A Critical Review of Spectral Models Applied to Binary Color Printing

David R. Wyble, Roy S. Berns

Munsell Color Science Laboratory, Chester F. Carlson Center for Imaging Science, Rochester Institute of Technology, Rochester, New York 14623-5604

Received 28 November 1998; accepted 1 May 1999

Abstract: A critical review of binary color printing models is presented. The goal is to provide an understanding of the application of color printer models as a component for device profiles within a color management system. A short description of a modern color management system is presented, followed by a brief explanation of the halftoning process. This leads into the discussion of the individual models, which takes an historical approach. The discussion starts with early models proposed in the 1930s by Murray, followed by Neugebauer, Yule and Nielsen, and other much more recent model forms. To aid in gathering the appropriate data for printer modeling, experimental techniques are then discussed, followed by an explanation of the model optimization methods needed for parameter fitting. The review concludes with procedures for model evaluation and a presentation of the results from an application of the models to a sample dataset for an electrophotographic printer.

Key words: color printing; color halftone; spectral models

INTRODUCTION

The field of color printing has long been dependent upon skilled operators of printing apparatus. These *colorists* visually inspect printed output and determine how to best adjust the system for optimum print and color quality. The proliferation of color desktop publishing has meant that most users of color reproduction systems are no longer experts in color printing. To assist nonexpert users, color management systems have been introduced,[1] which control the processing of color throughout the reproduction system (scanner or camera, monitor, and printer). Color management systems (CMSs) are based upon mathematical models representing the various subsystems involved. Figure 1 shows the information flow through part of a color reproduction process. On the left side of Fig. 1, color management is not used, and the information passed between subsystems is simply the input or output values used by each subsystem. Here, scanner red, green, and blue digital output values are sent directly to the printer, with no intervening transformation. The resulting color output by the printer cannot necessarily be expected to be the same as that of the scanned original. When the dashed arrow is replaced with the CMS on the right, all information flowing between the scanner and printer passes through the *Profile Connection Space*. The profile connection space (PCS) is a color space representing a default set of color viewing and measurement conditions. As its name implies, this connection space acts as the intermediate color space to and from which device profiles must transform input and output device values. In Fig. 1, the scanner profile transforms scanner RGBs into the PCS, a printer profile takes the PCS color as input and predicts the appropriate printer RGBs to reproduce the requested color. A similar process is applied to other subsystems in the color reproduction system. The resulting system can reproduce colors much more accurately than the original system of simply exchanging digital values without any transformations.

As color science evolves, the scope of CMSs is being extended to encompass more advanced types of color information, such as color appearance and cross-media reproduction. This review focuses on a subset of the models needed for a CMS, namely the models used to predict the color output of binary printers. Such printers are capable of depositing only one nominal amount of each colorant in any location. To reproduce colors other than those requiring full colorant or bare substrate, halftoning is used. Figure 2 shows a simple halftoning scheme, where a 2×2 pixel halftone cell is used to reproduce 5 levels of colorant on paper.

Correspondence to: David R. Wyble; email: wyble@cis.rit.edu

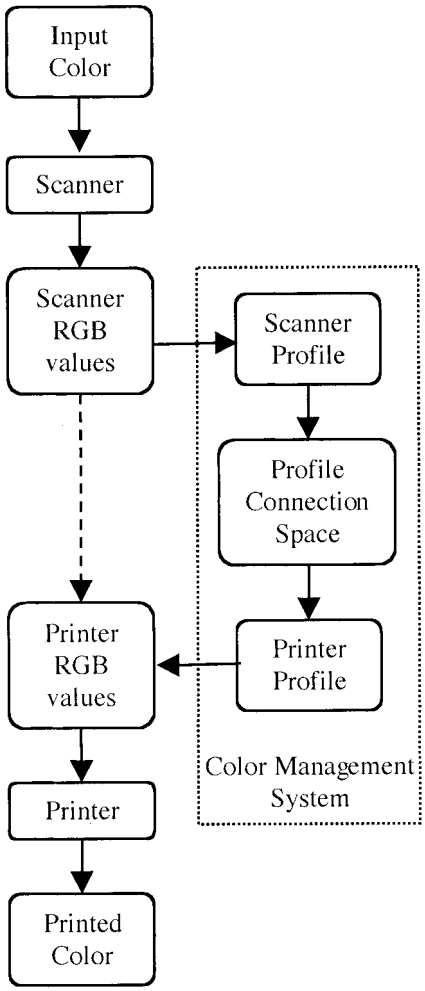

FIG. 1. Information flow in a sample color reproduction system. The left column is without any CMS. Adding the right column in place of the dashed arrow enables the transformations into the colorimetric based Profile Connection Space.

The input to the halftoning process is a grayscale image. Each pixel in this image consists of at least three values: cyan, magenta, and yellow, and often a fourth colorant, black. These values, representing an amount of each respective colorant, vary between 0–100%. In Fig. 2, the halftone cell is a group of four pixels. Any cell for which the input value exceeds the threshold is imaged; cells falling below the threshold are not. By applying the halftoning process, the number of effective colorant levels is increased from 2 to 5. Note that there is a corresponding decrease in image resolution. After halftoning with the sample halftoning

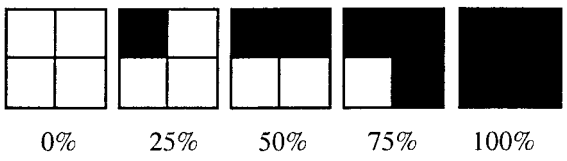

FIG 2. Simplified halftone scheme. Five cells are shown, each labeled with theoretical area coverage.

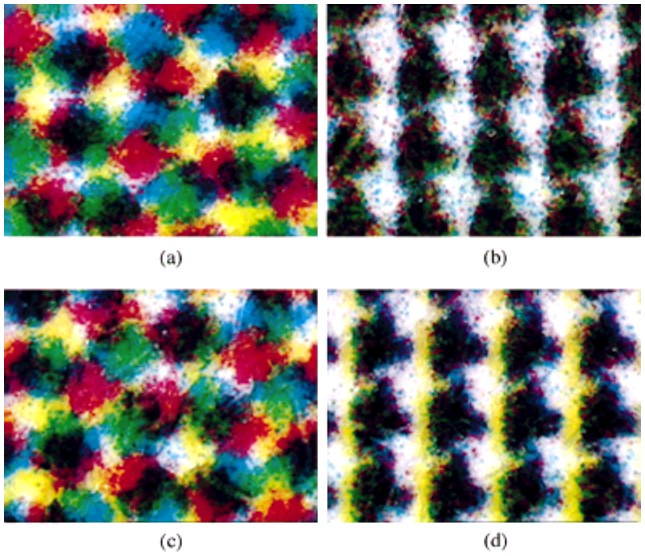

FIG 3. Photomicrographs of prints using the two halftone schemes described in the text. (a) and (c) are typical rotated screens; (b) and (d) show dot-on-dot halftoning. (a) and (b) are from the same, properly-registered page. (c) and (d) are from a page with severely misregistered separations. This misregistered yellow separation has imposed a color difference of about 12 ΔE^*_{ab} between patches (b) and (d). Patches (a) and (c) are only 1.7 ΔE^*_{ab} different. The actual width of each image is approximately 1 mm.

scheme, the smallest unit in the output image is equal in area to four pixels in the input image. One goal of the halftone design process is to balance image resolution and number of colorant levels such that reproduction quality is acceptable. Three types of halftoning are commonly used: rotated screens, dot-on-dot screens, and stochastic screens. Rotated screens place the dots for each separation at different angles, typically 75°, 15°, 0°, and 45° for cyan, magenta, yellow, and black, respectively, where 0° is to the right. Placing the separation dots at these angles reduces the color error that would otherwise occur due to slight misregistration of the various separations. Rotated screens, however, suffer from large-area spatial defects, such as moiré.[2] As the name implies, dot-on-dot screens place dots for all separations in the same place, on a fixed grid. Dot-on-dot schemes depend heavily on good registration, that is, accurately aligning dots from all separations. Color reproduction accuracy suffers significantly, if misregistration occurs. Figure 3 shows examples of rotated and dot-on-dot images. Figure 3(a) and (b) show rotated and dot-on-dot halftoning, respectively. Figure 3(c) is a misregistered form of (a), and Fig. 3(d) is a misregistered form of (b), each with the yellow separation no longer aligned with cyan and magenta separations. The color difference between patches (b) and (d) is about 12 ΔE^*_{ab}. The same shift in yellow registration, when applied to the patch in Fig. 3(a), results in a much lower color difference. Here, the color difference between patch (a) and (c) is only 1.75 ΔE^*_{ab}, which might be expected for page-to-page process variation. Stochastic screens do not cluster the area coverage in a single location in the halftone cell. Rather, area coverage is distributed over the entire cell.

Compared to the other halftoning methods, stochastic screens are relatively new, and the older models presented here were not generally intended for such imaging systems. For more specific details of halftoning see, for example, Fink[2] or Kang.[3]

Many models have been proposed to predict the color output of binary printers. These date to the 1930s and have been refined many times over the subsequent decades. This article is not intended as a complete review of every model ever proposed. Rather, an historical approach starts with early models and tracks their evolution. The models described here are selected, because they have been applied in many contexts over the years and most have been used with some success. Most of these models share a common "lineage" to the original Murray–Davies model, described below. For all models, a spectral form is used. The goal of these models is to predict the *spectral* output of the printer, as opposed to the *color*. This is to avoid metameric matches by predicting physical printer performance (the reflectance spectrum), not perceptual output (the color). Model performance should be reported in the form of RMS spectral difference and color difference in ΔE^*_{ab}. Both metrics should be reported, because a large spectral difference may correspond to small color error. Note that if spectral output is needed, spectral input data are also required.

