Color Management—How Accurate Need it Be?

G. Gonzalez, T. Hecht, A. Ritzer, A. Paul, J.-F. Le Nest* and M. Has*
*FOGRA, Munich, Germany; * INPG-EFPG, Grenoble, France*

Abstract

Color Management aims communication of colors assumes it can widely correct systematic errors introduced in prepress. Color measuring devices and the errors introduced in the production process are analyzed to gain information on their systematic/random behaviour. Origin, size and parameters influencing random errors introduced are discussed in terms of possible ways to suppress them and to understand error propagation in the production process. Noise on the printing signal, intentionally addressable color levels and errors due to image data compression are highlighted.

1. Introduction

Color management systems and their work have been described frequently and in great depth. A literature overview may be found in (1). Color management aims to be applied to devices located and operated in environments with non constant configurations—as they may not be calibrated as a whole configuration e.g. due to often occurring changes. The aim of color management is to avoid the characteristic systematic color reproduction ermrs which the devices under view perform.

However, color differences between originals and their reproductions may not only occure due to the different systematic behaviors but also due to stochastic errors on the signals as they are Hochred by the devices. Those errors are unavoidble and may have their origin in the physics of the materials involved[2,3] or in the characeristics of softwares.

In the past for users of color reproduction systems the question of color accuracy of the individual device or of software has been overlayed by the far more significant problem to communicate color doubtlessly. Now, since color communication works rather accurate with color management systems, color accuracy of the devices and softwares need to be evaluated.

Here we first address the accuracy of color measurement devices as they are in use in the field, then the accuracy of scanners and printers is discussed.

As there are some characteristics in the noise on the color signals of offset presses this aspect is highlighted in a separate chapter.

Ahead of comparing the results on noise in device data, error propagation in the publication process is discussed.

2. Color Measurement Devices

Color measuring devices perform different with regard to several aspects like heat stability, performance of aging sensors etc.. In the following we concentrate on the reproducibility of the measured data and the accuracy of the results obtained by the devices.

In our approach we evaluated devices of a quality as they may usually be found in professional reproduction and printing environments. A further criteria was whether the devices under view would permit to obtain results according to CIE standards[4] and the ICC specification.[5] The accuracy and fidelity of three devices have been tested: the Gretag SPM100-II, the X-Rite 938 and the Digital Swatch Book from X-Rite.

The devices are exact to the degree that they obtain numbers as small as hundreth of numbers represented in CIEXYZ-Coordinates. When computing spectral values in CIELAB color space coordinates they refer to different standards.[6,7] Thus values obtained from the tested color measurement devices were first collected in CIEXYZ coordinates and then converted in CIELAB using DIN reference values (X_0, Y_0, Z_0) for D50 conditions.

For better matching with perceptions of human observers all color differences calculated are given in ΔE^*_{94}.

2.1. Reproductibility of Color Measurement Devices

In order to evaluate the reproducibility of color measurement devices we did measure in different time sequences: the same target has been measured in six measuring cycles with 50 measurements per patch in a row with intermediate breaks of 24 hours and in three measuring cycles with 50 mesurements per patch in a row after 72 hours of use. In order to avoid problems due to warming up of the measuring devices under view, the devices remained to be switched on.

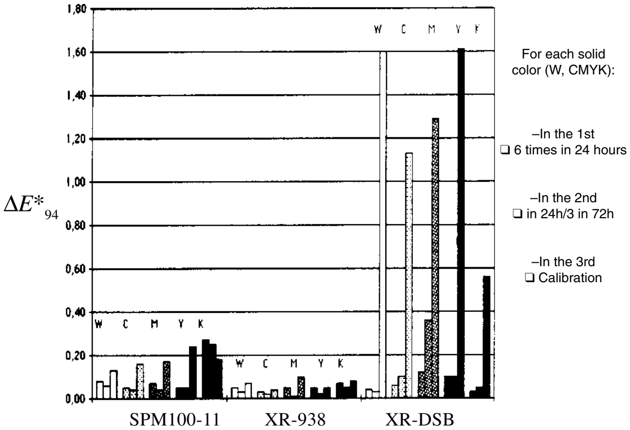

Figure 1. Reproducibility of color measurement. The different columns represent white reference and the primaries.

