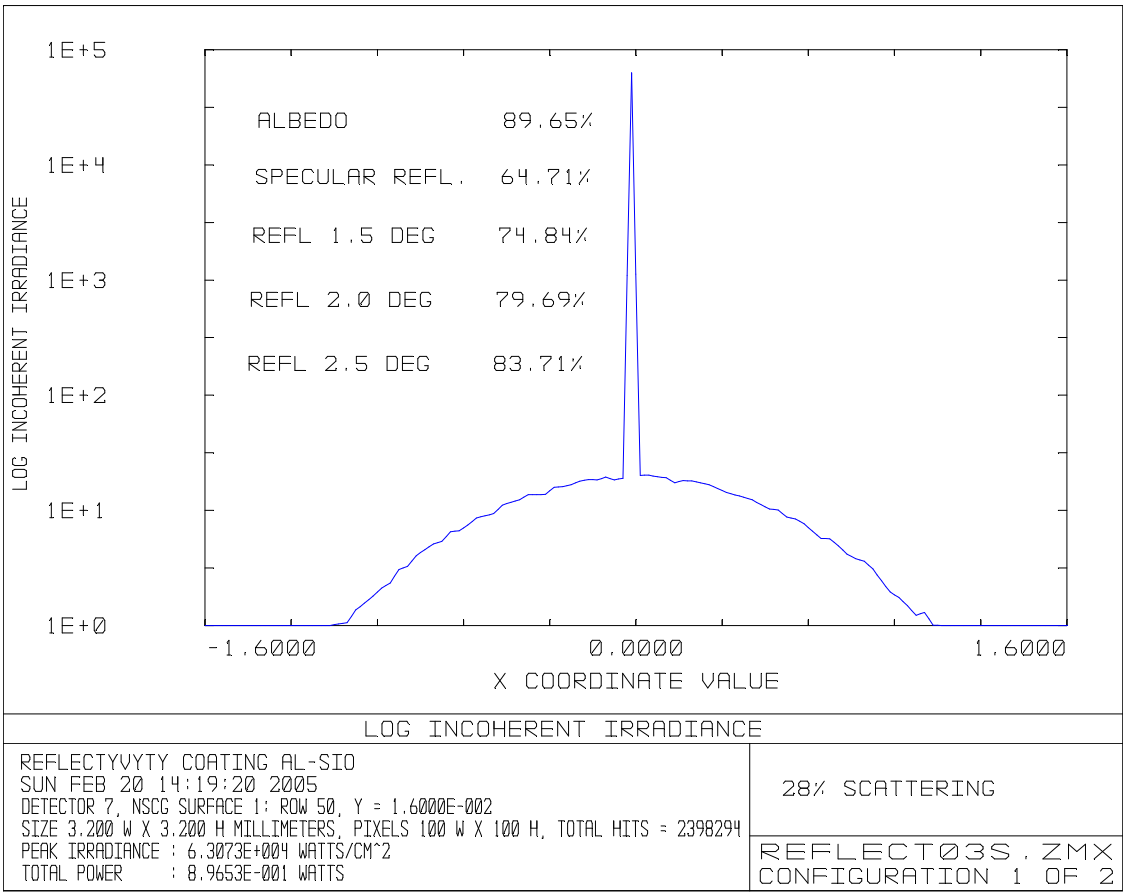


Fig.5. Sample reflectivity with 28% scattering and computed reflectivity for various angles of acceptance of the reflectometer.

From this example, one can also understand why mirrors apparently grossly degraded by dust diffusion, show little degradation when measured for reflectivity: if the used reflectometer has a large enough acceptance angle, it will always capture, measure and add most of the power which is scattered to the power which is truly reflected although the former would be lost for an astronomical image. It is also easy to understand why the Minolta instrument systematically shows results higher than those of the Cary 5: it is designed to measure the *albedo* and the *gloss*[6], not the TIS and the *specular reflectivity*.

[5] This is however a very difficult metrology problem because the TIS should be measured including angles as small as 1° from the specular angle. For instance the Minolta gloss meter measures the TIS excluding scattering angles < 8°, thereby excluding 20% to 60% of the true TIS depending on the cleanliness of the mirror. These are not small errors and show that this instrument is improper to perform measurements on an imaging mirror.

