AN EVALUATION OF ERRORS IN REFLECTION DENSITOMETRY

R. O. Colestock*

Abstract: The use of reflection densitometers is widespread in the Graphic Arts industry. These instruments are relatively inexpensive and simple to operate. They produce numbers which represent the thickness of ink on a substrate or a derived function such as dot gain. These and other related parameters determine the final acceptability of a printed job. Thus, judgments about the acceptability of a printed job are often made on the basis of densitometer numbers. Despite their widespread use and importance, relatively few studies have been made of the factors affecting the accuracy and precision of densitometers. (2),(3),(7),(8),(10),(12)

Densitometers, like other instruments, are subject to errors which fall into a few principle categories. The effect of these errors can become critical when the quality level of printing is near the borderline between acceptable and rejectable. Control of error sources is also important when information is obtained from, and transferred among different locations.

This paper examines common sources of errors in reflection densitometry. The results of a staggered nested designed experiment are discussed and recommendations are made for minimizing the effects of these errors.

Definitions

A clarification of terms is important to this discussion of errors. We will use the following terms in the sense advocated in the "Quality Control Handbook".(6)

Accuracy is used to describe close proximity to a "true" or target value. In reflection densitometry, accuracy is largely determined by the calibration plaque and the measurement procedure.

Precision has several meanings which relate to the ability of an instrument to produce similar results over smaller and larger periods of time.

Repeatability refers to measurements made over short periods of time with a single instrument. We will use it to describe successive measurements and to imply that no recalibration of the instrument has occurred during the measurement period.

* 3M Co., St. Paul, MN

Reproducibility refers to measurements made over longer periods of time which may include re-calibrations and/or multiple instruments.

Linearity refers to an instrument response that is proportional to input signals throughout its operating range. In densitometry, precision and linearity are largely determined by the instrument.

Absolute measurements are those which purport to give a 'true' or absolute value of a sample. A T-Ref is an example of a reference with an absolute measurement. (11) A value is assigned to a particular spot which is accepted as the true value of that spot. Accuracy and precision are primary concerns in absolute measurements.

Relative measurements are those which express a "larger than" or "smaller than" relationship between a sample and a reference material A SWOP HiLo Reference is an example of a relative measurement reference. (9) Sample densities are only required to be "in between" the densities of two reference spots. The magnitude of the numbers does not affect the decision, so calibration of the instrument is not required. Thus, accuracy is of little importance and a moderate level of precision is usually sufficient to provide a useful result.

A *density unit* is defined as 0.01 reflected optical density.

Sources of errors in densitometry.

Densitometer errors can be classified in three principal categories according to their source: plaques, instruments and procedures.

1. Plaques

Calibration plaques are intended to provide true values against which other samples can be compared. There is, at least, an implication that 1) the plaques themselves are accurately calibrated and 2) that different plaques will give the same results - i.e., that the members of a set of plaques are consistent. The calibration procedure and standards used by a plaque vendor should be traceable to a Primary Standard maintained at a national standards organization. In addition to investigating the accuracy and precision (consistency) of a set of plaques, we also checked them for mechanical damage, color uniformity and fading.

The set of plaques consisted of eight individuals purchased from a commercial densitometer vendor. The group had two distinctly different sets of serial numbers which are indicative of different production lots. Each plaque consisted of a paper base with six target areas: white, gray, black, cyan, magenta and yellow. The entire plaque was protected by a clear plastic layer.

A. Surface Blemishes

We first examined the plaques with a low power microscope. Most of the colored spots appeared to be uniform and free of defects. We did observe a small crack and several pits which may have originated from air bubbles

during application of the colored material. On several spots, we detected small specks of black material which could not be wiped away. In later testing, we observed that repeat measurements made on color spots with blemishes tended to have slightly higher variability. We believe the higher variability occurred from random inclusion or exclusion of the blemish due to different positioning of the instrument aperture.

B. Non-uniform spots

Color spots on plaques are slightly larger than the aperture of the densitometer. This allows small positioning errors to be made without altering the measured density. The assumption is made that the colored spot has the same density at all points on its surface. Experience has taught us that there is no such thing as a uniform printed surface so we examined one plaque for density uniformity.