In addition we measured the influence of a renewed calibration on the reproducibility of the device. The results are displayed in Figure 1.

As Figure 1 indicates the effect of a renewed calibration turns out to be the most crucial parameter having an influence on the accuracy of the measured data. Relative to this number the other variances are small - that is in the range between $\Delta E^*_{94} = 0.1$ and $\Delta E^*_{94} = 0.3$.

One should point out that the white references as they were used in the experiments, differ from device to device since the manufacurers of some of the measuring units provide the user with device specific targets (made of different materials). The exactness of the calibration depends on the sample and the range of wavelength taken into account by the devices.

There has been some indication that the most inaccurate measurements occur in the very light and very dark patches.

2.2. Fidelity of Color Measurement Devices

In order to acquire information on the fidelity of color measurement devices we used a printed sample of the KODAK 1TS.7/2 as target. Three independent measurements have been obtained in all color patches defined by ISO and, in addition the CMYK and RGB values have been measured.

The results obtained may be structured as follows:

1. The measurement results as they are derived with different devices differ significantly as can be seen when comparing Figure 2 and Figure 3.
2. The differences of the measurment results become larger as the saturation of the colors under view increases.
3. The inaccuracy of magenta turns out to be the largest.

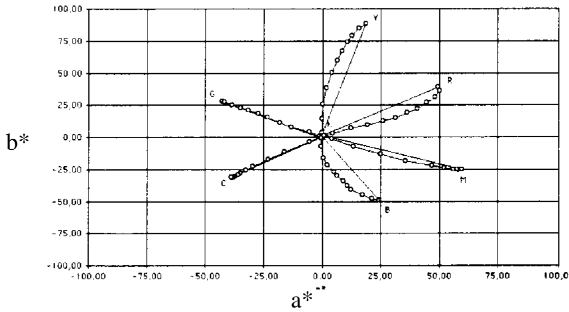

Figure 2. Gamut obtained with Gretag SPM100-II using the Kodak IT8.7/2 target fixed color patches as well as CMYK and RGB patches.

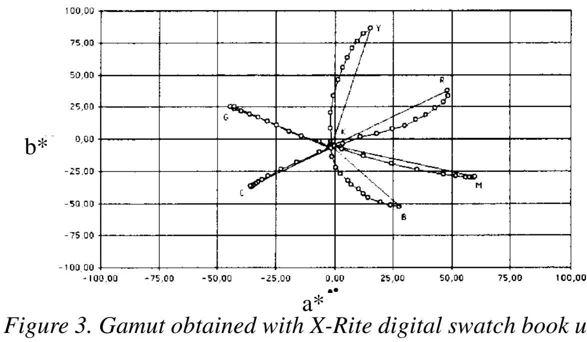

Figure 3. Gamut obtained with X-Rite digital swatch book using the Kodak IT8.7/2 target fixed color patches as well as CMYK and RGB patches.

3. Reproducibility of Color Signals of Devices in Color Reproduction

3.1. Scanners

The accuracy and fidelity of two flatbed scanners were tested: the Saphir (by Linotype-Hell) office scanner and the Topaz scanner (also by Linotype-Hell), which aims for professionaljobs as target groups. These two were chosen since they cover both the low and the medium price market segments and are equipped with scanner-adapted software.

The experiments were performed in the following way: the target (1T8.7/1 and 1T8.7/2 Kodak targets) has been fixed on the scanner and scanned 10 times under standard options and without any scanner profile. The transformation from RGB (of the CCD chip) to Lab is then performed by the scanner software, using an internal "generic" calibration.

The resulting data were saved as TIFFLAB files in LinoColor and analysed using a software tool. This software tool is a home made C program running on UNIX platform.