[6] The gloss of a diffusing surface (typically a sheet of paper) is the sharp increase in luminosity when looking a source of light close to the incidence angle. It is defined and measured in industry as the total amount of light diffused in a cone of 8° semi aperture around the specular direction.

The conclusion of this is that when a mirror shows an appreciable amount of scattering, the result of reflectivity measurement is *extremely sensitive* to the acceptance angle of the reflectometer and hence, the reflectivity measurements loose progressively their meaning as scattering is increased. One should then rely on the value of the BRDF (Bidimensional Reflective Distribution Function) to evaluate the damage to the mirror due to exposure. Usually, telescope mirrors are not in such a bad state and reflectivity measurements remain useful to monitor the decay rather than to perform absolute measurement. But grinding a test plate to compare instruments with unknown acceptance angles is certainly completely misleading.

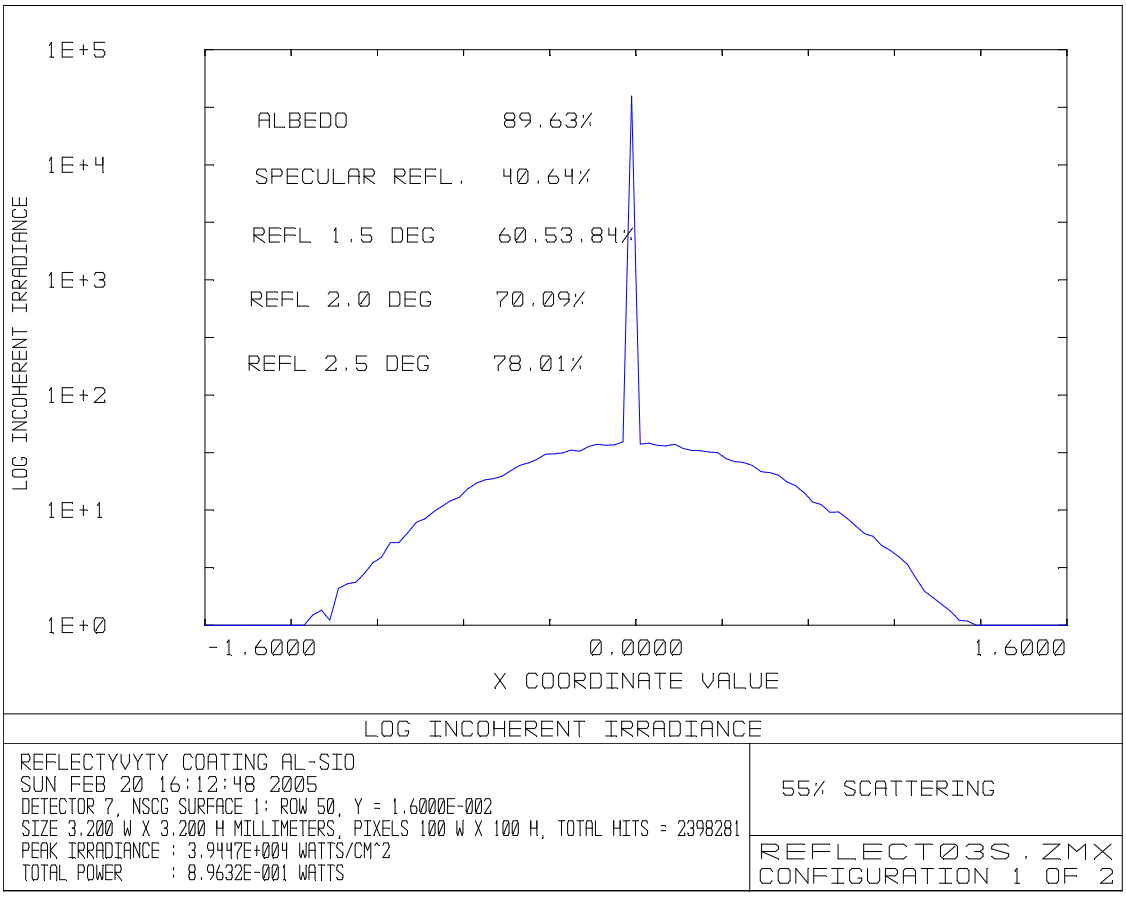


Fig. 6. Same computation for 55% scattering representative of sample C

For these computations, the albedo was computed with an acceptance angle of 5.1° and the specular reflectivity with an acceptance angle of 9.5 arcmin[7]. The level of scattering has been adjusted to yield a "reflectivity" of 80% for sample B and 70% for sample C when measured with the classical 2° acceptance angle. The scattering function was a Gaussian with σ=.035. These simulations are not intended to reproduce quantitatively the discrepancies shown in the paper because one does not know the acceptance angle of the

[7] It has also been checked that reducing further the acceptance angle to 5 arcmin did not produce any difference in the measured reflectivity.