REVIEW OF MODELS FOR HALFTONE COLOR REPRODUCTION

Printers are typically characterized by printing and measuring a variety of samples, and determining model coefficients, which predict physical properties of the output. The resulting model predicts spectral reflectance or color given the cyan, magenta, and yellow input values. A model thus derived is a *forward* model, where inputs are RGB or CMY levels, and outputs are in color coordinates: CIEXYZ, CIELAB, or another color space. The more useful model form is the *inverse* model, where the CMY printer requests are predicted from color coordinates. This is the form needed for the printer device profile, the input of which is the color request given in profile connection space coordinates. The printer profile can then determine print CMY values that best reproduce the color request. At one time, inverting a device model was a critical operation, and analytical forms needed to be chosen that could be inverted. Modern computing power and algorithm design make the use of a lookup table (LUT) practical. This table is populated by repeated runs of the forward model, and the inverse model is accomplished by interpolation. Interpolation is required, because the input to the device profile is a grid in color space, but the forward model generates a grid in device input values (e.g., CMY values in the case of a printer). The interpolation determines the device values that best reproduce the input color. Usually, this means determining the point on the grid that is nearest the input color. For color applications, the LUT required is multidimensional, and is often referred to as a CLUT (color lookup table). The use of a CLUT relaxes the mathematical constraints on the choice of forward models, freeing the modeler from the need to select an analytical form that is invertable. The complete description of the methods of CLUT creation or other methods of model inversion is beyond the scope of this review. More details and an expanded list of references can be found in Ref. 3

Models described here are limited to three colorants. Most are not difficult to extend to four-colorant printers. The chief complication is that the CLUT is now four-dimensional. The details of four-color models and their implementation can also be found in Ref. 3

What follows is an historical description of the development of certain halftone printer models. Due to the reasons cited above, emphasis is only on the creation of an accurate forward model. It is assumed that the inverse models can be implemented via multidimensional interpolation, and models are not evaluated for their ability to be inverted.

Model Overview

Models presented are of two general types. Most are *regression-based* models. These tend to be relatively simple, with parameters fit to a set of data. They are often useful for modeling printer output, because they tend to be reasonably accurate and their simplicity allows for short calculation times. Regression-based models do not necessarily strive for modeling of the physics of the process, but rather to emulate the behavior of the system. More physically plausible forms, called *first-principals* models, attempt to simulate the physical process. It is hoped that these models accurately estimate printer behavior, but a more important use is to increase understanding of the physical process itself. Often, the first-principals modeling approach cannot predict printer output as well as the simpler models. However, systematic trends should be apparent in the first-principals forms, even though the printer output might not be predicted with absolute accuracy. Also, variation of the parameters should produce output that assists in the qualitative description and understanding of the system.

REGRESSION BASED MODELS

Murray–Davies Model

Murray[4] first published a simple model to predict output density from input dot area. This model was originally proposed privately by a colleague of Murray at the Franklin Institute, E. R. Davies. Indeed, Murray himself refers to it as "Davies' formula." Instrumentation limitations at the time required Murray to work in density. The spectral reflectance form is presented here:

$$\hat{R}_\lambda = a_t R_{\lambda,t} + (1 - a_t) R_{\lambda,s}, \tag{1}$$

where $\hat{R}_\lambda$ is predicted spectral reflectance, a_t is fractional dot area of the ink (the *tint*), $R_{\lambda,t}$ is the spectral reflectance of the ink at full coverage, and $R_{\lambda,s}$ is the spectral reflectance of the substrate. The λ subscripts indicate all three reflectance values are a function of wavelength. Note how few

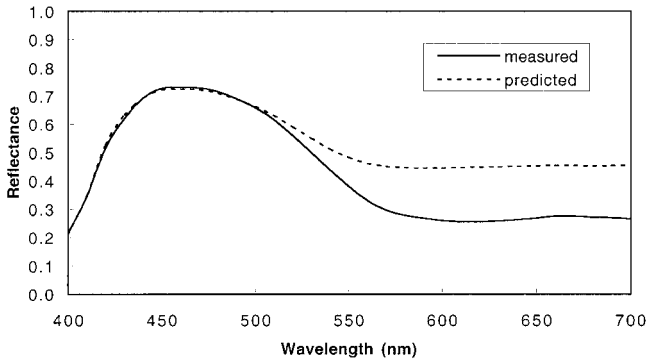

FIG 4. Predicted and measured spectral reflectance for 50% theoretical area coverage. Predicted data were made using the Murray–Davies model in Eq. (1) with theoretical area coverage .

measurements need to be made for this model: only the spectral reflectance of the solid ink and that of the bare substrate. The difficulty comes in estimating the area coverage a_t. Without an accurate understanding of the actual colorant coverage, this model diverges unacceptably from measured data.

Some distinction must be made about various types of dot areas. *Theoretical dot area* (or equivalently *theoretical area coverage*) is calculated from the actual binary image sent to the printer. While not always physically reasonable, a square pixel shape is assumed, and theoretical area coverage is determined as shown by the percentages in Fig. 2. (Pappas[5] presented an in-depth analysis of circular pixel shapes and the implications of overlaps between pixels of the same and differing separations.) *Effective dot area* (or *effective area coverage*) is always an estimated value. For example, if the spectral reflectance is accurately predicted by a particular area coverage, this is referred to as the effective area coverage, a_{eff}. This is equivalent to selecting the best proportion to scale $R_{\lambda,t}$ such that it most closely matches the measured spectral reflectance at area coverage a_t.

Use of the Murray–Davies model requires making the fundamental assumption that both the substrate and the ink are of uniform color. It is important to note that this is rarely the case. For typical lithographic printing on plain paper, dot densities are not uniform. Nevertheless, these assumptions apply to all the early models presented.

The Murray–Davies model makes a good deal of intuitive sense. From a conservation of energy viewpoint, the overall reflectance should be equal to the sum of the various components of light reflected off the print. If the light from each area is accounted for, the model should work well. However, if the theoretical dot area (that is, the printer input value) is used as a_t in Eq. (1), predicted reflectance for halftoned images is consistently higher than measured reflectance, as shown in Fig. 4. This effect is called *dot gain*, the phenomenon whereby measured prints are always darker than predicted; the dots behave as if they are larger than they are, hence dot gain.

The literature generally distinguishes between two types of dot gain. *Physical dot gain* is actual growth in colorant coverage due to the printing process. For example, when applying a drop of ink to a substrate, the drop might spread in flight or upon contact with the substrate. Ink also spreads upon penetrating the substrate. In these cases, the colorant covers more area than originally intended. Alternatively, *optical dot gain* is the scattering of light in the substrate. Here, light that enters bare substrate may be scattered, and exit underneath a dot. Likewise, light entering through a dot might be scattered and exit through bare substrate. These scattering effects cause the bare substrate to appear darker than expected, and the ink dots to appear lighter than expected. The overall effect is a darker image.

Estimating Effective Dot Area Another use of the Murray–Davies model is to predict the *effective dot area*. Effective dot area is found by rearranging Eq. 1, into

$$a_{eff} = \frac{R_{\lambda=min,meas} - R_{\lambda=min,s}}{R_{\lambda=min,t} - R_{\lambda-min,s}}. \tag{2}$$

Note that $\hat{R}_\lambda$ is replaced by the measured reflectance, $R_{\lambda=min,meas}$, and the calculated a_{eff} replaces the input, a_t. All reflectance values are subscripted $\lambda = min$ to signify that this is not a spectral calculation, but one performed at a single wavelength. The choice of λ is typically that of the minimum reflectance for either the measured or predicted tint. (This is the optimum wavelength, because it varies the most when dot area is changed.) In essence, a_{eff} is the factor by which the solid ink reflectance should be scaled to equal the measured reflectance of the sample. This equation can be used to estimate dot gain, improving the estimation of area coverage. To extend Eq. (2) to a spectral form, a_{eff} must be determined by a matrix calculation using least squares analysis:

$$a_{eff} = \mathbf{R}_{\mathbf{meas,adj}}\mathbf{R}_{\mathbf{t,adj}}^{\mathrm{T}}(\mathbf{R}_{\mathbf{t,adj}}\mathbf{R}_{\mathbf{t,adj}}^{\mathrm{T}})^{-1}. \tag{3}$$

Reflectance terms in Eq. (3) are row vectors, the length of which is the number of wavelengths in the spectral measurement. Here, $\mathbf{R}_{\mathbf{meas,adj}} = \mathbf{R}_{\mathbf{meas}} - \mathbf{R}_{\mathbf{s}}$, and $\mathbf{R}_{\mathbf{t,adj}} = \mathbf{R}_{\mathbf{t}} - \mathbf{R}_{\mathbf{s}}$, and the superscipts $^{\mathrm{T}}$ and $^{-1}$ indicate matrix transpose and inverse, respectively. This calculation needs to be repeated for each patch in the separation ramp. Note that this calculation assumes that the *cmy* area coverages are independent. Using Eq. (3), dot gain can be defined as the departure of effective dot area from the theoretical dot area, or $a_{eff} - a_t$. This is a reasonable definition, because the Murray–Davies model is theoretically, if not practically, sound. If dot gain were known, and, hence, a_{eff} could be calculated, monochrome printer output reflectance could be predicted well by using a_{eff} as a_t in Eq. (1).

The calculation of effective area coverage, shown in Eq. (3) represents the first stage of all subsequent models. The use of a_{eff} as input to later model stages is much more practical than simply using printer input CMY digital counts. Effective area coverage is shown for the cyan separation in Fig. 5. The calculated values can be stored, and used as input to more complex models by way of a one-dimensional LUT. Another possibility is to model the effective area coverages. Figure 5 also shows a simple cubic

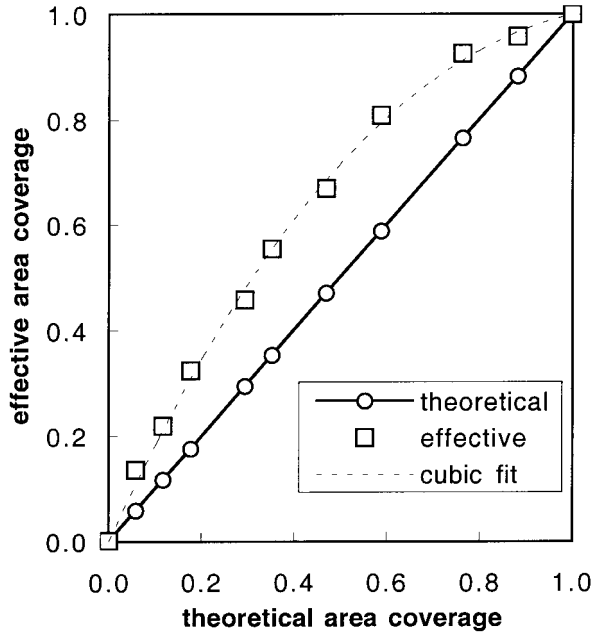

FIG. 5. Effective vs. theoretical area coverage for cyan separation ramp data. Effective area coverages are calculated by Eq. (3). Also shown is a cubic polynomial fit of the calculated data.

polynomial fit. For the sample data evaluated here, a LUT is used.

