Analysis and Evaluation of Magnetism of Black Toners on Documents Printed by Electrophotographic Systems

A. Biedermann[a,*], S. Bozza[b,a], F. Taroni[a], M. Fürbach[a], B. Li[c], W.D. Mazzella[a]

[a]*University of Lausanne, Faculty of Law, Criminal Justice and Public Administration, School of Criminal Justice, 1015 Lausanne-Dorigny, Switzerland*
[b]*University Ca' Foscari of Venice, Department of Economics, 30121 Venice, Italy*
[c]*China University of Political Science and Law, Institute of Evidence Law and Forensic Sciences, Beijing, China*

Abstract

This paper reports on a study to assess the potential of measurements of magnetism, using a proprietary magnetic analysis system, for the routine analysis of toners on documents printed by black and white electrophotographic systems. Magnetic properties of black toners on documents printed by a number of different devices were measured and compared. Our results indicate that the analysis of magnetism is complementary to traditional methods for analysing black toners, such as FTIR. Further, we find that the analysis of magnetism is realistically applicable in closed set cases, that is when the number of potential printing devices can be clearly defined.

Keywords: Forensic science, Document examination, Magnetic analysis, Black & white electrophotography, Magnetic toner, statistics.

1. Introduction

Due to the wide distribution and availability of business printing machines over the past 40 years, forensic document examiners now commonly face the need to analyze documents produced by electrophotographic printing processes using a dry toner. Black and white electrophotographic printing systems, that is laser printers and/or photocopiers, are part of this class of technology. Nowadays, forensic document examiners routinely examine documents printed by such systems by means of a stereomicroscope, by organic and inorganic chemical analyses or by using image analysis systems. They use these methods when they are requested to help, for example, with the issue of whether or not two or more documents were printed with the same laser printer or the same cartridge unit.

Analyses of toners and electrophotographic printing systems are well documented in forensic literature [1, 2]. Herlaar et al. [3], for example, extensively explored a fast method based on magnetic analysis, providing quantitative measurements. In this paper, we study the complementarity of this technique and traditional methods of analysis, such as FTIR, in view of routine examinations of magnetism of black toners on printed documents.

In order to assess the extent to which measurements of magnetic flux may be used in operational forensic casework, it is necessary to gain an understanding of several aspects related to magnetism of black toners

---

[*]Corresponding author
Email address: `alex.biedermann@unil.ch` (A. Biedermann)

and the measurement of this property. On the one hand, it is necessary to assess whether magnetism can be consistently measured by different operators using a commercially available measuring device. Such understanding is important to decide, for example, if part of the items (e.g., known and questioned) may be analysed by different operators. On the other hand, it is necessary to investigate magnetic properties of printed documents produced by the same and different printing devices (i.e., toner type, brand and model). In this paper, we approach these topics as follows. Section 2 explains the collection of reference documents produced under controlled conditions, that is the model and make of the printing device as well as the toner type (and FTIR properties) known from previous research [5, 6]. This section also introduces the measuring device for magnetism used in the experimental part of the study and the preliminary testing of this device. The results of the analyses of magnetism are presented in Section 3, including a study of prototype evaluative case examples, using established probabilistic interpretation schemes [e.g., 4]. Discussion and conclusions are presented in Section 4.

2. Materials and Methods

2.1. Collection of documents printed by black and white electrophotographic systems

This study used photocopied documents from 61 different multifunction black and white electrophotographic printing systems, which all use magnetic toner (Table 1). These photocopy printouts were collected earlier in 2009 for other research purposes (as part of a field study), not specifically related to the study reported in this paper. Among the 61 printing systems, only two main brands are encountered, that is Canon and Kyocera Mita. Table 1 summarises the different brand and models available, their original toners and FTIR group (named generically by a number), determined as part of previous research [5, 6]. Some combinations of brand/model and toner type were encountered several times. With each device listed in Table 1, the same standard text was photocopied and printed. For various reference purposes, additional documents were printed with a Canon SmartBase PC 1230D copier (see also Sections 2.4.2 and 2.4.3).