Cary 5 and of the SMS instruments (Minolta has 90° and DMO has 2.0°), and one has no idea of the scattering distribution function produced by the grinding operation. Moreover, as already seen, the difference of reflectivity of the three samples at 670 nm is a combination of interferometric effect due to due to the overcoat thickness and scattering due to the grinding. The computations simply show how much the measured reflectivity is sensitive to these parameters when a significant amount of scattering is present and hence how inadequate the sample making procedure was. These effects completely explain the differences obtained between the measurements : they are not due to instrument shortcomings, but entirely to wrong measurement procedures.

The Cary 5 measurements

It can be seen in Figure 8 of the paper and it is mentioned in the text that the spectral measurements of the three samples by the Cary 5 at ODA show a "glitch" at 800 nm caused by a grating change. This glitch is interpreted by the authors of the paper as a reflectivity offset more pronounced on the degraded samples, with a maximum of .5%. It is then concluded that, in view of the smallness of this discrepancy it has no significant impact on the results of the test.

It is not understandable how a change of grating can lead to a shift in the measured reflectivity; the VW technique is robust and is not sensitive at all to the processing of the entering beam: it measures exactly the reflectivity for the entering beam as it is. What happened most probably is that the change of grating introduced a shift in *wavelength*. The relation between the angle of the grating and the wavelength scale was not the same for both gratings. This is an invariant shift, which means that the spectrum is displaced by a fixed amount on the scale corresponding to the value of the angular error. Of course one does not know which part of the spectrum is displaced with respect to the other or even worse whether both parts are displaced by a slightly different amount. The fact that the displacement varies from one run to the next is more worrying; it means that there is some backlash in the mounts of the grating. This throws a shadow of doubts on the pertinence of the absolute calibration used as reference in the paper.

Note that the effect of the improper and variable wavelength scale has greatly varying effects on the reflectivity results: where the slope of the spectrum is low it has little effect, while in places where the slope is large the effect is accordingly important. As roughly computed from Figure 8 of the paper the shifts are -6 nm (out) for sample A, +17 nm (in) for sample B and +38 nm (in) for sample C. For the latter, if you shift the wavelength by the same amount at 530 nm the reflectivity shift would be 2.7%, far from negligible as this error combines with all the others that we have mentioned so far. Again, note that if the samples would have a gentle spectral reflectivity curve the slope would be everywhere quite small and the glitch would be of little consequence. Again the perversity of this error stems from the very bad choice of the samples mainly samples B and C that have a much too thick SiO2 coating.

It is clear from this discussion that it would have been much better to use pure Al coatings. These are naturally protected by a very thin (5 to 10 nm) layer of Al2O3 formed

by exposure to the atmosphere as soon as the vacuum is broken after aluminization. The stability of such a coating is sufficient to insure excellent measurements for several weeks and much more if they are kept in a safe atmosphere. The various values of reflectivity would be more difficult to obtain with pure Al, the best way being to produce samples with thinner Al layers.

What measurement does exactly the MNLT instrument

As already stated, the MNLT instrument is not a reflectometer (i.e. an instrument aimed at measuring the specular reflectivity of a mirror); its use is to measure the albedo and the color of more or less diffusing surfaces like printed images. It is an error to use it for measuring telescope mirrors, because in this case, the albedo is exactly what one does *not* want to measure. The MNLT can be configured to exclude from measurements a cone of 8° semi aperture. This is obtained by removing a port at the specular angle. By substracting the measurement 'port open' from the measurement 'port closed', one measures the 'gloss' of the diffusing sample. Minolta user's manual calls the measurement with the port open 'specular component excluded' (SCE) and the measurement with the port closed 'specular component included' (SCI). It is incredible that the authors believe that "*the difference of the 2 (SCI-SCE) should equal the specular reflectivity*." This difference is equal to the gloss and it is equal to the specular reflectivity only when the TIS is equal to zero, that is when there is no scattering. In practical cases of exposed telescopes the gloss differs from the specular reflectivity by several percent (amount unknown between 0.5% and 12%, depending on the cleanliness of the mirror, that is exactly what one wants to measure!). The measurement 'port open' is also very different from the TIS depending on the shape of the BRDF which is totally unknown. It is usually comprised between 20% and 65% of the TIS. It can in no way be accepted as a 'measurement' of the TIS.