As for the tested color mesurement devices, the results are presented in ΔE^*_{94} for the IT8.7/2 target with an accuracy of only one thenth of CIELAB coordinates. This accuracy has been chosen to take into account that the data obtained in one scan vary from pixel to pixel within the same patch scanned. We were unable to learn wether this variance occurs due to the scanner or to the target. The limits of the accuracy of the data collected by the scanner ranged within the boundaries of tenth of CIELAB coordinates.

ΔE^*_{94}	mean	max.	min.
Saphir	0.2	0.9	0.0
Topaz	0.2	1.0	0.1

As the results indicate, both scanners have an alike accuracy. Nevertheless, when looking at the color gamut obtained from the ITS.7/2 taget, the white and black point of the Topaz scanner are slightly better centered in an (a*,b*) representation than the data generated with the Saphir Scanner we used. This results into a better fidelity in the color dataset acquired with the Topaz.

3.3. Printers

In order to gather insight into the reproducibility of color signals of printers the following experiments have been performed with several presses and printers. The ability of three printing processes (sheet-feed offset, gravure and non impact-Xeikon) to reproduct 8 bit colorlevels in CMYK as well as in secondary and tertiary colors (C+Y, C+M, M+Y and C+M+Y) was evaluated.

The test target we used contained 256 patches with increasing color level of the primary, secondary and tertiary color under view. The field contained patches which have been placed in a square of 16 by 16 fields with different colors. A sketch of the target it has been printed with the different devices is displayed in Figure 3.

Here the results for sheet-feed offset with AM and FM screening and paper class 1 as well as for non impact-Xeikon with paper class 2 (as specified in (8)) an: presented. As Figure 6 indicates, none of the printing processes tested were able to reach the 256 color levels contained in the initial digital document.

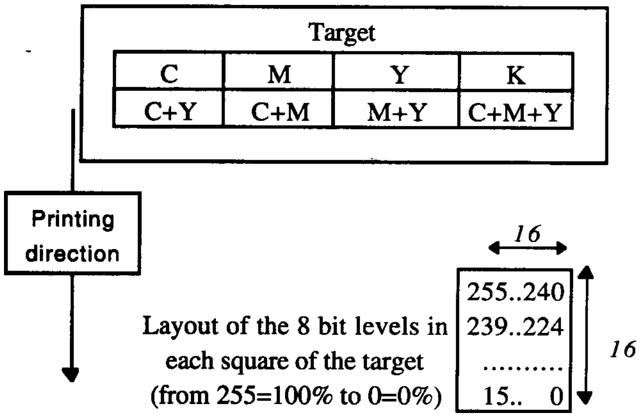

Figure 4. Organisation ofthe test target

Parameters influencing those results are numerous: screening (e.g. screening method, dot shape and size), properties of paper involved (e.g. gloss, whiteness) and of course the way the printing process runs.

The coordinates of the colors measured in the patches do not develope linearily in the CIE space. In order to evaluate the real length in ΔE^*_{94} from level 0 (color at 100%) to 255 we used a step of 5 levels which estimates the shape of the curve rather well. Due to artefacts in the color levels, steps smaller than 5 could not be choosen.

We used the ΔE^*_{94} calculation to take the difficulties and limits of the human observers to differenciate colors in some parts of the CIELAB color space into account.