We measured four locations on each spot. The measured areas were ninety degrees apart and as close as possible to the edge of the spot. (Fig. 1)

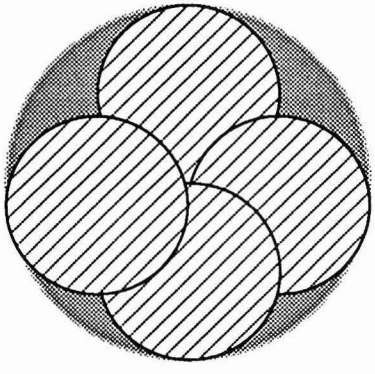

Fig. 1. Locations of aperture on color spot.

Ten measurements were made on each of the four locations and the maximum difference among the locations calculated. The total ranges within color spots were small - the largest being 0.004 density units for Yellow. This indicates that, for a sample of one, uniformity of the colored spot is not a major source of error.
Note: Multiple measurements allow errors to be expressed at greater precision than the stated value for the instrument.

C. Fading

Experience has also taught us that there is no such thing as a permanent ink color. Our plaques were stored in paper envelopes in filing cabinets at ambient lab conditions. We have observed the densities of these plaques for approximately eighteen months without significant changes. Our procedure will detect changes of 0.005 density unit.

D. Accuracy

Determination of plaque accuracy is difficult and time consuming. One procedure consists of measuring the colored spots with a 45°/0° spectrophotometer and calculating the filter densities from spectral data. This requires that the spectrophotometer be accurately calibrated and that statistically significant data be carefully gathered.

We checked the accuracy of the plaque calibration values by measuring one plaque with a 45°/0° reflection spectrophotometer (Gretag SPM-100). The reflectance values were used to calculate the Status T densities which were then compared to the absolute values assigned by the vendor. The absolute values calculated by our test were different from the assigned values by the following amounts: 0.04 density units lower for the cyan and magenta spots, 0.03 units higher for the yellow spot and 0.06 units lower for the black spot. These differences are not unexpected since our spectrophotometer is different from the one used to calibrate the plaque.

Our purpose is not to suggest that our calibration values are more accurate than the vendor's values. We wish the reader to understand that accurate calibrations require complex absolute measurements. There remains significant potential for error between different calibration plaques.

E. Precision

Precision can be thought of as consistency of calibration among a group of plaques. We began by hypothesizing that all of the plaques were consistent with each other. If this were true, we should be able to calibrate on any one of the plaques and then measure any other plaque with close agreement to the calibration values on that plaque. When we performed this experiment, we found close agreement among the eight plaques. The average difference among the eight plaques was 0.005 density unit and the maximum was 0.02 for two black spots. The effects of a 0.02 plaque error will be considered in the Discussion section. (Table I)

Table I. Calibration Error of Eight Densitometer Plaques.

PLAQUE SER. NO.	Yellow Delta	Magenta Delta	Cyan Delta	Black Delta	Gray Delta	White Delta
5586	This plaque was used to calibrate the densitometer.					
5585	0.00	0.00	0.00	-0.01	-0.02	0.00
3731	-0.01	-0.01	-0.01	-0.02	0.00	0.00
3726	0.00	-0.01	-0.01	-0.02	0.00	0.01
3730	-0.01	0.00	0.00	0.00	0.00	0.00
3732	0.00	0.00	0.00	0.00	0.00	0.01
3646	0.00	-0.01	-0.01	-0.01	-0.01	0.00
3725	0.00	-0.01	0.00	0.00	0.00	0.01
AVERAGE	-0.003	-0.006	-0.004	-0.009	-0.004	0.004

2. Instruments

This category includes a study of the relative importance of factors affecting instrument precision. We also measured the difference between two identical densitometers from the same vendor. The experiment was repeated with two identical densitometers from a different vendor.

We have also included in this category two other error sources: 1) mismatches between sample spectra and densitometer filters and 2) errors due to data handling. Mismatch errors can occur because a Status T densitometer is not intended to provide consistent numbers for various hues of ink. The wavelength of the filters is fixed, so the reflection minima of the ink spectra will not always be ideally located within the filter window. The instrument will continue to provide numbers, however the relationship of the numbers to visual appearance will no longer be consistent.

Round-off errors occur when an instrument measures and calculates to three or more significant places internally but only reports two significant places. This effect is small in everyday applications but can be a measurable source of error in high precision work.