Interpretation of Murray–Davies Another way to think about the Murray–Davies model is that it simply performs linear interpolation between the reflectance of the paper and that of the full coverage ink, as shown in Fig. 6. The predicted reflectance is the sum of the weighted reflectance of substrate and ink, the weights being the distances along the line in colorant space. When considering Eq. (1), this is not really a new idea, merely an alternative interpretation. This idea is expanded with the Neugebauer model, where the weights are distances along the three axes of colorant space.

Neugebauer Model

The monochrome Murray–Davies model was extended to color by the 1937 landmark work of H. E. J. Neugebauer.[6] When modeling a color print, multiple colorants must be considered. Neugebauer accounted for multiple colorants by summing the product of the fractional area coverage of each colorant and its reflectance at full area coverage:

$$\hat{R}_\lambda = \sum_i F_i R_{\lambda,i,max}\,. \tag{4}$$

Here, the summation is over i, representing the eight *Neugebauer primaries*: bare substrate, single separations, and two- and three-color overlaps. For three-color subtractive printing, these are white, cyan, magenta, yellow, red, green, blue, and three-color black. $R_{\lambda,i,max}$ are the spectral reflectance values of each i^{th} primary at full colorant coverage, and F_i are the fractional area coverages of each respective colorant.

A common assumption used when calculating these area coverages is that the halftone dots are printed randomly on the substrate. (Rogers[7] presents a detailed consideration of randomness of halftone screens. For dot-on-dot printing schemes, Balasubramanian[8] presents alternative interpolation methods.) As shown in Fig. 3(a), dot randomness is fulfilled by the use of rotated halftone screens. This implies that the fractional area coverages of the eight Neugebauer primaries are closely approximated by the probability equations shown below in Eq. (5). Area coverages F_i are estimated for $I = w,c,m,y,r,g,b,k$ using the Demichel[9] equations:

$$\begin{aligned}
F_w &= (1 - a_{eff,c})(1 - a_{eff,m})(1 - a_{eff,y}),\\
F_c &= \mathrm{a}_{\mathrm{eff,c}}(1 - a_{eff,m})(1 - a_{eff,y}),\\
F_m &= (1 - a_{eff,c})a_{eff,m}(1 - a_{eff,y}),\\
F_y &= (1 - a_{eff,c})(1 - a_{eff,m})a_{eff,y},\\
F_r &= (1 - a_{eff,c})a_{eff,m}a_{eff,y},\\
F_g &= a_{eff,c}(1 - a_{eff,m})a_{eff,y},\\
F_b &= a_{eff,c}a_{eff,m}(1 - a_{eff,y}),\\
F_k &= a_{eff,c}a_{eff,m}a_{eff,y}.
\end{aligned} \tag{5}$$

Here, $a_{eff,c}$, $a_{eff,m}$, and $a_{eff,y}$ are the effective area coverages of cyan, magenta, and yellow, respectively, calculated from the separation ramp spectral reflectance data using Eq. (3). These and the spectral reflectance measurements for the eight primaries at maximum area coverage are the only inputs to the Neugebauer model. It is a relatively straightforward extension of Murray–Davies, and, as such, it also makes the same assumption discussed under Murray–Davies, above.

Interpretation of Neugebauer Model Since the Neugebauer model makes such intuitive sense, some interesting implications are often overlooked. When treated simply as a mathematical analysis, the Neugebauer model reduces to performing trilinear interpolation across the printer reflectance space. The nodes for this interpolation are the spectral reflectances of the eight Neugebauer primaries. This is shown graphically in Fig. 7, after Yule,[10] in two dimensions for simplicity. Like Murray–Davies, the inputs to the model are distances along colorant axes and spectral reflectance of

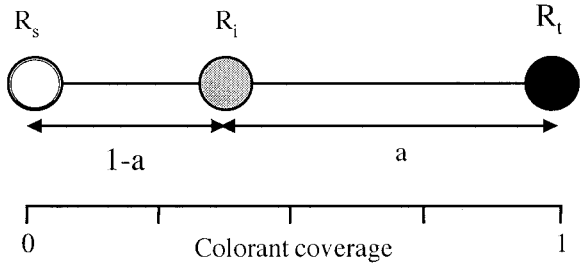

FIG. 6. Murray–Davies model, explained in terms of linear interpolation between substrate and full area coverage ink.

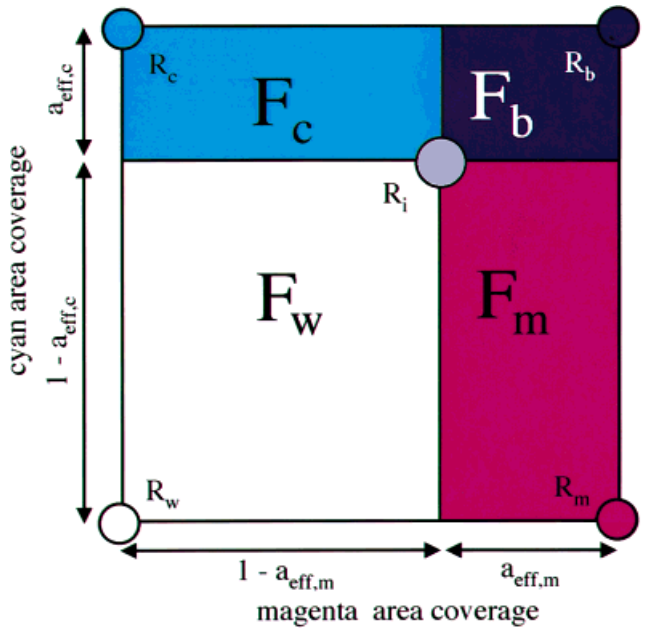

FIG. 7. Graphical interpretation of the Neugebauer model. The weights for each primary are the areas of their respective rectangles.

the primaries. In this case, the weights for the spectral reflectance summation are the areas of the respective rectangles defined by the effective area coverage boundaries. These areas are equivalent to the Demichel weights calculated by Eq. (5). For example, the white area (F_w) is calculated as $(1 - a_{eff,c})(1 - a_{eff,m})$. When extended to three-colorant space, the weights become volumes of rectangular polyhedra, and F_w is calculated as $(1 - a_{eff,c})(1 - a_{eff,m})(1 - a_{eff,})$, which is identical to the calculation in Eq. (5). Given that this interpolation is done across the entire output color space, the performance of the Neugebauer model is going to be quite limited, because it assumes that reflectance space is linear with respect to colorant coverage. In reality, this is not the case, due to effects such as optical dot gain. As such, methods have been proposed to linearize the printer output space with respect to colorant coverage. The most successful of the early attempts of linearization was Yule and Nielsen's now classic article on light scattering in the substrate.

Yule–Nielsen Model

In 1951, Yule and Nielsen published their important work on light penetration and scattering in paper.[11] Their theoretical analysis showed that the nonlinear relationship between measured and predicted reflectance could be well-described with a power function. Starting from the base Murray–Davies model in Eq. (1), they added the exponent $1/n$ to the reflectance values:

$$\hat{R}_\lambda^{1/n} = a_{eff} R_{\lambda,t}^{1/n} + (1 - a_{eff}) R_{\lambda,s}^{1/n}, \tag{6}$$

or equivalently:

$$\hat{R}_\lambda = [a_{eff} R_{\lambda,t}^{1/n} + (1 - a_{eff}) R_{\lambda,s}^{1/n}]^n, \tag{7}$$

where n is a parameter accounting for light spreading in paper, and all other variables are as described as previously. n is referred to as the *Yule–Nielsen n-value, n-value, n-factor,* or often simply *n.*

Fitting *n-value* requires nonlinear optimization. Such techniques are briefly described below, under Model Optimization. To determine the optimum *n-value* for a set of data, the effective area coverages, $a_{eff,i}$, must be calculated using Eq. (3) on reflectance data that have been raised to the $(1/n)$ power. This is done to transform the data into a space in which the Yule–Nielsen model is assumed to be linear. The exponentiated reflectance values are then scaled and summed, as in the Murray–Davies model, and then raised to the n power to inverse transform them back into reflectance space. It is important to understand that the effective area coverages must be updated with each choice of *n-value* during the optimization process. If this is not done, the effective area coverages are being applied in the incorrect space, and there is no guarantee that the optimum value of *n-value* has been found. Alternatively, *n-value* can be chosen by simple iteration. As described below in the section on Evaluating Printer Models, by plotting an error metric against *n-value* a reasonable choice can be made. Note that the need to update effective area coverage still applies.

Interpretation of n-value. Much work has been published on the physical meaning of *n-value*. An argument can been made that the Yule–Nielsen model is inherently nonphysical, due to the exponentiation of reflectance (or density).[12] This is not an unreasonable point. Remember the physical basis for Murray–Davies, and by extension Neugebauer: the reflected light from various areas of the patch are added to predict overall reflectance. Once a nonlinear transform has been applied to the reflectance values, the elegance of the conservation of energy has been lost. In spite of this, the Yule–Nielsen model is often used, because it tends to predict printer output somewhat better than the Murray–Davies model. Likewise, the Yule–Nielsen extension of Neugebauer, described below in Eq. (11), performs better than Neugebauer. The difference between regression-based models and first-principals models is now apparent. When the goal is to derive an accurate printer model, applying a nonphysical transformation is reasonable, if model performance is improved. First-principal modelers cannot take this route, but must delve deeper into the underlying physical process to uncover new aspects not yet accounted for by their model.