The discussion about the acceptance angle shows that, for a diffusing mirror, the difference between specular reflectivity and albedo can be very large and will grow larger as the scattering of the mirror goes up. That is the main reason why the procedure of making the reflectivity sample was not acceptable.

The way the MNLT works in view of the angular acceptance and of the incidence angle is somewhat complicated and could be difficult to understand by people not familiar with optical metrology. It is explained hereafter.

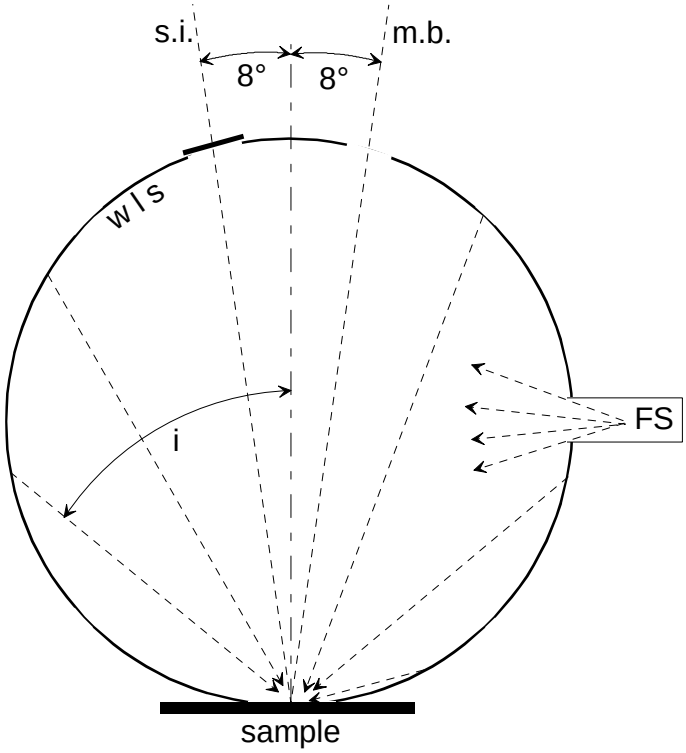


Fig. 7 Optical structure of the MNLT albedo meter.

The main features of MNLT lighting system is represented in Fig. 7. It is very important to well understand this technique in order to see its implications on reflectivity measurements.

The lighting is obtained by a flash lamp (FS) lighting the inside of an integrating sphere. The latter is a hollow sphere with the inside wall painted with a white lambertian scatterer (wls). The sphere is provided with several ports in order to introduce light, position a sample or extract light to be measured. It is a well known property of an integrating sphere that whatever the lighting, the flux at any port not seeing directly the source of light is uniform with position and angle. This means, in the case of the MNLT instrument, that the sample receive the same flux from any direction **i**. The measured beam is fixed at 8° and exits through a dedicated port. Finally, there is a port symmetric to the measuring beam port, which can be used for suppressing the incidence light in a cone of 8° semi aperture around the specular angle, that is for measuring the albedo minus the gloss. This is intended for measuring paper or paint scattering, not a specular mirror unless the latter is so deteriorated that it appears glossy and not reflecting, or absolutely clean (no scattering).

To describe the consequences of this type of lighting on our problem, i.e. measuring the specular reflectivity of slightly scattering mirrors, one has to analyze the situation in still more details.