A. Staggered Nested Designed Experiment

Errors inherent in the instruments were investigated by using a staggered, nested designed experiment. In this experiment, we examined the relative importance of seven factors which are significant in reflection densitometry. These were 1) simple repeatability of the instrument in a series of successive measurements, 2) errors in positioning an instrument on a target area, 3) the error arising from re-calibration of the instrument, 4) the 'agreement' of instruments of the same type and brand, 5) the 'agreement' of instruments of the same type from different manufacturers and 6) the effects of different sample colors. (Fig. 2)

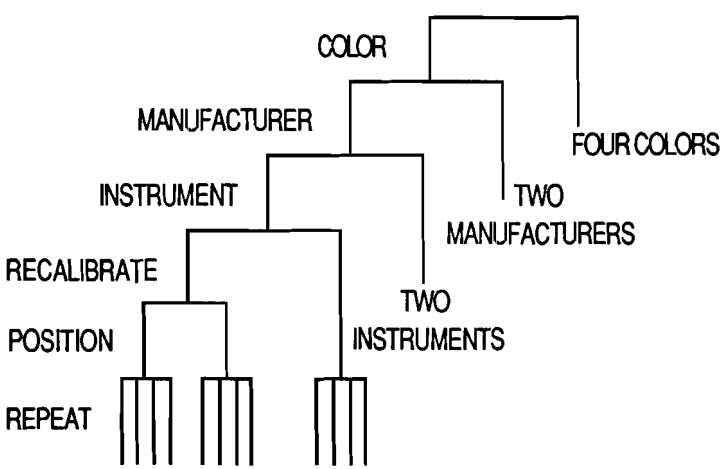

Fig. 2. Schematic of Staggered Nested Design

In this experiment, we initially chose to use BCRA ceramic tiles because of their durability. We were concerned that multiple measurements on the ink-on-paper plaques might cause damage which would inject false errors in the experiment. When we analyzed the data, the errors were considerably higher than expected. We believe this to have been caused by a combination of sample gloss and of mismatches between the tile spectra and the Status T filter windows. We later repeated the experiment several times with ink-on-paper plaques with no sign of damage.

The experiment consisted of positioning a densitometer on a color spot of a T-Ref and making four successive measurements ten seconds apart. The instrument was not moved between these measurements. This provided an estimate of short-term precision (repeatability). Next, the unit was lifted off the spot and repositioned, visually, in the same location. Four more measurements were made to provide an estimate of positioning error. The instrument was then re-calibrated, repositioned on the spot and four more measurements made. After correcting for positioning error, this provided an estimate of errors caused by calibration.

The entire nest was repeated using a second unit of the same type from the same manufacturer. This provided an estimate of instrument-to-instrument variability. We then repeated the entire nest with a pair of Status T densitometers from a different manufacturer. This provided an estimate of the variation between different manufacturers of the same type of instrument. Finally, each instrument was used to measure the other three plaque colors to provide an estimate of the consistency of the filters. The analysis of the nested design is shown in Tables IIA-D. This analysis identifies the percentage of the total variance that can be assigned to each of the factors in the design. The individual and total standard deviations are also given.

TABLE II-A. IMPORTANCE OF ERROR SOURCES - PROCESS YELLOW

SOURCE	VENDOR A		VENDOR B	
	STD. DEV.	%	STD. DEV	%
INSTRUMENT	0.0000	0.0	0.0011	13.5
CALIBRATION	0.0008	69.8	0.0000	0.0
POSITION	0.0004	15.1	0.0000	0.0
REPEAT	0.0004	15.1	0.0029	86.5
Total Std. Dev.	0.0009		0.0031	

TABLE II-B. IMPORTANCE OF ERROR SOURCES - PROCESS MAGENTA

SOURCE	VENDOR A		VENDOR B	
	STD. DEV.	%	STD. DEV	%
INSTRUMENT	0.0001	4.6	0.0005	5.3
CALIBRATION	0.0004	41.6	0.0000	0.0
POSITION	0.0000	0.0	0.0007	10.5
REPEAT	0.0005	53.8	0.0020	84.2
Total Std. Dev.	0.0006		0.0022	

TABLE II-C. IMPORTANCE OF ERROR SOURCES - PROCESS CYAN

SOURCE	VENDOR A		VENDOR B	
	STD. DEV.	%	STD. DEV	%
INSTRUMENT	0.0000	0.0	0.0011	13.5
CALIBRATION	0.0008	69.7	0.0000	0.0
POSITION	0.0000	0.0	0.0000	0.0
REPEAT	0.0005	30.3	0.0029	86.5
Total Std. Dev.	0.0010		0.0031	