2.2. Methods and instrumental techniques used

2.2.1. Fourier Transform Infrared spectroscopy (FTIR)

Fourier Transform Infrared spectroscopy is a suitable method for the analyis of organic binders present in toners. Analyses of toners were performed, as part of previous research (see Tables 1 and 3), by microscopic Attenuated Total Reflectance (ATR) with an Internal Reflection Element (Germanium crystal) using a Digilab$^\circledR$ Excalibur spectrometer, Canton, USA. FTIR spectra were acquired from 4000 to 650cm^{-1} with a resolution of 4cm^{-1} and 64 scans were co-added for each spectrum. For shortness of notation, we summarise FITR group assignments generically by a single number.

2.2.2. Analysis of magnetic properties using a magneto-optical visualizer

Single-component toner powders, which contain magnetic material, incorporated into the toner particles when fixed onto the paper, exhibit magnetic properties similar to other forms of magnetic printing. The magnetic properties, particularly the magnetic flux (nWb), were measured using the Regula Magmouse Model 4197. All the measurements with this magneto-optical visualizer were carried out under standardized and stable conditions, on the same position on a wooden table covered with a glass plate (to avoid interferences with measurements of magnetism). Preliminary test measurements showed that three measures per examined item are adequate [5, 6]. Note that magnetic flux is assumed to be area dependent. The Regula manual mentions that to achieve correct measurements, the measured areas on examined items should be the same. In order to verify this assumption, preliminary analyses as a function of the printed area were conducted as a first step in the experimental procedure (see also Section 2.4.2).

Brand and Model	Toner type	Nb. of devices	FTIR group
Canon iR 2000	Canon Toner C - EXV 5	1	1
Canon iR 2200	Canon Toner C - EXV 3	3	2
Canon iR 2800	Canon Toner C - EXV 3	4	2
Canon iR 3300	Canon Toner C - EXV 3	3	2
Canon iR 2230	Canon Toner C - EXV 11	15	3
Canon iR 2270	Canon Toner C - EXV 11	1	3
Canon iR 2870	Canon Toner C - EXV 11	5	3
Canon iR 3025 N	Canon Toner C - EXV 11	6	3
Canon iR 3225 N	Canon Toner C - EXV 11	1	3
Canon iR 3035 N	Canon Toner C - EXV 12	8	3
Canon iR 3045 N	Canon Toner C - EXV 12	1	3
Canon iR 3530	Canon Toner C - EXV 12	8	3
Canon iR 5000	Canon Toner C - EXV 1	2	4
Kyocera Mita KM 2530	Kyocera/Mita Toner 5PLPXLMAPKX	2	4
Kyocera Mita KM 3530	Kyocera/Mita Toner 5PLPXLMAPKX	1	4

Table 1: Summary of the brand and model, the toner type, and number of distinct printing devices from which photocopied documents were collected for analysis of magnetism. The column on the far right-hand side indicates the FTIR properties as determined in previous research [5, 6].

2.3. Data analyses and study of prototype case examples

The measurements of magnetic flux obtained in this study were summarised using standard descriptive statistical techniques. For the analysis of prototype case examples, likelihood ratio based Bayesian inference methods were used [4, 7] as well as graphical probability models [8].

2.4. Experimental

2.4.1. Main research questions

In this study, analyses have been carried out in order to study the following aspects related to the measurement of magnetic flux on printed documents:

- differences in measurements within a given document depending on, for example, the area of measurement;

- differences in measurements made by different operators, at distinct times, on the same set of printed documents, using the same measuring device;

- differences in measurements carried out on different documents produced under controlled (i.e., known) conditions.

2.4.2. Preliminary testing of measuring device (magneto-optical visualizer)

As noted in Section 2.2.2, the measurement of the magnetic flux with the magneto-optical visualizer used in this study appears to depend on the area. To investigate this aspect, a Canon SmartBase PC 1230D copier was used to print documents featuring repetitions (of different lengths) of the same number '5', using Times New Roman (TNR) 12 pt. The magnetic flux was measured for photocopies including series of 3, 5, 7 and 14 repetitions the number '5'. For example, on a document featuring 3 instances of the number '5',
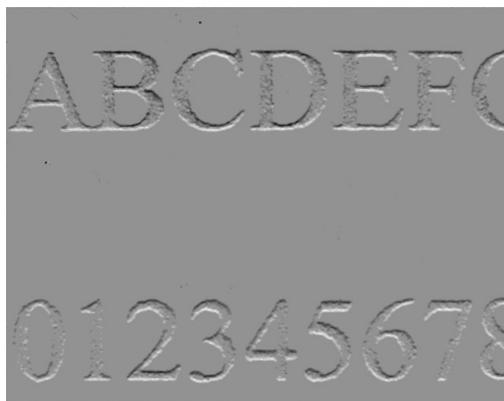
Figure 1: Standard text on which magnetic flux was measured on all documents examined in this study.