Ruckdeschel and Hauser[13] presented an in-depth physical analysis of light spreading in paper and its relationship to n. By inspection, we see that for $n = 1$, Yule–Nielsen reduces to Murray–Davies for monochrome prints. Like Yule,[11] Ruckdeschel and Hauser concluded that $n = 2$ corresponds to highly scattering substrates. Values of n between 1 and 2 are all physically meaningful, while values greater than 2 represent other effects, such as variations in dot density. After experimenting with a variety of substrates, halftone screens, and area coverages, Pearson[14] recommended using

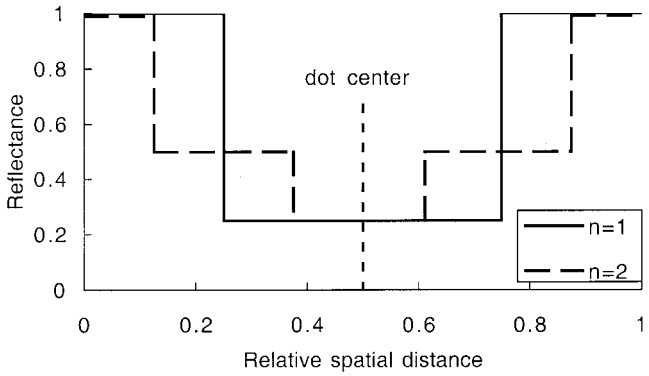

FIG 8. Schematic diagram of reflectance resulting from Murray–Davies and Yule–Nielsen models. Reflectance is shown as a function of spatial distance across a halftone cell. The assumed effective area coverage is 0.5. The solid line shows the two-part Murray–Davies calculation. The dotted line shows the effect of using $n = 2$, effectively adding an area of intermediate reflectance.

$n = 1.7$, unless a better choice is known. Pearson also noted that using $n = 1.7$ always improves performance over $n = 1$. While both of these analyses are sound, is must be noted that values of n greater than 2 are often required for modern, high-resolution printers.

Another interesting interpretation of n was made by Shiraiwa and Mizuno.[15] Their data were fit well by n of 3.0. Once an integer value of n is assumed, additional mathematical expansions can be done that further clarify the role of n in printer models. Suppose that the single separation case is studied, e.g., Eq. (7) with $n = 2$. For this analysis, $R_{\lambda,t}$, the tint reflectance, is replaced by the product $R_{\lambda,s}T^2_{\lambda,i}$, where $R_{\lambda,s}$ is paper reflectance and $T_{\lambda,i}$ is ink transmittance. Equation (7) becomes:

$$\hat{R}_\lambda = [a_{eff,i}(R_{\lambda,s}T^2_{\lambda,i})^{1/2} + (1 - a_{eff,i})\,R^{1/2}_{\lambda,s}]^2, \quad (8)$$

and the following expansion can be performed:

$$\hat{R}_\lambda = a^2_{eff,i}R_{\lambda,s}T^2_{\lambda,i} + 2a_{eff,i}(1 - a_{eff,i})\,R^{1/2}_{\lambda,s}T_{\lambda,i} + (1 - a_{eff,i})^2\,R_{\lambda,s}. \quad (9)$$

The coefficients (areas) and color (reflectance) can be readily seen in this expanded form. What can be learned from this expansion is how the Yule–Nielsen model changes the predicted reflectance compared to the Murray–Davies model. There are intermediate terms corresponding to intermediate reflectances. In Eq. (9), the Murray–Davies area coverage for the primaries is reduced (since both $a_{eff,i}$ and $1 - a_{eff,i}$ are always less than 1), while adding area coverage for the $R^{1/2}_{\lambda,s}T_{\lambda,i}$ term. The overall effect of this is to make the predicted reflectance lower, shown graphically in Fig. 8 for an example, where $a_{eff} = 0.5$. In Fig. 8, there are three levels of reflectance for the $n = 2$ case. It can also be seen that the overall reflectance is lower when using $n = 2$. This is consistent with the earlier statement that the Murray–Davies model overpredicts reflectance. In general, completing this expansion is not recommended, because the model is valid only for integer values of *n-value*. Since this is rarely the case, this section is presented only to increase the understanding of the interpretation of *n-value*.

Arney, *et al.*[16] have found that *n-value* can be approximated as:

$$n \cong 2 - e^{-Ak^{\nu}_p}, \quad (10)$$

where A is a constant relating to dot geometry, k_p is a constant relating to MTF (modulation transfer function, described briefly below), and ν is the halftone dot frequency in dots per millimeter. So, a theoretical analysis of the process by which light spreads in paper can result in an approximation of *n-value*.

Yule–Nielsen Modified Neugebauer Model

The next extension is made by combining Yule–Nielsen and Neugebauer:

$$\hat{R}_\lambda = [\sum_i F_i R^{1/n}_{\lambda,i}]^n. \quad (11)$$

Fractional area coverage F_i is exactly as calculated above in Eq. (5). The density form of this model was proposed by Yule and Colt[17] in 1951, and it was further explored in spectral reflectance space by Viggiano.[18] All the discussion above from both Yule–Nielsen and Neugebauer apply to this model as well. In particular, note that, as before, when fitting *n-value*, effective area coverages must be recalculated and then put in the Demichel equations each time that *n-value* is changed.

Cellular Neugebauer Model

As discussed, the use of any of the above forms of Neugebauer is essentially interpolating through the entire printer gamut from a few points on its surface. Given the nonlinearities imposed by light spreading in paper and many other effects, it is no surprise that interpolating over this space fails, sometimes dramatically. Up to this point, the nonlinearity was accounted for by using effective area coverages, and applying the Yule–Nielsen *n-value* correction. Another approach is to reduce the space over which interpolation is performed by providing more "primaries," some of which are inside the printer gamut. For example, by measuring each colorant at 0, 50, and 100% area coverage, the printer gamut is subdivided into eight smaller cells. This approach, called *Cellular Neugebauer*, was first presented by Heuberger, *et al.*[19] and later explored by Rolleston and Balasubramanian.[20] A simplified case, with two divisions per separation, results in four cells for two-color printing. This is shown graphically in Fig. 9. When using noncellular Neugebauer, reflectance of the test patch, marked with an X, is predicted by interpolating between R_{00}, R_{02}, R_{20}, and R_{22} [Fig. 9(a)]. The use of cellular Neugebauer decreases the size of the interpolation cell. The reflectance of the test patch is now calculated from the reflectance of R_{01}, R_{02}, R_{11}, and R_{12} [Fig. 9(b)]. If more accuracy is required, more

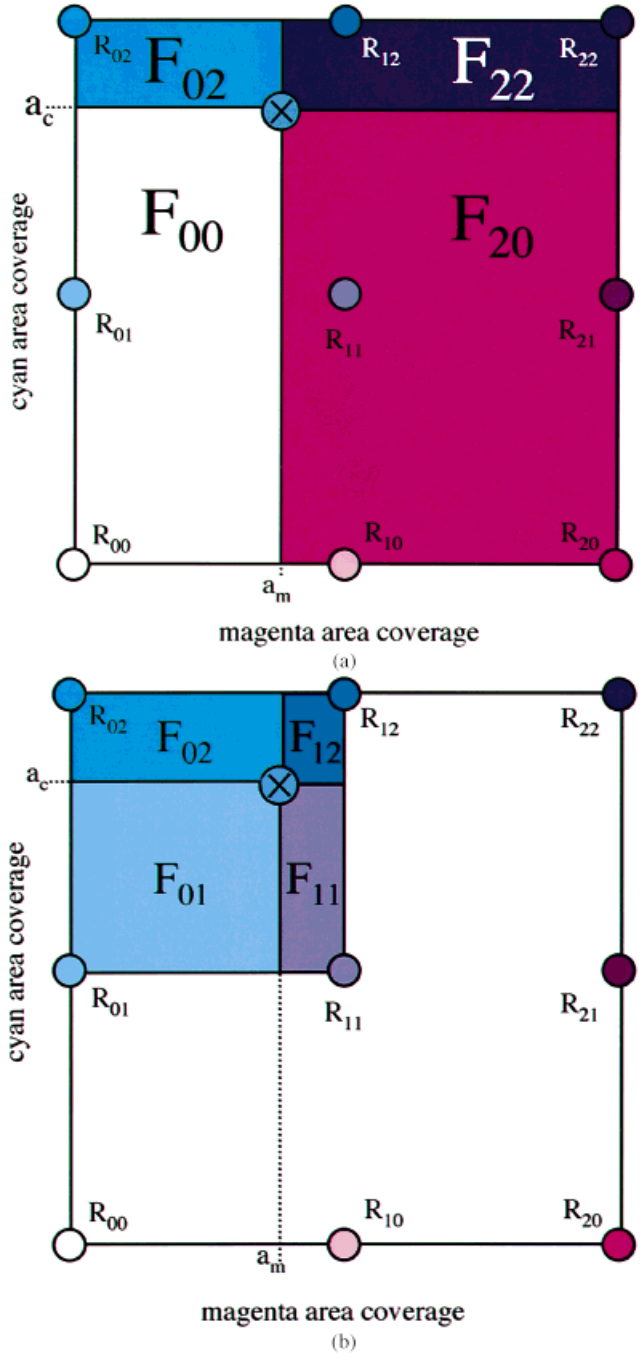

FIG 9. Graphical interpretation of the cellular Neugebauer model: (a) shows the primaries selected for noncellular Neugebauer, as in Fig. 7; (b) shows the primaries selected for the cellular case.

nodes (primaries) can be added at the expense of computation and measurement time.

When using cellular Neugebauer, the interpolation is identical to Eq. (4). However, additional computation is required to determine which primaries to use, or equivalently, selecting in which cell the interpolation is to be performed. This is done iteratively, searching all cells until the following requirements are met:

$$0 \le a_{c,l} \le a_c \le a_{c,u} \le 1, \tag{12}$$

where $a_{c,u}$ and $a_{c,l}$ are the upper and lower area coverages bounding a cell, respectively; and a_c is the theoretical cyan area coverage. A similar search is needed for magenta and yellow. Once the cell is identified, the effective area coverage of the colorant is calculated. Equations (13) and (14) are analogous to Eqs. (2) and (3). They show the scalar form and the least squares analysis on the spectral reflectance data:

$$a_{eff,\,c} = \frac{R_{\lambda=min,c} - R_{\lambda=min,c,l}}{R_{\lambda=min,c,u} - R_{\lambda=min,c,l}}, \tag{13}$$

and

$$a_{eff,c} = \mathbf{R}_{\mathbf{c,adj}}\mathbf{R}^{\mathrm{T}}_{\mathbf{c,u,adj}}(\mathbf{R}_{\mathbf{c,u,adj}}\mathbf{R}^{\mathrm{T}}_{\mathbf{c,u,adj}})^{-1}. \tag{14}$$

This is precisely the same as Eq. (3), with slightly different terms relating to the cellular case. For this case, adjusted values are those that have had the reflectance of the lower primary subtracted, $R_{\lambda=min,c,l}$ in Eq. (13). So, $a_{eff,c}$ represents the effective cyan area coverage in Eq. (5), which, with similarly adjusted magenta and yellow coverages, is used to calculate the effective area coverages of the relevant Neugebauer primaries. (The eight corners of this cell.) At this point, Eq. (4) can be applied to predict spectral reflectance. The above model description comes largely from Ref. 20; interested readers should look there for further information.