TABLE II-D. IMPORTANCE OF ERROR SOURCES - PROCESS BLACK

SOURCE	VENDOR A		VENDOR B	
	STD. DEV.	%	STD. DEV	%
INSTRUMENT	0.0000	0.0	0.0010	9.1
CALIBRATION	0.0016	94.9	0.0000	0.0
POSITION	0.0000	0.0	0.0022	42.4
REPEAT	0.0004	5.1	0.0024	48.5
Total Std. Dev.	0.0016		0.0034	

Variability among instruments of the same type and manufacturer is the sum of small differences in the lamps, physical filters, detectors and compensating systems. These combined errors are expressed in the specifications for densitometers - typically "+/- 0.01" within instruments and "+/- 0.02" between instruments. Vendor literature usually does not specify whether these limits are for 95% of the population (+/- two standard deviations) or for the stricter 99% (+/- three standard deviations). (Fig. 3) Our interpretation will be made with the stricter three-sigma limits.

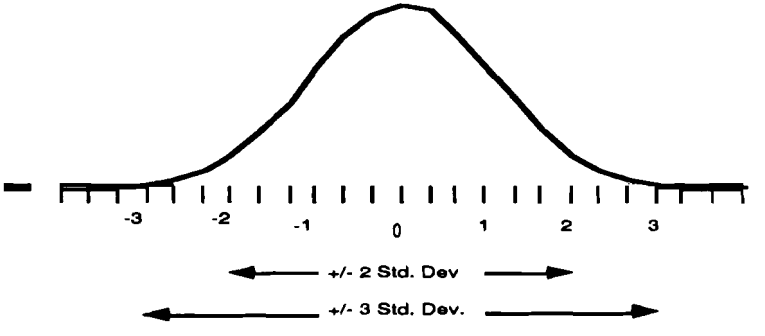

Fig. 3. Two and Three Standard Deviation Limits

We make no claim that the factors included in our evaluation are the same as those used by the manufacturers in establishing their specification limits. Typical usage does include repositioning and recalibration the instruments and would tend to enlarge the error due solely to repeat measurements.

When the standard deviation due to REPEAT is multiplied by six (+/- 3 S.D.) we obtain an estimate of each instrument's internal consistency. The data from vendor A instruments had three-sigma ranges from 0.0024 to 0.0030 for the four process colors. The data from the vendor B instruments had three-sigma ranges from 0.0120 to 0.0174 density units. Therefore, the instruments from both vendors met the manufacturer's specification of +/- 0.01 for repeatability.

It is unlikely that the manufacturer's specification contains the errors due to POSITION and CALIBRATION, however, an inspection of Table IIA-D shows that substantial fractions of the total variation came from these two sources. Metrology programs should make provision for these errors.

Tables IIA-D also show that the percentage of the total variation due to instrument-to-instrument error was zero for the two units from vendor A and approximately ten percent for the units from vendor B. When the three- sigma error from all error sources is considered, the instruments from vendor A ranged from 0.0036 to 0.0096. The instruments from vendor B ranged from 0.0142 to 0.0204. Therefore, the instruments from both vendors met the manufacturer's specification of +/- 0.02. The next comparison in this study is between the two vendors. Our analysis showed that the instruments of Vendor A and vendor B had very close agreement in absolute numbers when calibrated to the same plaque. The instruments from vendor A had significantly less variation than those of Vendor B. We chose to configure the instruments from vendor A to display three decimal place data because the two decimal place data often showed zero variation. The final error source was sample color. All four instruments showed the smallest variation when measuring magenta and the largest when measuring black.

B. Filter and ink mismatches

Filter sets in densitometers are selected to measure a relatively narrow range of printing ink spectra. This means that the spectral 'window' of each filter in the densitometer is centered at or near the minimum reflectance of the spectral curve of the ink. Changes in either the width of the filter window or the spectral curve of the sample can result in unwanted reflection occurring within the filter window. Fisch (4) and Huntsman (5) report that this effect may be large enough to cause measured densities to be inversely proportional to the visual perception and colorimetric correlates (C^*) of color intensity.

C. Data display Errors

Some densitometers can be set to display three decimal places. We measured the T-Ref twice - once at two decimal places and once at three decimal places using the same instrument from vendor A. The two-place data showed no variation in the twelve successive readings of four process colors. The three place data typically showed a range of +/- 0.001 to +/- 0.002 depending on the sample color. While three-place data are useful in high precision work, two-place data are sufficient for most applications. Round-off error is usually not a significant source of error.