the total magnetic flux for these three numbers was measured. Next, the documents on which the magnetic flux for the numbers '5' was measured were scanned at 600 dpi using a Canon® Lide 110 scanner, and measurements of the total area represented by the series of numbers '5' on each document were carried out using the Personal IASLabTM software made by Quality Engineering Associate Inc. (QEA®), Burlington, USA. This enabled a study of the magnetic flux as a function of the area covered by toner.

2.4.3. Analysis of magnetic flux on documents printed under controlled conditions

On all 61 printed documents available in our collection (Section 2.1), the measurement of magnetic flux was carried out on the same printed text which consists of TNR 12 (Figure 1). This combination of font type and font size was chosen because it is regularly encountered in casework and because it is used on all the specimens in our database. The area of this printed text, as captured by the magneto-optical visualizer, corresponds to 512 x 640 pixels. From preliminary tests it was judged that this area presents a good choice to ensure a proper measurement process. One operator measured magnetic flux on all the 61 printed documents on the same day. Two further operators, each of them on a distinct day, repeated this procedure. When going through the 61 documents, a reference document, printed with a Canon iR 2230 model, was measured as an 'internal standard' after every fifth analysed document. This is a measure to monitor the stability of the measurement procedure when processing the collection of 61 specimens from our database.

3. Results

3.1. FTIR spectra

FTIR spectra have been grouped according to similarities in their general aspect, resulting in a limited number of groups. For brevity, these groups have been denoted with a number (see Table 1 and Section 2.2.1). Table 1 shows that some toner types cannot be distinguished on the basis of their FTIR spectra. For example, toner types 'C - EXV 11' and 'C - EXV 12', both from the manufacturer Canon, are found to belong to the same FTIR group. But there are even toners from different manufacturers that exhibit the same FTIR properties. This is the case, for example, with the toners 'C - EXV 1' from Canon and '5PLPXLMAPKX' from Kyocera/Mita. Thus, for situations in which FTIR analyses do not help discriminate between documents, it is of interest to consider further analytical features, such as magnetic flux. This

4

property is chosen in this paper because, from general experience, it is thought to be largely influenced by individual settings of the printing device, so that detectable differences may be expected on documents printed at different instances on the same or different machines (even if they are of the same model and from the same manufacturer).

3.2. Preliminary analyses: magnetic flux as a function of measurement area

Table 2 summarises the mean magnetic flux measured on different sequences of repetitions of the number '5'. These sequences represent different areas covered by black toner. The measurements exhibit a fairly high linear relationship. From these preliminary analyses, confirmed also by other comparable tests[1], we conclude that for the main experimental part of this study, as well as for any future uses in operational casework, it will be important the select constant areas of measurement in order to ensure comparability between measurements made on different documents.

Measured text	Area (mm^2)	Mean magnetic flux measured	Magnetic flux per mm^2
55555555555555	25.32	14.00	0.55
5555555	12.58	7.33	0.58
55555	8.76	6.00	0.68
555	5.57	3.00	0.54

Table 2: Results of the preliminary analyses of magnetic flux as a function of the area of measurement.

3.3. Magnetic properties of documents printed under controlled conditions: comparison of measurements from different operators

The question dealt with regarding magnetic flux in this study is not one of calibration where one measures *known* quantities, because for each of the documents printed under controlled conditions (Section 2.1), the true value of magnetic flux is unknown. Instead, the focus here is on the agreement between measurements made by different operators. A first step to examine the data is to plot results for pairs of operators as shown in Figure 2, in addition to a straight line which shows where observations would lie if the compared operators would produce exactly the same measurements. Operators 1 and 2 in this study appear to produce values that correspond fairly close, as they scatter densely around the straight line. In turn, operator 3 exhibits a tendency towards higher values when compared to operators 1 and 2, although these values are still close to the straight line. Note that the corresponding correlation coefficient is high. In both statistical and medical literature [9, 10, 11] is has been pointed out, however, that this way of displaying the data may hide some aspects, such as how much the measurements differ from another, that is between-operator differences. To help explore this issue, it is useful to plot – for each examined document – the difference between the mean of replicate measurements by different operators against the mean of those measurements. This type of plot, known also as Bland-Altman or Tukey mean-difference plot, is shown in Figure 3. This way of representing the data reveals considerable disagreements that are not readily conveyed by Figure 2. For example, operators 1 and 3 diverge up to approximately four units of measurement, as shown by the datapoint in the lower right-hand side of the plot in the middle of Figure 3. Similar observations hold for the comparison between operators 2 and 3. Generally, the lack of agreement is summarised by calculating the