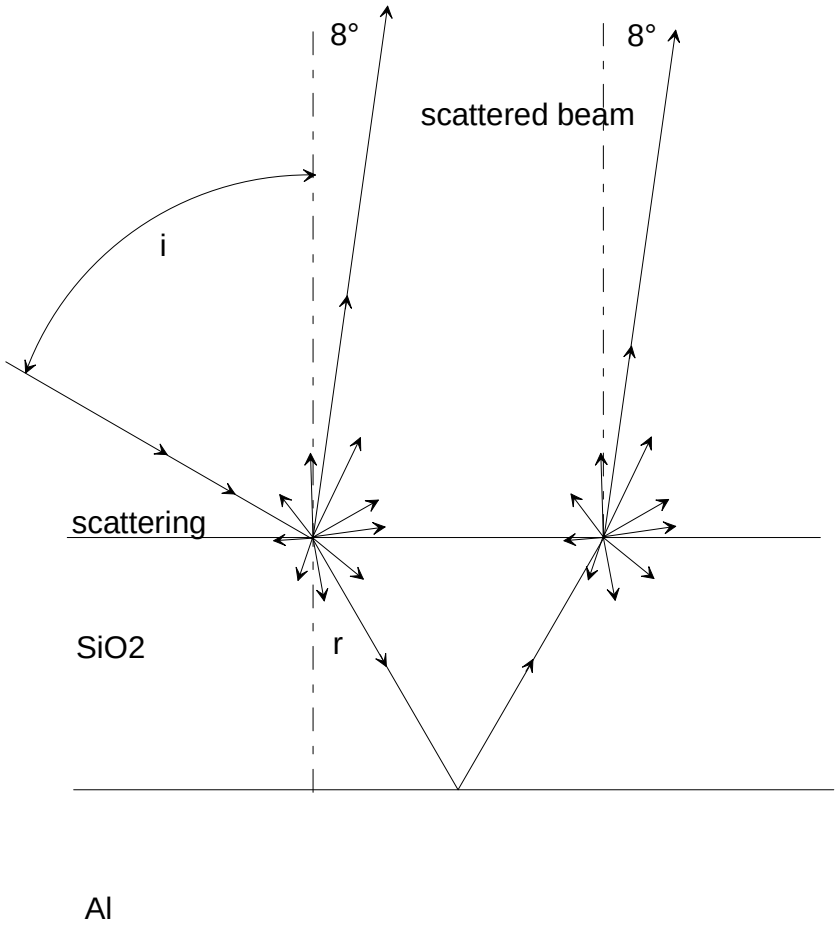


Fig. 8 Light reflection and scattering on a SiO2 over coated Al coating

Fig. 8 shows the beam path of an incident beam at angle **i** reflected on an Al coating and diffused on a SiO2 over layer. The scattering of interest occurs at 8° since only this angle is measured by MNLT. As we have seen, the angle **i** can take any value over the hemisphere (each angle with the same flux). The total measured scattered light results from the integration of all the incidence beams over the hemisphere, including the beam incident at 8° that corresponds to the specular reflectivity. Hence, the MNLT measures exactly what we call the albedo of the mirror, which can be an arbitrary value off the correct specular reflectivity, depending on the scattering properties of the SiO2 layer.

Moreover, the dielectric layer acts as an interference filter if it is not very thin, as we have discussed at length. The retardation due to the layer depends on its thickness **e**, its refractive index **n** and the incidence angle. It can be computed as:

$$OPD = 2 \times e \times \frac{[n - \sin(r) \times \sin(s)]}{\cos(r)}$$

This value is longer than the OPD of the specular beam (i=s) for any value of i≠s.
This means that the spectral features of the over coating are shifted to longer wavelength for increasing angles of incidence. It is a long and tedious computation work to evaluate the total result of these shifts and it has not been done; it is an integration on all incidence angles and would result in a mixture and a spreading of the sharp spectral features that one observes at one particular incidence[8].

The total effect of the thick SiO2 scattering over coat is very complicated in the case of the Minolta instrument. It shows sharp spectral features which are shifted differently for various angle of incidence. The integrating sphere of the MLNT instrument washes out the effect of spectral features but yields surely a result which is closer to the albedo than to the reflectivity (i.e. too large). Each of these effects can have a magnitude of the order of +/- 5% and combine in various ways so that sometimes they reinforce, sometimes they counteract each other, giving in all a very large uncertainty in the measurements.

Data processing

Normalization procedure

The normalization[9] procedure as explained in the paper is difficult to understand, and we redefine the terms as follows:

There are three samples A, B and C that are calibrated at ODA; let us call the result of these calibrations A_c, B_c and C_c.

Then the three samples are measured by instrument x (x being MNLT, SMS or DMO); let us call these measurements A_x, B_x and C_x.