3.Procedures

A functional, well-calibrated densitometer can still be used in ways which provide erroneous data.

The densitometer geometry requires that the sample be essentially flat and in close contact with the foot of the instrument. Measurements made with an instrument that is slightly tilted will be significantly different from those made with a correctly placed unit.

Show-through is another procedural issue that can cause significant errors. ISO/ANSI standard 5.4 requires that samples to be placed on a matte black material having a density of 1.5 or greater. This eliminates the error due to show-through of differently colored work surfaces and white back-up sheets of different color.

Small, but measurable, errors can arise from warm-up effects. Twenty-five successive measurements made with a 'cold' instrument can show a difference of up to one density unit (0.01) between the first and last value. The effect does not occur equally in all instruments.

Errors of several density units can arise from measurement of materials which polarize light. Colored materials on tensilized polyester substrates may cause density measurements to vary with the circular orientation of the instrument - i.e., north-south versus east-west. The solution to this problem is to make measurements at several locations on the sample and average the results.

Discussion

Errors in densitometry arise from the calibration plaques, the instruments and the procedures used to make the measurements. In the simplest case of a single instrument being used to make relative measurements, these errors may have minimal impact on decisions to accept or reject. See Fig. 4. When sample #1 is centered in the acceptance range, the +/- 0.01 error of the instrument is small when compared to the +/- 0.07 of the SWOP HiLo Reference. When sample #2 is near a limit, the uncertainties combine to increase the risk of 'wrong' decisions. The assumption is made that the HiLo Reference has been accurately calibrated. If the Reference is inaccurate, even this simplest of measurements cannot produce useful results.

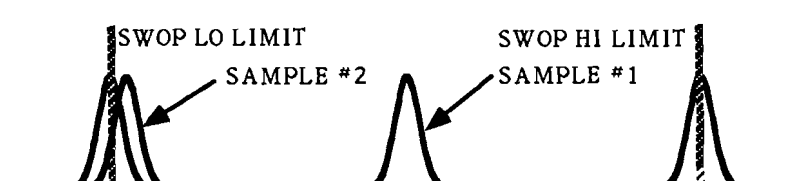

Fig. 4. Densitometer Error in Proportion to HiLo Reference

In the case of a single instrument used to establish conformance of a sample to an absolute specification, the sum of all the errors can easily cause a wrong decision to be made. In this example, instrument, plaque and procedural errors are combined to give a total error of 0.03. See Fig. 5. To circumvent this uncertainty in the true sample value, every large number of samples must be tested and/or increased risks of a wrong decision must be accepted. Note that the 'wrong' decision can be of either type - to reject good material or to accept bad material. Both decisions will cost time and money.

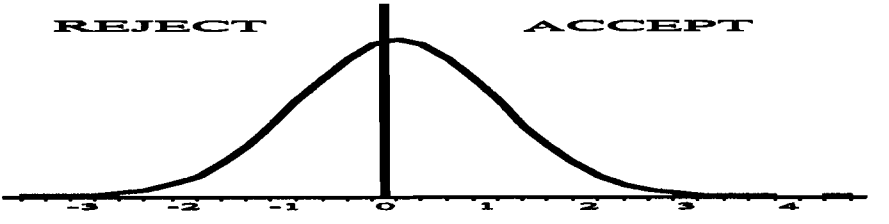

Fig. 5. Effect of Uncertainty in Vicinity of Specification Limit

In the more extreme, but increasingly common, case of a group of instruments which are expected to agree with each other, it is safe to say that agreement cannot be obtained nor maintained without meticulous control of plaque, instrument and procedural factors. See Fig. 6 In this example, two instruments with individual errors of 0.01, are calibrated on separate plaques that are 0.02 apart. The acceptance limit falls midway between the values measured by the instruments. Which instrument is giving the "right" value?

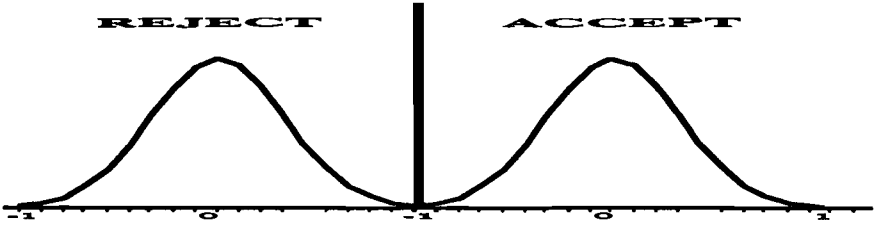

Fig. 6. Effects of Uncertainty in Two Densitometers at a Limit.