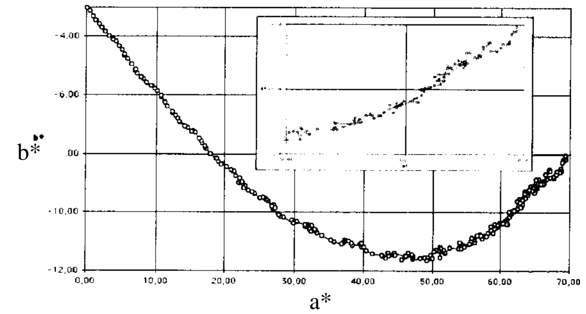

Figure 5. Adressed 8 bit Magenta levels in sheet-feed offset with AM screening

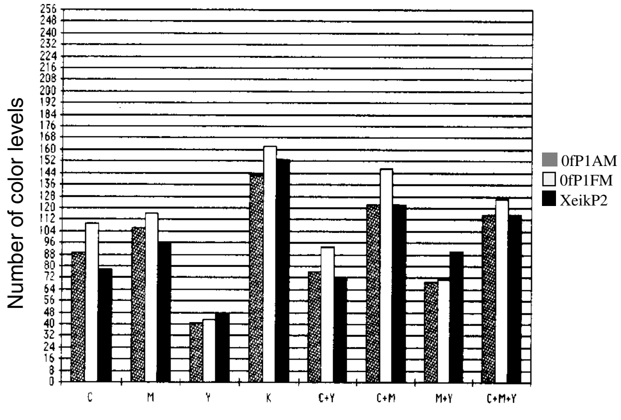

Figure 6. Number of colorlevels obtained from adressed 8 bit levels in C, M Y, K, C+Y, C+M, M+Y and C+M+Y

Figure 5 displays the number of color levels meaaned on the patches on one sheet under view. As can be seen the curve is not linear and—as displayed in the enlarged segment of the curve—the measurements do not follow each other in a clear direction—meaning one color level may e.g. be darker than the following one despite its encryption in the data set would indicate a darker point. This interpretation also holds in the three dimensional representation of the measurement results.

This effect leads to the insight that, despite a large number of color levels may physically be printed on a printer or a press, only a limited number within a certain accuracy may be *intentionally* reproduced. Figure 6 displayes the number of color levels as they may intentionally be reproduced in this sense by the different devices (see tables in appendix). As can be seen the numbers of color levels displayed in the primary colors ranges around 200 depending on the screening method and the paper under view. The overprints of different primaries as it leads to secondary and tertiary colors gain a number of around 80 to 100 color levels which may intentionally be addressed.

However there is also a further source of inaccuracy on the color signal produced by a press—the time dependent noise.

Figure 7 displays the density measured in solid tones on sheets as it developes after the ink flow has been changed in a given zone in an offset press. It can be seen that the signal is not constant but varies significantly. Analyzing the data it can be seen that the noise is of random character.[9]

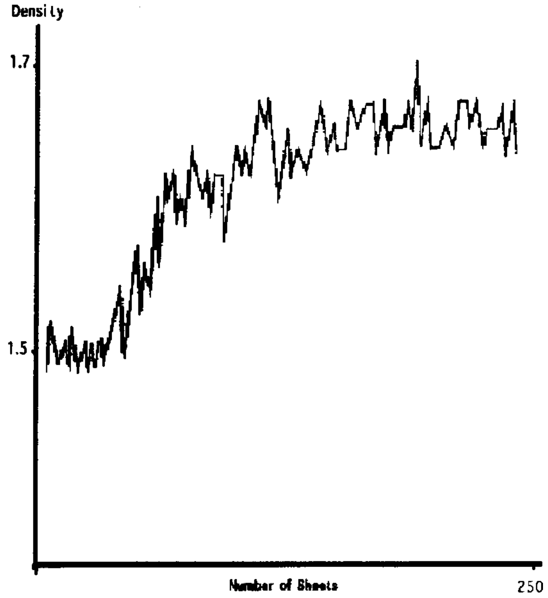

Figure 7. Fluctuation of density on sheets printed one after the other

Analyzing the size of the noise in solid tones published data indicate (10) led to the conclusion that the accuracy depends on the printing method under view and that tolerances as they should be permitted for solid tones should be rather high:

Method	Tolerance*
Rotogravure	$\Delta E^*_{ab} = 3.0$
Offset	$\Delta E^*_{ab} = 5.5$

(Tolerence id defined in the sense that *90% of all colors measured on patches are within the range given).

In order to varify those numbers we measured the color signal dependent from color level and time—the results are given in Figure 8. As can be seen in this Figure the time dependence is less developed as the ability of the device to reproduce intended values for certain grey levels.