The *calibration* of measurements B and C vs. measurement of standard A means:

$$B_{nx} = B_x * A_c/A_x \quad \text{and} \quad C_{nx} = C_x * A_c/A_x \qquad (1)$$

But in the paper, Equation 2 seems to state that

$$B_{nx} = B_x * A_x/A_c \quad \text{and} \quad C_{nx} = C_x * A_x/A_c \qquad (2)$$

[8] In the same way as an interference filter would show washed out transmissions features (less sharp and less deep) if used in a fast converging beam.

[9] This term is improper for the operation defined by equation 2 of the paper (where the calibration factor has been inverted.) This is in fact a 'calibration' of the instrument scale. A normalization operation on the data would require $B_{nx} = B_x * A_c/B_c$ in such a way that if you normalize the B calibration measurement it becomes identical with the A calibration measurement, that is $B_{nc} = B_c*A_c/B_c = A_c$. Normalization of several curves means equating the values of all of them to the chosen reference at one place (and loosing the scale) while calibration means putting all the curves on the same scale as the chosen reference. It is practically the contrary of the former.

The latter equation however makes no sense and is probably a printing error (although it is also wrongly described in the text).

Anyhow the authors utilize the formulation (1), the correct one for calibration, for plotting the results in Figure 10 of the paper since the values obtained by this formula are close to the plotted ones except for a few cases which are difficult to understand. Note that this equation operates a scale correction; it is a fixed factor for each instrument. It is clear however that one cannot correct errors due to the incidence angle or to scattering originating in a bad and non representative sample by a scale correction of the instrument, except if one does not understand the least what is going on in a measurement process.

The plots of Figure 9 and 10 of the paper are the measured departures of the measurements made by the three tested instruments from the true value measured by ODA during the calibration process; the mathematical expression for this operation is:

$$EAx = Ax-Ac; \qquad EBx = Bx-Bc \quad \text{and} \quad ECx = Cx-Cc$$

With evident notations: EAx = Error on measurement of sample A with instrument x etc.

Raw data taken from Figure 8 of the paper												
	ODA A	MNLT A	SMS A	DMO A	ODA B	MNLT B	SMS B	DMO B	ODA C	MNLT C	SMS C	DMO C
460	0.859	0.875		0.897	0.891	0.909		0.866	0.78	0.808		0.717
530	0.899	0.916		0.914	0.864	0.888		0.81	0.741	0.777		0.654
660	0.902	0.925	0.894	0.893	0.784	0.805	0.801	0.819	0.705	0.735	0.725	0.687
880	0.846			0.77	0.856			0.796	0.755			0.658
1300	0.926		0.958		0.91		0.95		0.79		0.834	

Differences with ODA values measured on Figure 8 of the paper												
	ODA A	MNLT A	SMS A	DMO A	ODA B	MNLT B	SMS B	DMO B	ODA C	MNLT C	SMS C	DMO C
460	0%	1.6%		3.8%	0%	1.8%		-2.5%	0%	2.8%		-6.3%
530	0%	1.7%		1.5%	0%	2.4%		-5.4%	0%	3.6%		-8.7%
660	0%	2.3%	-0.8%	-0.9%	0%	2.1%	1.7%	3.5%	0%	3.0%	2.0%	-1.8%
880	0%			-7.6%	0%			-6.0%	0%			-9.7%
1300	0%		3.2%		0%		4.0%		0%		4.4%	

The tables show the raw data directly taken from Figure 8. of the paper, and the differences with the ODA measurements measured on the same figure. All the numbers of the table of differences are slightly higher when drawn on figure 9. of the paper (by 4 to 8%) except three of the figures of sample C measured by DMO; the 9.7% figure is plotted as 12% (24% increase) on figure 9 and is again increased to 14% (increase of 44%!!) in the text[10] analyzing the very same figure! The 8.7% is plotted as 11% (increase of 27% and the 6.3% is plotted as 7.7% (increase of 22%). Moreover, in view of what we have seen about the difficulties raised by the samples, it would be wise (and fair) to compare the instruments only at common wavelength because nobody knows what error a given instrument would suffer at a wavelength where it does not operate. If we compare

[10] As a matter of facts, the text says “The DMO unit exhibits errors of up to 9% for standard A, and up to 14% for standard B and C”. Hence the 6% error on sample B is thereby increased by a factor of 2.3

the three instruments at 660 nm (interpolating DMO from 650 to 660 and SMS from 670 to 660 for the sake of comparison) one obtains the following table.