One important point is that the cell boundaries do not need to be spaced uniformly. For some printers, it is best to sample the space more frequently in areas where reflectance changes rapidly with respect to input area coverage. The choice of cell boundaries could be an input to model optimization. Although this would be an extremely complex, nonlinear optimization, it would most likely result in a very good cellular model.

Another useful extension of the cellular case is the addition of a Yule–Nielsen factor. This is a straightforward modification and is also explored in Ref. 20.

Spectral *n-value*

Another regression-based model allows n to vary with wavelength.[21,22] The model shown here is spectral Neugebauer with Yule–Nielsen, analogous to the single *n-value* form of Eq. (7):

$$\hat{R}_\lambda = [\sum_i F_i R_{\lambda,i}^{1/n(\lambda)}]^{n(\lambda)} \tag{15}$$

Clearly, this model requires more computation than earlier, single-n models. However, it can improve the model fit significantly. Note that no claim is being made as to the physical significance of the fitted $n(\lambda)$. Modelers may be tempted to draw conclusions about the relationship of light scattering to wavelength, but $n(\lambda)$ is only a fitted parameter.

Regressing the Neugebauer Primaries

So far, models have focused on the nonlinear transformation of the Neugebauer primaries, and linear regression to determine the effective area coverage. Recent work has been done by Balasubramanian,[23] which also includes the fitting of the reflectance of the spectral primaries themselves. This procedure can help avoid the measurement noise that might be associated with the reflectance data required for previous models. However, to train the model as described in Ref. 23, more input spectral reflectance data must be measured than simply the eight primaries and the CMY ramps.

To implement this model, an iterative technique must be applied. The author of the model recommends first using the least squares method, described above, to determine the initial choice of effective area coverages. The reflectance of the primaries is then determined by linear regression transformations using the input training set. Because the primaries have been changed, the effective area coverages must be found again, which may require another iteration to find new primaries. In Ref. 23, only two iterations were needed for a convergence.

FIRST PRINCIPALS MODELS

From this point, the models take fundamentally new forms. New methods are presented here: Arney, Engeldrum, and Zeng;[12] Arney, Wu, and Blehm;[24] and Kruse and Wedin.[25] The first attempts to increase the physical basis of Murray–Davies; this case is extended to color in the second. The last is of a somewhat different spirit than all the previous models. Up to now, ease of measurement and calculation have been important. Now, however, both data collection and model fitting are more difficult, in some cases relying on theoretical rather than empirical parameters. This is to be expected, because these models attempt to better account for physical phenomena that have so far limited the performance of the simpler models. Also note that these models suffer somewhat in terms of absolute accuracy in their prediction of printer output. While a decrease in accuracy is never desired, it is acceptable in this case, because the chief utility here is to learn about the physical process, not necessarily to derive a model suitable for a printer profile. Hence, these models are presented in an introductory form only. Interested readers should seek the references for more detailed information.

Expanded Murray–Davies Model

Arney, Engeldrum, and Zeng[12] did not want to lose the physical plausibility of the Murray–Davies model, but they wanted to make improvements to account for its known deficiencies. It was shown that the reflectance of paper between the halftone dots varied continuously as area coverage is increased.[26] So, paper reflectance is not fixed, as the R_s of the older models, but rather it is a function of area coverage as well. It was also shown that dot reflectance is a function of area coverage. The final form of the expanded Murray–Davies model is identical to Eq. (1). However, reflectance of the ink, $R_{\lambda,t}$, and substrate, $R_{\lambda,s}$, are functions of area coverages $a_{eff,t}$, and $a_{eff,s}$, respectively. The equations for these functions are:

$$R_{\lambda,t} = R_{\lambda,g}[1 - (1 - T_{\lambda,i})a_{eff,t}^{w}][1 - (1 - T_{\lambda,i})a_{eff,t}^{v}], \quad (16)$$

and

$$R_{\lambda,s} = R_{\lambda,g}[1 - (1 - T_{\lambda,i})(1 - a_{eff,s}^{w})] \times [1 - (1 - T_{\lambda,i})(1 - a_{eff,s}^{v})]. \quad (17)$$

Several parameters are introduced here: $T_{\lambda,i}$, the ink transmittance, is equal to $(R_{\lambda,i}/R_{\lambda,g})^{1/2}$ at full area coverage; $a_{eff,s}$, the effective area of the substrate, is $1 - a_{eff,t}$; $R_{\lambda,g}$ is the measured substrate reflectance; and w and v are empirically fit parameters. Note that $R_{\lambda,g}$ and $T_{\lambda,i}$ serve similar functions as $R_{\lambda,s}$ and $R_{\lambda,t}$ in Murray–Davies: they represent the influence of the bare substrate and the ink at maximum area coverage. This model is linear with reflectance, and does not use the nonphysical exponentiation of reflectance. The parameter v is intended to model softness of the dot edges. (According to the printer's terminology, a "hard dot" is very sharp, with a clean distinction between dot edge and substrate; while soft dots are feathered, and ink thickness decreases near their edges.) w represents the effect of light scattering, as with *n-value*.

Extending to Color: the Probability Model

Just as the Murray–Davies model must be correct due to conservation of energy, so too should the Neugebauer model work, if the correct reflectances of ink and paper can be estimated. Arney, Wu, and Blehm[24] reasoned that it should be possible to determine the probability that light would enter and exit any given pair of color areas. In our examples, only cyan and magenta ink are considered; hence, there are four possible routes of entry into the paper and four possible routes of egress. The four colors present are cyan, magenta, blue, and white. So, there are sixteen light paths that need to be considered. After deriving forms for all sixteen probabilities, the reflectance of the primaries can be estimated, and mixture reflectance can be calculated.

For this model description, the subscripts i and j indicate the color of the ink through which light leaves the paper, and the ink through which light enters the paper, respectively. First, the area coverages of the primaries are calculated using Eq. (5), the Demichel probability equations. P_{jj}, the probability that light will exit the same color area it entered, is calculated as:

$$P_{jj} = 1 - (1 - f_j)[1 - (1 - f_j)^{w} + (1 - f_j^{w})], \quad (18)$$

where f_i are the primary area coverages, and w is an empirically fit parameter. For light entering and exiting different color areas, the probabilities are calculated using Eq. (19):

$$P_{ji} = (1 - P_{jj})\left(\frac{f_i}{1 - f_j}\right). \quad (19)$$

The sixteen probabilities calculated above are essentially used as weights in a summation to determine the reflectance of each primary. The primary ink reflectance is calculated by

$$R_{\lambda,i} = R_{\lambda,s} T_{\lambda,i} \sum_j \left(T_{\lambda,j} P_{ji} \frac{f_j}{f_i} \right), \tag{20}$$

where all terms are as previously defined. Recall that the subscript i is where light leaves the paper, and j is where light enters the paper. Equation (20), then, simply sums the color of all the light that will exit the paper through color i, reflects this light off the substrate R_s, and then filters it with transmittance T_i on the way out. Once the reflectances of all primaries are known, they can be used directly in Eq. (4), the original Neugebauer form.

Like Murray–Davies, this model retains a sense that conservation of energy is obeyed. The model still requires the empirical exponent w, but a physical interpretation has been made,[24] relating w to halftone geometry and substrate properties. Because of its physical plausibility, this is a useful tool for exploring the complex process of the interaction of light with ink and paper.

Modeling Paper Spread Function

Taking a step away from the Neugebauer and Yule–Nielsen forms, research has been published addressing the theoretical physics of light scattering within the paper and the dots.[25–29] Also, attempts have been made to directly measure the paper spread function or paper MTF, the modulation transfer function.[30] (Here, paper spread function and MTF are treated together. Paper MTF is the paper spread function after transformation into the frequency domain. Hence, either function can be derived from the other by processing with a forward or inverse Fourier transform.) For the purposes of this article, MTF can be considered the amount a binary input image is blurred by light spreading in the paper. For example, a sharp edge between the substrate and ink is softened somewhat. Whether measuring or calculating the paper spread function, the result is applied to the halftoned image in the form of a convolution:

$$\hat{R}_{x,y,\lambda} = [I_{x,y,\lambda} T_{x,y,\lambda} * P_{x,y,\lambda}] T_{x,y,\lambda}, \tag{21}$$

where $I_{x,y,\lambda}$ is the incident light, and $P_{x,y,\lambda}$ is the point spread function of the substrate, the mathematical description of light spreading. Here, the predicted reflectance, $\hat{R}_{x,y,\lambda}$, and ink transmittance, $T_{x,y,\lambda}$, are both functions of spatial location, and, therefore, the halftone dots are no longer necessarily uniform. The desire to model nonuniform dots is shown graphically in Fig. 10. Here, there are two sets of arrows, each depicting a ray of light. The arrows of each pair pass through the same set of inks. For example, the pair of arrows on the left enter in through magenta and exit out through blue. These paths cannot be distinguished by the probability model. It should be clear that the amount of absorption is going to be different for these two light paths, the solid arrow passing through much more cyan than the dotted arrow. Similarly, the arrows on the right pass through only cyan, but very different amounts of cyan. This shortcoming has been addressed by Rogers,[31] who has proposed a method for incorporating an MTF-based model and the Neugebauer equations. Since the MTF model can account for nonuniform dots, it has a more physically plausible basis. This gain is made at the expense of greatly increased computational complexity. When the MTF is empirically characterized, measurement complexity increases significantly as well.