Figure 8 displays two sets of information.

Figure 8(a) displays the difference between two measuring results of immediately following patches and Figure 8(b) the noise on the time dependent color signal as it occurs in the individual color levels.

It can be seen that the time dependent noise is significantly slower than the differences between following points.

Thus both effects, the time dependent noise on the color signal and the unability of the device to intentionally display color levels limit the number of colors that may intentionally be addressed with printers and presses which may have caused the literature data to be higher than the values we obtain.

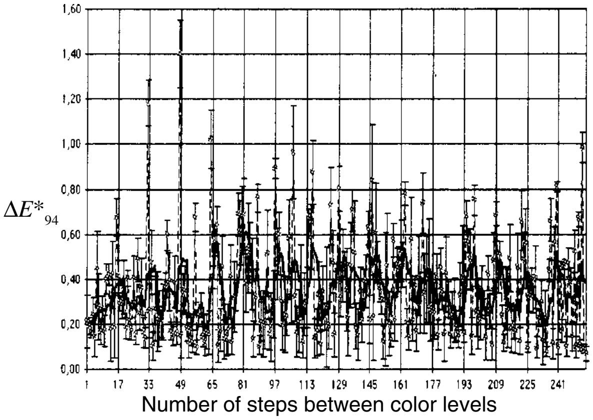

Figure 8a. ΔE^*_{94} between one color patch and the next as well as standard deviation for the measurement displayed for the color level under view

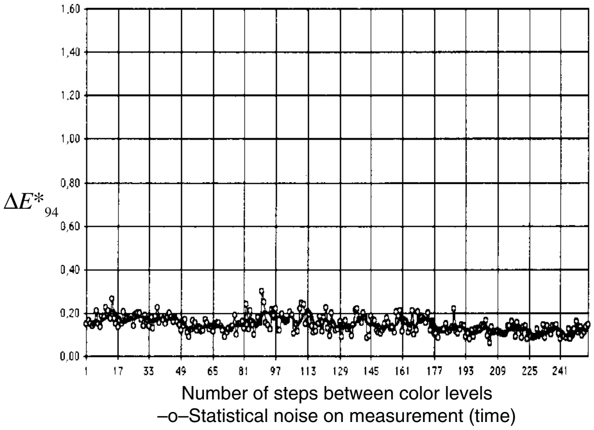

Figure 8b. Color level versa time related noise on the signal

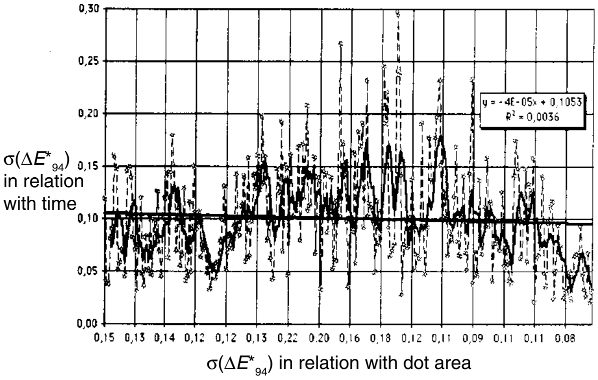

Figure 8c. Independence of the noise relating with dot area and the noise related with time

4. On the Fluctuations of the Color Signals of Offset Presses

Color fluctuations in the primary colors have effects of varying strength in different areas of the color space. In the case of mixed colors, the fluctuations in the individual primary colors can combine to form an overall fluctuation. In areas where there are low dot percentages, fluctuations in layer thicknesses have consequences that differ from what occurs in solid tones (see tables in appendix). The accuracy with which colors can be reproduced varies depending on the particular area of the color space therefore.

It is particularly in the midtone range that the secondary colors have a decisive effect on the appearance of images. This is why fluctuations in the secondary colors are of great significance to image reproduction.[11]

In experimental studies, it turned out that the shapes, sizes and alignments of the ellipsoids in the color space—they are the factors which characterize the fluctuations between different colors—differ markedly from one another.