	Sample A	Sample B	Sample C
MNLT	2.3%	2.1%	3.0%
SMS	-0.8%	1.7%	2.0%
DMO	-0.9%	3.5%	1.8%

This fair comparison is very far from the message one reads in the paper.

The normalized (actually calibrated) differences are computed in the following table

Differences recalibrated to ODA A												
	ODA A	MNLT A	SMS A	DMO A	ODA B	MNLT B	SMS B	DMO B	ODA C	MNLT C	SMS C	DMO C
460	0.859	0		0	0.891	0.1%		-6.2%	0.78	1.3%		-9.3%
530	0.899	0		0	0.864	0.8%		-6.7%	0.741	2.2%		-9.8%
660	0.902	0	0	0	0.784	0.1%	2.4%	4.3%	0.705	1.2%	2.6%	-1.1%
880	0.846			0	0.856			1.9%	0.755			-3.2%
1300	0.926		0		0.91		0.8%		0.79		1.6%	

Again here, one does not see any discrepancy approaching the 14% stated in the text of the paper.

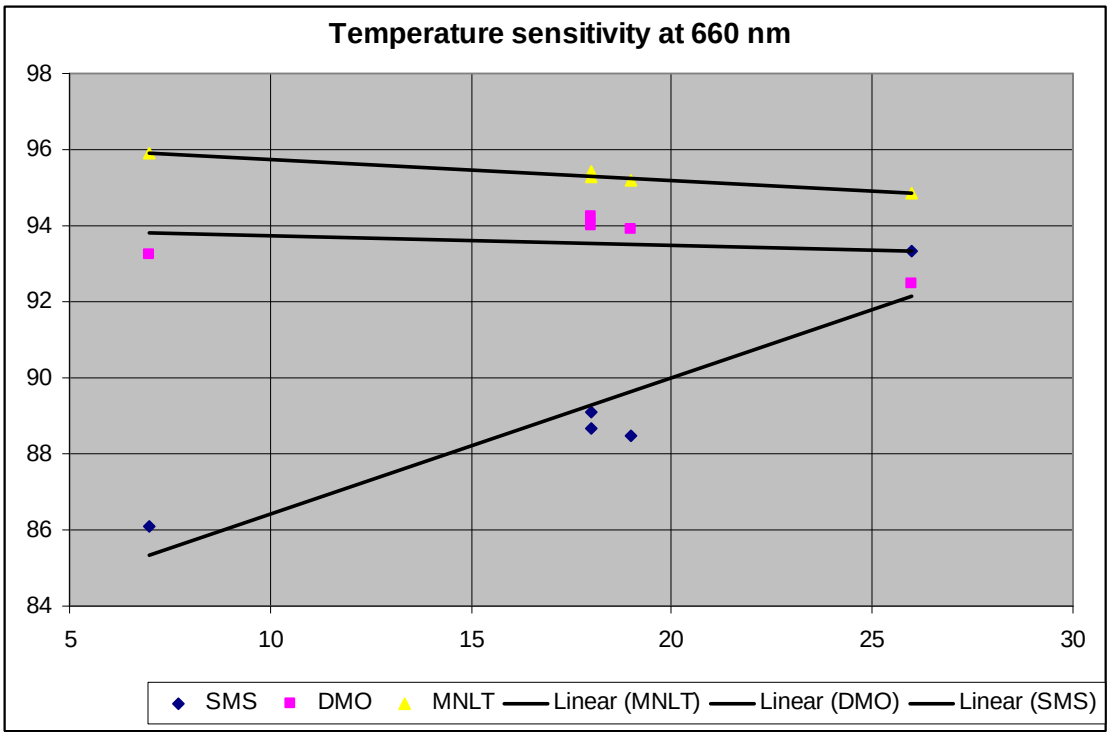
Let us insist on the fact that the significance of all the discrepancies observed in these measurements has little to do with the quality of the tested instruments, but are to be traced to the quality of the samples as we have shown. Then applying a recalibration procedure in such conditions make no much sense, because the spectrum of sample A is very different from the spectrum of sample B and C and telescope mirrors show no spectrum if they are well coated. Recalibration (normalization as the authors call it) corrects the scale of an instrument by measuring a known sample; but when the differences are not due to a scale error as it is the case here, recalibration makes no sense. Imagine that you measure the 'reflectivity' of a ruled grating; you would certainly find strange discrepancies when comparing the results given by two instruments (ODA and any of the x), and it would be still stranger that you attempt to correct the results by a normalization. If an instrument certified by its vendor to measure within 1% yields a discrepancy of 10 or 15%, a scientist would first ask "what did I do wrong?" and if he discovers nothing wrong in his procedure, he would ask the vendor or an expert in metrology what is wrong with this measurement, and not simply write publicly that the instrument makes an error of 10% or 15%[11].