Conclusions

In this study, the group of eight plaques was found to be consistently calibrated - that is, the individuals agreed closely with each other. We recalibrated the absolute values of the plaques with a 45°/0° spectrophotometer and found differences of .03 to .05 density units from the assigned values. We assert no claim for the superiority of our calibration. Color uniformity of the spots and resistance to fading were good. A small number of colored spots showed blemishes of different types which appeared to have small effects on measurement precision.

The precision for simple repeat measurements was within the manufacturer's specification of +/- 0.01 density unit for all four instruments. The instruments from vendor A were well within the spec limits while those of vendor B were just within the limit. These differences would be apparent only in high precision work.

Positioning error tended to be small (< 15% for both instruments). Recalibration, however, consistently introduced additional error in instruments from vendor A while instruments from vendor B were unchanged by recalibration.

When the densitometers were calibrated to the same T-Ref, the averages of the measurements were essentially identical. This indicates that the Status T response is serving its purpose. It should be noted that the agreement of Status T instruments from different vendors might be reduced when measuring non-process colors which are not as well centered in the filter windows.

Procedural errors can be large or small. This type of error is within the control of the individual user to a larger degree than errors in plaques and instruments. Relative measurements with single instruments minimize the effects of plaque and instrument errors. Absolute measurements increase the risk of wrong decisions. Measurements made with a group of instruments require careful, persistent control of errors in all three categories.

Recommendations

For relative measurements with single instruments:

1. Use black backup material
2. Keep instrument flat and in contact with sample.
3. Use fresh SWOP HiLo reference.

For absolute measurements with single instruments add:

3. Use fresh, undamaged calibration plaques.
4. Calibrate only when instrument drifts outside limits.
5. Use control charts to monitor process for trends and excursions.

For absolute measurement with a group of instruments add:

6. Select or re-calibrate plaques to maximize plaque agreement.
7. Use current instruments of the same type and manufacturer.
8. Use round-robin tests to establish agreement between sites.

Acknowledgments

The author gratefully acknowledges the careful work of Anne Krause who made the measurements and was responsible for data handling. The suggestions of Dr. James R. Huntsman are also gratefully acknowledged.

Literature Cited

1) Bartels, S.A., Fisch, R.S., Henderson, T.A.
1990 A Study Involving the Use of Densitometry and Colorimetry for Process Control of Color Inks and Graphic Arts Color Printing, J. Imaging Technology, Vol.17, No. 1, pp 9-17.

2) Bowden, D.M.,
1989 "Narrow-band vs Wide-band Densitometers: An Update", TAGA Proc.,pp 1-10.

3) Brehm, P.V.
1990 "Introduction to Densitometry", Revised Edition, Graphic Communications Association, 62 pp.

4) Fisch, R.S., Colestock, R.O.
1990 "A Contribution Towards a Standard Metrology for the Colorimetric Specification of Printing Ink Color", TAGA Proc., pp 497-515

5) Huntsman, J.R.
1990 "Colorimetric Analysis Methodology for Graphic Arts", SPSE 6th Int. Cong. on Advances in Non-Impact Printing.

6) Juran, J.M.
1975 "Handbook of Quality", Third Edition, McGraw Hill

7) Rehm, L.H.
1965 "Densitometer Standardization in the Gravure Industry", TAGA Proc., pp235-244.

8) Southworth, M.
1983 "Quality Control Scanner", Graphic Arts Publishing Co., Vol. 1, No. 2

9) SWOP, Recommended Specifications Web Offset Printing, 1988 Ed., International Prepress Assn., So. Holland, IL

10) Vogelsong, W.F., Swartz, C.B.
1975 "Specification of Variables for Proper Control of Densitometers", J. App. Photographic Eng., 1, pp 47-53.

11) Vogelsong, W.F.
1989 "Standard Reference Materials for Densitometry", TAGA Proc., pp 132-141.

12) Vogelsong, W.F.
1990 "Capability Studies of Densitometers and Densitometry", TAGA Proc., pp 1-10.