Considerable differences are seen, in size, shape and alignment of the ellipsoids. The fluctuations in primary colors mainly tend towards varying chroma, whereas fluctuations in mixed colors appear more evenly distributed in chroma, in hue and in lightness. The way in which primary colors fluctuate can be plausibly explained by fluctuations in ink layer thickness and in dot gain. Fluctuations are therefore mainly to be expected along the line running between the solid tone and paper white. This shows up very clearly in the case of yellow, where the ellipses are particularly elongated. However this shape is also caused by a distortion in the CIELAB color system. The human eye does not perceive differences in chroma in the yellow range as weightes in the CIELAB color system.

The fluctuation ellipses shown in Figure 9 grow larger as the dot percentage increases. It is clearly apparent that the fluctuations are greater at a dot percentage of 80% than they are in the case of solids. The reason is, that in three quarter tones, fluctuations in ink-layer thickness occur as well as fluctuations in dot gain.

The primary colors can fluctuate independently of one another. This is why, in the case of secondary colors, the fluctuations do not perform such a marked dominant trend. As expected, the overall fluctuation of the secondary colors are markedly greater than with the primary colors. The fluctuation ellipses in the ranges of the three secondary colors red, green and blue show that particularly strong fluctuations in the hue are to be expected here. Between these three secondary colors considerable differences also occur in the shape and alignment of the ellipses. In the case of blue, the ellipses are relatively small and elongated. They show that in the case of blue the fluctuations in hue are mainly to be expected. The largest fluctuation ellipsoid occurs with green when the yellow and cyan dot percentages are 80%.

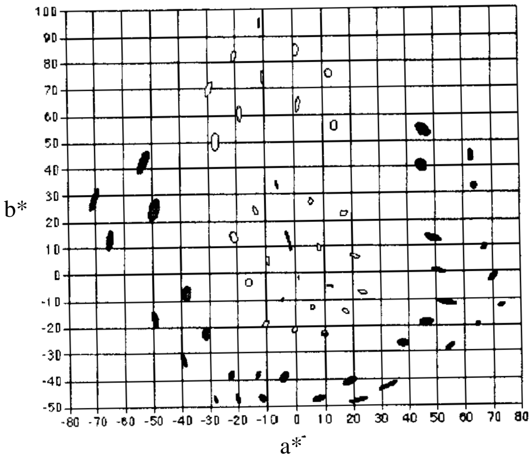

Figure 9. Fluctuation ellipsoids in various parts of the color space obtained by about 100 measurements each.

In order to provide a theoretical estimate of color fluctuations, it was initially necessary to develop an arithmetical model for calculating the mixed colors from the spectral curves of solid primary colors. The formula is based on a model which describes the individual printed dot. A highly variable ink layer thickness in different parts of the screened dot is observed upon examination with a microscope, and this contradicts any assumption that the ink is evenly distributed across the paper. Much more, the ink layer thickness is markedly lower when the dots se small than when they are large. These observations led to the following description of screened primary colors, which serve as a model.

A screen dot is covered by a layer with a certain thickness over an area that matches that of the dot on the printing plate. The ink layer thickness, and thus the color density, depend on the size of the dots, which—at any rate in printing with conventional screening—is a measure of the tonal value. Apart from this portion, there is also a marginal area which renders the dot greater than the dot on the plate, where the ink layer thickness is markedly lower. For the purpose of the calculations in the model, it is assumed that the color density in is half that of the central area The color of ink layers with varied thickness, based on a reference thickness may be predicted with high percision if the ratio of the ink densities is defined and used to calculate the changed spectral reflection factors.[11]

It is a simple matter to extend the method to the estimation of color fluctuations. In this case, it is not the absolute color values that are significant, but only the variations occurring in the color values when the ink density and/or the dot gain changes. It turned out that the calculated color fluctuations corresponded well to the masured fluctuations. Success was achieved in reproducing arithmetically some typical properties of the fluctuation ellipsoids in the various parts of the color space.[12]

5. Accuracy in Color Management, Comparison with Noise Related Data and Conclusion

Our search did not indicate any literature available pointing out adaptable methods to understand error propagation between the devices used in the color reproduction chain. Traditional methods used to evaluate the resulting errors in interacting signals overlayed by random errors do not apply in our case since they base on the assumption that one can add the fourier transforms of the.[13] Since the noise under view is both time and dot area dependent such an is inappropriate and we can just estimate the resulting errors in the process.