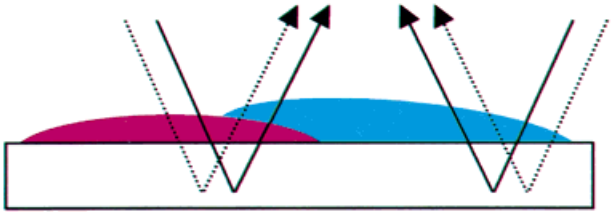

FIG.10. Justification of the need to model nonuniform dots. When applying the probability model, each pair of light paths are treated the same.

A more simplistic explanation of the color models can aid in the understanding of the MTF model. Consider that Neugebauer attempts to predict spectral reflectance by using the eight Neugebauer primaries, which are combined in varying ratios to predict output spectral reflectance. The probability model increases the number of primaries to 64, and describes better the physics of the system. The MTF model, because of its use of convolution, has essentially an infinite number of primaries. The nonuniform dots in most printing processes imply that an infinite number of primaries are present. The MTF model, therefore, best represents the physical process. However, this should not be confused with the best model choice for a color management device profile, which depends on many other factors, briefly described below.

EXPERIMENTAL TECHNIQUES

There are several important considerations to make regarding the creation and printing of a test target for fitting and verifying the models described here. This section is intended to provide guidelines to aid in creating appropriate data to which these models can be applied.

Printer Description

One of the difficulties often encountered in deriving a printer model is the lack of control over printer hardware. Printers often have built-in *Gray Component Replacement* (GCR), where areas containing cyan, magenta, and yellow colorant are replaced by black. For example, suppose that a dark red patch is desired, consisting of exactly 20% cyan, 50% magenta, and 50% yellow area coverage. In many printers, GCR is applied to these input values, resulting in a slightly different request sent to the actual printing hardware. In the example above, the printer software might

choose a more efficient set of colorants to print, such as 20% black, 30% magenta, and 30% yellow area coverage. (This is more efficient, because total area coverage applied has been reduced from 120% to 80%, out of a total of 400% possible for a four-color printer. We assume that the printer calibration is predicting that the observed color is the same in either case, so it makes economic sense to choose the request resulting in less ink usage.) Other changes might be made even if only a single colorant is requested. For example, printer software might be set up to automatically account for dot gain, and input area coverages would be adjusted accordingly. While these limitations should not affect the ability to derive a printer profile, when implementing a model it is most helpful to know actual area coverage specifications sent to the imaging hardware.

Unfortunately, most modern printers do not allow the user to bypass these internal control algorithms, making the job of the modeler somewhat more difficult. The most important concern is that it is usually impossible to guarantee that no black is printed, requiring that the four-color form of a model be implemented. This does complicate model derivation, but not so much as to make the process intractable for those willing to apply the statistical techniques described here.

Test Target Description

There are three main considerations in the design of the test target: printer stability, separation registration, and required test data. Printers are available in many technologies, and, therefore, have differing print-to-print and run-to-run stability. Likewise, separation registration (alignment) differs among these various technologies. For all printers, the optimal test data depends on the model selection. The description that follows assumes that stability and registration should be monitored.

Each print should consist of three sections: control patches to assess stability, alignment cross-hairs to assess registration, and sample patches. The size of the patches depends on the instrumentation used for the spectral measurement, described below. The control patches should be sets of seven patches: cyan, magenta, yellow, red, green, blue, and three-color black. If within-print instability is expected, the control patches should be placed on each end of the print. There should be alignment cross-hairs in two corners of the print, one cross-hair each of red, blue, green, and black. Data patches can be placed in any fashion on the print, but care should be used to not overburden the printer with too high a total area coverage. For some technologies, this might affect process stability.

To monitor printer stability, the color of all control patches should be measured. Corresponding colors across multiple prints are averaged, and the control values of individual prints are compared to this average. Depending on the required accuracy, a threshold color difference should be set. Prints exceeding this threshold should be discarded. Separation registration is evaluated by the alignment cross-hairs. This can be done either visually or instrumentally. Note that the importance of registration depends heavily on the halftoning scheme used, as discussed in the introduction above.

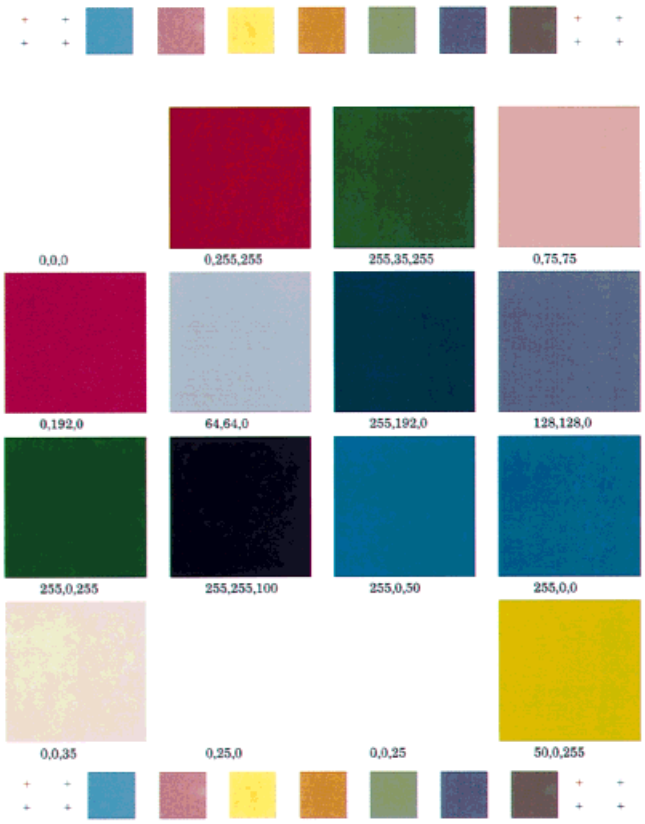

FIG. 11. Example test target, reduced.

The selection of test data is dependent on the choice of models. In all cases, fitting effective area coverages requires several levels in each separation ramp. The exact number of levels depends on the number of levels in the halftone, but practicality must enter the process as well. As a guide, eight to ten steps should be sufficient. There should be three or four separation ramps, one for each colorant.

Other data to be printed are the Neugebauer primaries resulting from overlays, i.e., red, green, blue, and three-color black. For a four-color model, each of these, along with the cyan, magenta, and yellow primaries also need to be combined with black, for a total of 16 primaries.

To verify the model, test data needs to be generated. Consideration needs to be paid to colors in the interior of the printer gamut space. The ramp and primary data described above all lie on the gamut surface. A uniform sampling of printer space is useful to apply, such as $5 \times 5 \times 5$ sampling, where each colorant is given 5 levels, and all 125 combinations are printed and measured. Note that some of the more complex models, in particular the cellular models, require such data for model fitting, and even more for proper model verification.

An example test target is shown in Fig. 11. Alignment cross-hairs are in each corner, control patches are at the top and bottom, and samples patches are the large areas in the center. Sample patches are labeled with their respective

cyan, magenta, and yellow digital values. Depending on the choice of measurement device, the sample patches can be made significantly smaller, thus fitting more per page and requiring fewer overall pages to be printed. This is desirable, because fewer pages means less opportunity for printer variability to impact the data. Data for the examples presented here were measured on a large-area integrating sphere device, which required a sample area of at least 25 mm in diameter.

Sample Measurement

The type of measurement depends both on the accuracy desired and the time needed for measurement. An automated process is always useful, but not typically available for integrating sphere devices. For the best materials analysis, integrating sphere, specular excluded (SPEX) should be used. Since all the models here are spectral-based, spectral data need to be gathered. Typically, the wavelength range is 400–700 nm, at 10-nm intervals. A finer sampling might improve the model performance, but 10 nm has been used with success in many applications.

MODEL OPTIMIZATION

Optimization tools

There are several tools and techniques commonly used for optimizing the halftone models described here. All of the mathematical analysis needed can be done in a spreadsheet, such as Microsoft Excel®. However, models requiring larger datasets might find the spreadsheet application too slow, and more specialized mathematical packages could be applied, such as IDL™, from Research Systems Inc., or Matlab™, from The Math Works, Inc.

Optimization Techniques

As model forms have evolved, the complexity of the optimization required has typically increased. Murray–Davies and Neugebauer require linear optimization to estimate effective area coverage. The Yule–Nielsen *n-value* requires nonlinear optimization, as does the extended Murray–Davies model. The cellular Neugebauer model could be used with only linear optimization, by arbitrarily choosing the cell boundaries, or with nonlinear optimization, using the tools to select cell boundaries. In both cases, linear optimization is needed for estimating effective area coverages. All these types of optimizations can be done using any of the tools listed above. For a mathematical introduction to nonlinear estimation techniques, see Ref. 32.

Optimization Objectives

The ultimate goal of these optimizations is to minimize the spectral error of the models. Therefore, a useful objective function is the root mean square (RMS) of the difference between the measured and predicted spectral reflectance. To retain the spirit of minimizing the measurement requirements, fitting can be done against separation ramp data only. However, some of the models, e.g., cellular, require the measurement of much more data, and these too should be included in the model optimization.

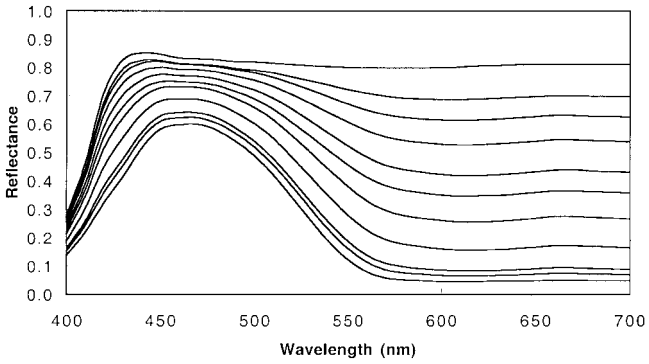

FIG. 12. Reflectance data from cyan separation ramps.