[11] He should also think that his conclusion is most unlikely to be the correct one since it is practically not possible to build such a bad instrument. Such large errors can happen accidentally in case of malfunctioning of an instrument. It was discovered that the IRIS instrument lent to the authors had a problem with the 880 nm channel when used with the battery loader connected. This can be clearly seen in the long term stability tests.

Improper use of Standard deviation (stdv)

In § 3.2.2. of the paper, the stability of the instruments vs. temperature is examined. One measurement was performed at 26C, one at 7C and three interlaced measurements at room temperature.

It is the funniest idea I have ever encountered to analyze these results by computing the standard deviation (stdv) of the five measurements. These deviations are not accidental errors; they are functions S(T) or sensitivity to temperature and have to be analyzed in term of slope dS(T)/dT. This procedure is equivalent to plotting the measured values as a function of temperature and measuring the slope on the plot; a good device will show a zero slope i.e. no sensitivity variation with temperature. This is illustrated in the following graphs for the red wavelength common to the three instruments.



The stdv of the five results is meaningless and shows again that the authors do not know what they are doing; they apply an EXCEL function to a series of numbers and they think they are doing scientific analysis of a problem. The stdv applies to an unbiased population or to independent measurements of the same thing. Here the measurements are biased by the temperature! Imagine that instead of doing one measurement at 7° three at 18° and one at 26° you perform three at7° and at 26° and one at 18° with the same results, say 86.1% at 7°, 88.75° at 18° and 93.3° at 26°; the stdv is 2.6% in the first case and 3.6% in the second case, with the same instrument and the same measurements. Try the series one-five-one and you get σ=2.1% and so on.

Conclusion

This paper has been written in 2005 as scientific arguments in response to a non sense paper presented at a SPIE meeting on telescope maintenance. It has been sent to the authors and remained unanswered. I thought that the errors in this paper were so crude that anyone would put it aside and that I had not to worry anymore about this. But it seems that idiocy is more attractive than intelligence to many, and I am astonished to see that so many fairly intelligent people are still trying to use a gloss meter like the Minolta CM508D or the Surface Optics 410-Vis in order to measure the TIS or, even worse, the specular reflectivity of the coating of their telescope. Of course those instruments yield 'results', that is some figure of measurement that vary in one direction when the mirror seems to get worse; but this is not sufficient for metrology; in fact, it is not metrology at all because the operator does not know exactly what he is measuring. *Any instrument sensitive to light* would yield some result if presented to the mirror. But those results, if they can sometimes be compared to themselves on a short time scale, cannot be compared to other results, either for different telescopes or for the same telescope on a long period of several years. As a consequence it might happen that mirrors are re-coated while they need only to be flushed with dry CO2 or are flushed while they need urgently being coated anew. So the very reason why to monitor the reflectivity and scattering of the mirror is completely missed and it would be as useful to use a searchlight to inspect the coatings, because metrology is exact or is useless.

Until today, there was no portable instrument capable of measuring the specular reflectivity, i.e. the reflectivity excluding all scattered light outside a cone of <1° semi aperture. This is however the main parameter that drives the quality of a coating.

IRIS has now been replaced by a new instrument, Coatest 7L, measuring the specular reflectivity in a cone of 0.8° semi aperture and in seven bands extending from 365 nm to 970 nm. This allows tracking chemical pollution that can deteriorate the reflectivity at certain wavelengths while leaving it untouched at others. This is particularly true at short wavelengths (<500 nm). The new instrument comprises also of a 'dust detector', that is a scatterometer measuring the light scattered between 9° and 19° in the seven bands.