As pointed out, spectrometers and scanners perform both reproducible to a degree of a mean ΔE^*_{94} =0.2.

In an ongoing study we evaluate the accuracy of the ICC approach.[14] It turns out that the inaccuracy obtained between using and not using color management systems is significant. However, using color management for the scanners under view we still obtained a medium ΔE^*_{94} =5.0 using an ICC profile and up to $\Delta E^*_{94} = 9$ without.

That is why, studying scanners, the calculation of a profile should also be taken into account. In addition to the outpointed errors, there may be quantization errors of the software used. We have no numbers describing those errors.

Thus one cannot not expect the ICC approach to lead to an accuracy better than $\Delta E^*_{94} = 0.4$ in scanner reproduction. In addition, even when not taking quantization errors into account, significantly higher errors may randomly occure for this device type. The fact that the medium errors obtained are signifivantly higher than that may point to weaknesses in the softwares used or weaknesses in the ICC approach.

Dealing with presses and drivers, our results indicated that they were able to reproduce between 80 and 200 color levels on the basis of 8 bit data. Tests with profiles for offset printing conforming to the BVD/FOGRA standard[8] resulted in an average color deviation of $\Delta E^*_{ab} = 3$ to $\Delta E^*_{ab} = 5$ and a maximum deviation of $\Delta E^*_{ab} = 9$ to $\Delta E^*_{ab} = 10$ for the basic patches of the 1T8/7.3 test plate[14] when simulated on a dye sublimation printer. The results published in reference (15) of the bibliography result into the same range of color deviation. They are also comparable with results presented in reference (16) of the bibliography.

As for scanners, the calculation of a profile should also be taken into account. The profile should be based on overaged measurements in order to reduce the error due to deviation during a run. As indicated by the reproducible color levels, the presses are not able to print color depths of 8 bits.

Given measurement errors and errors occuring when producing the profiles are to be taken into account in the case of printers, too, one has to at least average the random errors in the production of the profile by measuring a significant amout of targets or estimate an inaccuracy of the profile in the region of at least a medium $\Delta E^*_{ab} = 3.0$. Averaging over a larger amount of sheets may lead to the creation of more accurate profiles, but to expect color management for printers to lead to an accuracy significantly higher than a medium $\Delta E^*_{94} = 5 \pm 2$ on an individual sheet under view depending on the CMM and printer or presses involved in reproduction is asking too much.

Concluding one may say that the results presented for printers indicate that if appropriately using ICC profiles and the CMMs and if the profiles were generated with the fidelity of printers, the printers used in the color reproduction may produce a reproducibility close to the random error. Scanners perform better in this respect. However, there is a need to create theory that results into means to calculate error propagation in subsequent processes like in the reproduction of color data.