EVALUATING PRINTER MODELS

Model Evaluation Technique

Nearly all the models presented above require relatively few measurements to be made. Of course, this means that the models must then perform some type of interpolation to predict printer output for input values not in the measured dataset. Since the focus here is on spectral-based models, it is important to examine the reflectance spectra of the measured data. By viewing the data in this way, an intuitive sense can be gained as to the utility of the respective models. Figure 12 shows the reflectance data for cyan separation ramp data. The simplest model presented, Murray–Davies, interpolates between the highest curve, for area coverage of 0, and the lowest one, for area coverage of 1.0. Due to the great difference between reflectance of paper and the solid area patches, it is difficult to interpret how well these data can predict the intermediate reflectance values of cyan area coverage. Therefore, a normalized form of reflectance is used, after Berns, *et al.*[33] The equation used for normalizing cyan reflectance is:

$$R_{\lambda,normalized} = \frac{R_{\lambda,c} - R_{\lambda,w}}{(R_c - R_w)_{\lambda=minimum}} \tag{22}$$

where $\lambda = minimum$ corresponds to the wavelength at which full area coverage cyan reflectance is minimum. Similar forms are applied to magenta and yellow. Looking at the normalized cyan reflectance data in Fig. 13, the accuracy of the Murray–Davies interpolation can be visually assessed more easily. For the Murray–Davies equation to hold, all these spectra would be coincident. Since this is obviously not the case, it is clear that some enhancement needs to be made. (Note that for this and all the other normalized spectral plots, there is no reason why values cannot be greater than one or less than zero. This is simply

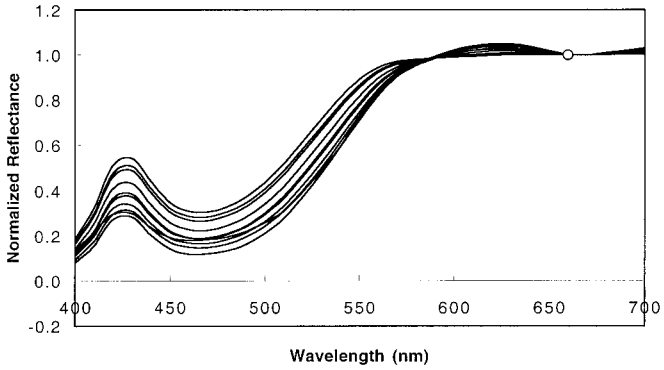

FIG. 13. Normalized cyan reflectance data. Normalized reflectance is described in the text, and calculated using Eq. (22). The circle represents the wavelength to which the data were normalized, λ = minimum in Eq. (22).

an indication that the minimum point of the ink reflectance spectrum is not at the same wavelength as that of the paper.)

Metrics reported for model evaluation should be the RMS difference of the predicted and measured reflectance spectra, and color error measured as ΔE^*_{94}. RMS difference is calculated as follows: for each test patch, spectral difference is calculated as

$$\left\{ \sum_{\lambda} [(\hat{R}_i - R_{i,meas})^2] \right\}^{1/2} \tag{23}$$

These spectral differences are then averages for all patches. RMS difference may be reported for mixture data, while the normalized reflectance data can be reported for separation ramp data. This is because the ramps are measured and used as interpolation anchors, while the mixtures represent the predictive performance of the model. Model performance is best reported as spectral and color error of mixture data (data not used in the derivation of the model).

For the sample data evaluated here, reporting is done only for cyan and magenta separation ramps, and cyan and magenta mixtures. A summary of model performance for all data is shown in Table I for separation data, and in Table II for mixture data. Note that the results summarized here should not be taken to apply to all printers and media. Rather, they are shown to give the reader an idea of the relative performance of the models for a particular set of reflectance data, and to aid in gaining an understanding of how much the performance of the various models should be expected to differ.

TABLE I. Summary of model performance for separation ramp data

		ΔE^*_{94}		
	Separation	Ave	Max	Spectral RMS
Murray–Davies	Cyan	2.2	3.8	0.137
	Magenta	2.8	4.7	0.086
Yule–Nielsen corrected Murray–Davies (n = 5)	Cyan	0.5	0.7	0.024
	Magenta	0.5	0.9	0.025
Yule–Nielsen corrected Murray–Davies, $n(\lambda)$	Cyan	0.6	1.1	0.000
	Magenta	0.4	0.8	0.000
Expanded Murray–Davies	Cyan	1.3	2.3	0.070
	Magenta	1.8	3.1	0.058

TABLE II. Summary of model performance for mixture data

	ΔE^*_{94}		
	Ave	Max	Spectral RMS
Neugebauer	2.1	4.4	0.095
Yule–Nielsen corrected Neugebauer (n = 5)	1.2	3.1	0.050
Cellular Neugebauer, 2 divisions	1.2	3.2	0.002
Yule–Nielsen corrected Cellular Neugebauer, 2 divisions, n = 5	1.1	3.8	0.002
Yule–Nielsen corrected Neugebauer, $n(\lambda)$	1.3	2.9	0.053
Regressing spectral primaries	1.5	3.9	0.060
Probability Model (w = 0.8)	5.9	13.8	0.294

Sample Data

The data used to evaluate these models were measured from Xerox MajestiK™ electrophotographic printer. The printer was modified for research purposes such that it printed the requested c, m, y levels without any attempt at color management. To limit the impact of printer variability, all images were made on a single print run. Multiple prints of each image were made, and all control patches were measured. Prints near the overall average color of control patches were used; those exceeding a ΔE^*_{ab} of approximately 1.0 were discarded.

Murray–Davies Model

Normalized spectra for the cyan separation ramps are shown in Fig. 13. As mentioned, they are not coincident, and it is clear that further improvements must be made. Recall that the input area coverages for these calculations are effective area coverages estimated by least squares analysis. Since this model is so simple, it should not be expected to perform very well.

Neugebauer Model

Since the Neugebauer model is a simple three-color extension of Murray–Davies, the normalized reflectance spectra are the same as those in Fig. 13. As with the Murray–Davies model predictions, the input area coverages are effective area coverages estimated by least squares analysis. As seen in Table II, the mixture fits are good, but not exceptional.

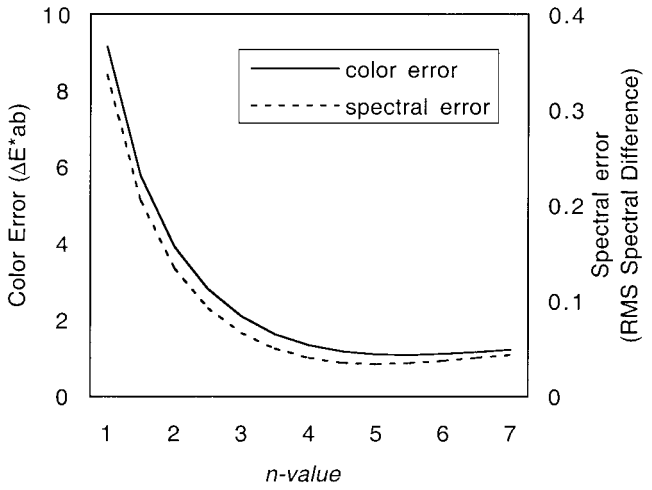

FIG 14. Color error vs. *n-value* and RMS spectral error vs. *n-value*.

Yule–Nielsen Model

We now proceed to the models for which we expect some improvement in the normalized reflectance data. For Yule–Nielsen type models, the measured spectral reflectances are first transformed into the "1/*n*" space, the space in which the Yule–Nielsen model is assumed to be linear. To do this, reflectance at each wavelength is raised to the 1/*n* power before normalization. First, the choice of *n-value* needs to be made. To avoid complications associated with optimizing *n-value*, these data were generated by simply repeating the calculations while varying *n* from 1.0–7.0 in increments of 0.5. An example of results from such an iterative selection of *n-value* is shown in Fig. 14. For spectral modeling, *n-value* should be chosen to minimize the RMS spectral error between the measured data and model predictions. The normalized "1/*n*" space plots are shown in Fig. 15 for cyan ramp data. As with the Murray–Davies normalized data, the extent to which the ramp data are coincident indicates the validity of the model, and Yule–Nielsen is clearly better than Murray–Davies. The Yule–Nielsen corrected Murray–Davies (Table I) and Neugebauer (Table II) both perform substantially better than their uncorrected counterparts.

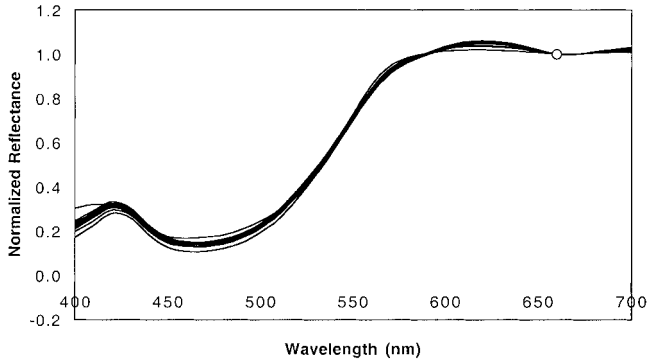

FIG 15. Normalized cyan reflectance in Yule-Nielsen 1/*n* space for $n = 5.0$. The circle represents the wavelength to which the data were normalized, λ = minimum in Eq. (22).

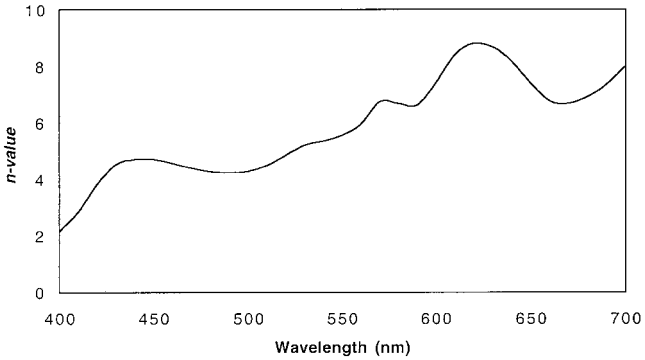

FIG 16. Example $n(\lambda)$ as regressed for the magenta and cyan separation ramp data.

Cellular Neugebauer Model

The normalized reflectance data are the same as the Murray-Davies model, shown in Fig. 13. The cellular model should show improved performance over Neugebauer, depending on the number of cellular divisions. Rolleston and Balasubramanian[20] showed a noticeable improvement by using 3 or 5 divisions per separation in their evaluation. This would be equivalent to dividing the two-dimensional cyan-magenta colorant space, shown in Fig. 7, into 9 or 25 areas rather than only four.

The Yule–Nielsen model can also be applied to cellular Neugebauer. This is explored fully in Ref. 20. However, once a relatively large number of cells are used, the incremental improvement gained by applying the Yule–Nielsen model is small. This is because the nonlinearity in the colorant space is adequately compensated for by small cells, and the *n-value* correction is not needed. Results from the uncorrected and Yule–Nielsen corrected cellular cases are in Tables I and II. Applying the correction improved the predicted reflectance somewhat, but not significantly. It is interesting that even in this limited example, with two cellular divisions per separation, the improvement gained by the correction is slight.

Spectral *n-value*

Separation ramp reflectance data should be used to fit $n(\lambda)$. Since there are now many *n-values*, an iterative technique must be applied, typically by applying a nonlinear optimization routine. Here, the 31 $n(\lambda)$ were fit simultaneously to both cyan and magenta separation ramp data using a nonlinear search routine. Using these optimized *n-values*, effective area coverages were calculated and used for predicting ramp and mixture reflectance data. Some improvement might be found by iterating this process (repeating the nonlinear optimization of $n(\lambda)$ with the new effective area coverages), but no further optimization was done for this sample data. The optimized $n(\lambda)$ is shown in Fig. 16. The shape of the curve is difficult to interpret; as mentioned earlier, caution should be used when attempting to extract physical meaning from the curve shape.

The normalized reflectance plots for cyan are shown in

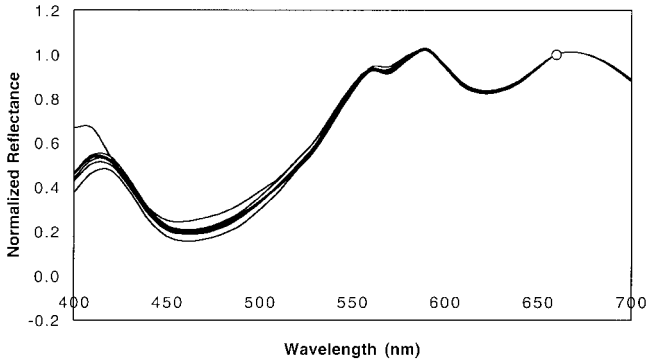

FIG 17. Normalized cyan reflectance in Yule–Nielsen $1/n(\lambda)$ space. The circle represents the wavelength to which the data were normalized, λ = minimum in Eq. (22).

Fig. 17. Note that these are again in $1/n$ space, although here, n varies with wavelength. The model should be expected to perform well considering the good consistency shown by the normalized plots. Like the Yule–Nielsen model, there can still be only one n for all colorant combinations. In this case, one n for each wavelength is applied to predict mixture reflectance data. The predicted separation ramp data results, shown in Table I, are very accurate. This is due to the fact that 31 parameters (*n-values* for each wavelength) were fit to about 20 samples of spectral reflectance data. Applying this $n(\lambda)$ to mixture data predicted the input reflectance very well. Note that, due to rounding, the RMS spectral error for this model is listed as zero. It is also interesting to note that the RMS spectral error is lower and the average color error is slightly higher for this model than for the single-n Yule–Nielsen. It is not clear why this is the case, but it provides a good example of why both metrics are useful for evaluating color device models.

Regressing the Neugebauer Primaries

For this model, the reflectance of the Neugebauer primaries were adjusted to minimize the spectral error of the mixture predictions. This is the only model evaluated in this review that uses the mixture data for parameter fitting. Ideally, model performance should be evaluated using test samples that are not used for fitting the model. Even with these considerations, this model performed quite well. Note that there is no Yule–Nielsen correction applied here; the results shown in Table II are for the uncorrected Neugebauer model using the fit primary reflectances.

Expanded Murray–Davies Model

Since this model can only be applied to separation data, results are shown in Table I. For the sample data, the parameters w and v were fit to minimize the RMS spectral error between the measured and predicted cyan and magenta separation ramp reflectance data. Performance is improved over Murray–Davies, but still not to the level of the empirical forms. This is not surprising, but rather encouraging that we can improve the performance of the older models and retain the physical plausibility of a physics-based model.

Probability Model

The probability model was applied to the mixture data; the performance is also shown in Table II. The parameter w was chosen by averaging the values needed to minimize RMS spectral error on cyan and magenta ramps reflectance data. The performance is somewhat worse than the empirical models, but, again, this is not really the most important metric by which the first-principals models should be evaluated. This model provides a means to explore the effects of dot geometry and paper spreading.

CONCLUSIONS

The need for accurate printer models has increased due to the proliferation of color desktop publishing. The typical users of such a system are not experts in color, and, even if they were, the printers used do not usually have the controls available that are required to fine-tune the printing process. Good color reproduction is, therefore, dependent on the implementation of color management systems. The heart of a CMS is the device profile transforming device input values to color. This review has explored and explained many models that can be used as a component of the printer device profile. Most of the models presented are reasonably accurate. However, accuracy is only one metric that should be considered when selecting a model. Other important characteristics are the number and type of measurements required, and the computational complexity and time needed.

When selecting a model, one should generally choose the simplest model that meets the requirements of the application. It is likely that the very simple models presented, Murray–Davies and Neugebauer, are not suitable for any practical use. However, reasonable performance might be achieved using the Yule–Nielsen form. This should be investigated before the software and measurement investment is made in the more complex models. Perhaps in the future, such a solution will be fully automated, but for now, the modeler must still understand the details of the available techniques to properly implement a printer model.

ACKNOWLEDGMENTS

This article was completed under the support of the Munsell Color Science Laboratory. Special thanks are due to Dr. Jonathan Arney, of Rochester Institute of Technology, for many helpful discussions regarding its presentation and content. The authors also express their appreciation to anonymous reviewers for numerous suggestions, which significantly improved the work.

1. International Color Consortium, ICC Profile format specification ver. 3.4. Published www.color.org (1997).

2. Fink P. Postscript screening: Adobe accurate screens. Mountain View: Adobe; 1992.
3. Kang H. Color technology for electronic imaging devices. Bellingham, Washington: SPIE; 1996.
4. Murray A. Monochrome reproduction in photoengraving. J Franklin Inst 1936;221:721–744.
5. Pappas TN. Model-based halftoning of color images. IEEE Trans Im Proc 1997;6:1014–1024.
6. Neugebauer HEJ. Die theoretischen grundlagen des mehrfarbendrucks, Zeitscrift fur wissenschaftliche Photographie 1937;36:73–89 [Reprinted as Neugebauer memorial seminar on color reproduction, Proc SPIE 1989;1184:194–202.]
7. Rogers GL. Neugebauer revisited: random dots in halftone screening. Col Res Appl 1998;23:104–113.
8. Balasubramanian R. A printer model for dot-on-dot halftone screens. Proc SPIE 1995;2413:356–364.
9. Demichel ME. Procédé 1924;26:17–21, 26–27.
10. Yule JAC. Principles of color reproduction. New York: Wiley; 1967.
11. Yule JAC, Nielsen WJ. The penetration of light into paper and its effect on halftone reproductions. TAGA Proc 1951;3:65–76.
12. Arney JS, Engeldrum PG, Zeng H. An expanded Murray–Davies model of tone reproduction in halftone imaging. J Imag Sci Tech 1995;39:502–508.
13. Ruckdeschel FR, Hauser OG. Yule–Nielsen effect on printing: a physical analysis. Appl Opt 1978;17:3376–3383.
14. Pearson M. n value for general conditions. TAGA Proc 1980;32:415–425.
15. Shiraiwa Y, Mizuna T. Equation to predict colors of halftone prints considering the optical properties of paper. J Imag Sci Tech 1993;37:385–391.
16. Arney JS, Arney CD, Katsube M, Engeldrum PG. The impact of paper optical properties on hard copy image quality. Proc. IS&T Non-impact Printing 12: Int Conf Digital Print Tech; 1996. p 166–168.
17. Yule JAC, Colt R. Colorimetric investigations in multicolor printing. TAGA Proc 1951;3:77–82.
18. Viggiano JAS. The color of halftone tints. TAGA Proc 1985;37:647–661.
19. Heuberger KJ, Jing ZM, Persiev S. Color transformations and lookup tables. TAGA/ISCC Proc; 1992. p 863–881.
20. Rolleston R, Balasubramanian R. Accuracy of various types of Neugebauer model. Proc IS&T SID Col Imag Conf; 1993. p 32–37.
21. Iino K, Berns RS. Building color management modules using linear optimization I. Desktop color system. J Imag Sci Tech 1998;42:79–94.
22. Iino K, Berns RS. Building color management modules using linear optimization II. Prepress system for offset printing. J Imag Sci Tech 1997;42:99–114.
23. Balasubramanian R. The use of spectral regression in modeling color halftone printers. Rochester, NY: IS&T/OSA Optics Imag Info Age; 1996.
24. Arney JS, Wu T, Blehm C. Modeling the Yule–Nielsen effect on color halftones. J Imag Sci Tech 1998;42:335–340.
25. Kruse B, Wedin M. A new approach to dot gain modeling. TAGA Proc 1995;47:329–338.
26. Engeldrum PG. The color between the dots. J Imag Sci Tech 1994;38:545–551.
27. Gustavson S, Wedin M, Kruse B. 3D modeling of light diffusion in paper. TAGA Proc 1995;47:848–855.
28. Rogers GL. Optical dot gain in a halftone print. J Imag Sci Tech 1997;41:643–656.
29. Maltz M. Light scattering in xerographic images. J App Photo Tech 1983;9:83–89.
30. Engeldrum PG, Pridham B. Application of turbid medium theory to paper spread function measurements. TAGA Proc 1995;47:336–352.
31. Rogers GL. Effect of light scatter on halftone color. J Opt Soc Am A 1998;15:1831–1821.
32. Draper NR, Smith H. Applied regression analysis. New York: Wiley; 1981.
33. Berns RS, Bose A, Tzeng DY. The spectral modeling of large-format ink-get printers. Munsell Color Science Laboratory Internal Report; 1996.