7. Literature

1. M. Has, T. Newman, M. Stokes: On the ICC Approach to Color Handling, CIE Color Expert Symposium on Color, Wien, 1996.
2. F. Dolezalek, T. Fuchs, J. Schliitken: Reprlsentanz mitgedruckter Kontrollfelder, FOGRA-Forschungsbericht 3.280, FOGRA, Miinchen, 1991.
3. M. Has, H. Wordel: Charakterisierung der Farbziigigkeit von Offsetdruckfarben, Applied Rheology, Hannover, Oktober 1996.
4. CIE Publication 15.2-1986 “Colorimetry”, Vienna, 1986.
5. International Color Consortium ICC Consortium Profile Specification Version 3.0, Boston 1995.
6. DIN 5033.
7. ISO 13655.
8. BVD/FOGRA, Manual for Standardisation of the offset printing process, Wiesbaden, 1992, ISO 12647-2, ISO 2846-1.
9. Michael Has: Ink Control in Sheet Fed Offset Printing, *Proceedings of the 22nd Research Conference of the International Association of Research Institutes for the Graphic Arts Industry Advances in Printing Science and Technology*, Edited by W.H. Banks, Volume **22**, p. 414 ff, Wiley & Sons, 1994.
10. K. Schlidpfer und E. Widmer: Color deviations and variations in Gravure and Web Offset Printing, *TAGA Proceedings* 1995, p. 607 ff
11. A. Paul: Color Fluctuations in Offset Printing, *Proceedings of the IARIGAI*, 19971.
12. A. Paul: Drucktechnische Farbschwankungen im Offsetdruck und ihre farbmetrische Bewertung in verschiedenen Farbraumbereichen, FOGRA Forschungsbericht 52.017, München 1997.
13. V. Beers: Introduction to the Theory of Error, Wesley Publisching Inc, New York, 1962
14. ISO/DIS 12640.
15. A. Ritzer, M. Has: Bewertung von Color Management Systemen, FOGRA Forschungsbericht, Dez.
16. 1997 Florian Süß1: Der gute Ton, MACup, S. 80, November 1996.
17. K. Traber, F. Dolezalek: Anpassung von Digital prüfdruckgeräten an den Offsetdruck, (Adaptation of digital proofing devices to offset printing), Forschungs-bericht (research report) 10.034, BvD/FOGRA, Munich/ Wiesbaden 1996.

8. Appendix

Table I. Number of Colorlevels Obtained from Addressed 8 Bit Levels in C, M, Y, K, C + Y, C + M, M + Y and C + M + Y for Sheet-Feed Offset with AM Screening

	C	**M**	**Y**	**K**
Global ΔE^*_{94}	47.74 ± 0.48	56.53 ± 0.27	41.25 ± 0.52	77.02 ± 0.24
Nb. of levels	89 ± 5	106 ± 3	41 ± 5	142 ± 2
	C + Y	C + M	M + Y	C + M + Y
Global ΔE^*_{94}	68.00 ± 1.57	78.59 ± 0.69	69.14 ± 0.52	95.60 ± 3.43
Nb. of levels	76 ± 13	122 ± 7	69 ± 5	115 ± 18

Table II. Number of Colorlevels Obtained from Addressed 8 Bit Levels in C, M, Y, K, C + Y, C + M, M + Y and C + M + Y for Sheet-Feed Offset with FM Screening

	C	**M**	**Y**	**K**
Global ΔE^*_{94}	46.72 ± 0.13	55.60 ± 0.12	40.56 ± 0.49	78.51 ± 0.52
Nb. of levels	109 ± 2	116 ± 2	43 ± 6	162 ± 5
	C + Y	C + M	M + Y	C + M + Y
Global ΔE^*_{94}	65.02 ± 0.61	73.19 ± 0.47	65. 31 ± 0.54	85.37 ± 1.54
Nb. of levels	93 ± 7	147 ± 4	71 ± 5	126 ± 10

Table III. Number of Colorlevels Obtained from Addressed 8 Bit Levels in C, M, Y, K, C + Y, C + M, M + Y and C + M + Y for Sheet-Feed Offset with non-impact Xeikon

	C	**M**	**Y**	**K**
Global ΔE^*_{94}	48.64 ± 1.58	54.11 ± 0.58	42.02 ± 0.35	68.65 ± 0.38
Nb. of levels	78 ± 6	96 ± 3	47 ± 3	153 ± 2
	C + Y	C + M	M + Y	C + M + Y
Global ΔE^*_{94}	96.61 ± 3.54	75.85 ± 2.15	67.43 ± 3.14	97.68 ± 5.38
Nb. of levels	72 ± 8	122 ± 6	90 ± 12	115 ± 12

☆ This paper was previously published in *IS&T/SID 5th Color Imaging Conference Proc.,* p. 270 (1997).