---

[1]Full data records are available from the corresponding author upon request.
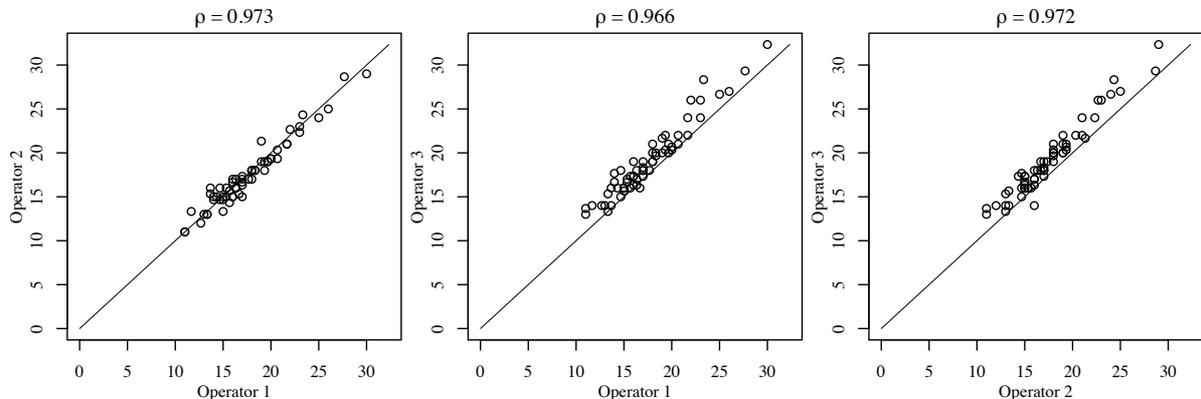
Figure 2: Comparison of the measurements of magnetic flux made by three different operators. For each couple of operators, the correlation coefficient ρ is reported.

mean of operator differences, shown in Figure 3 in terms of the dashed line, as well as the standard deviation of these differences (dotted line). Note that the latter quantity, $\hat{\sigma}_d$, has been obtained by means of the correction of the standard deviation of differences between means of several repeats (i.e., $s_{\bar{d}}$), as proposed by Bland and Altman [11].[2]

These considerations, however, are only descriptive. They do not answer the question of how much discrepancy is acceptable, which is a criterion that should be defined in advance, independently of actual measurements. In view of the observations made in this section, it seems advisable that documents of a given case ought to be examined by the same operator.

3.4. Distribution of magnetic flux on documents printed under controlled conditions: statistical model and notation

The available measurements are denoted by x_{ij}, where i refers to the analysed page and j refers to the measurement repeat number. It is assumed that measurements are normally distributed with mean θ and variance σ^2, $(\bar{X}_i \mid \theta) \sim N(\theta, \frac{\sigma^2}{n})$, where n represents the total number of repeated measurements made on page i. The parameter θ is supposed here to follow a normal distribution with prior mean μ and prior variance τ^2, written $\theta \sim N(\mu, \tau^2)$.

3.5. Case example 1: Alleged page substitution

3.5.1. Case description and definition of probabilistic model

To illustrate how measurements of magnetic flux of black toner on printed documents may help discriminate between competing propositions of interest, it is instructive to analyse and discuss exemplary cases.

---

[2]According to this, and considering two hypothetical operators, 1 and 2, the standard deviation of differences between single measurements is estimated by

$$\hat{\sigma}_d^2 = s_{\bar{d}}^2 + \left(1 - \frac{1}{n_1}\right) s_{1w}^2 + \left(1 - \frac{1}{n_2}\right) s_{2w}^2, \tag{1}$$

where $s_{\bar{d}}^2$ represents the variance of the differences between means of several repeats, and s_{1w}^2 and s_{2w}^2 represent the observed within-document variances from measurements by operator 1 and by operator 2, respectively, while n_1 and n_2 represent the number of repeated measurements.

6
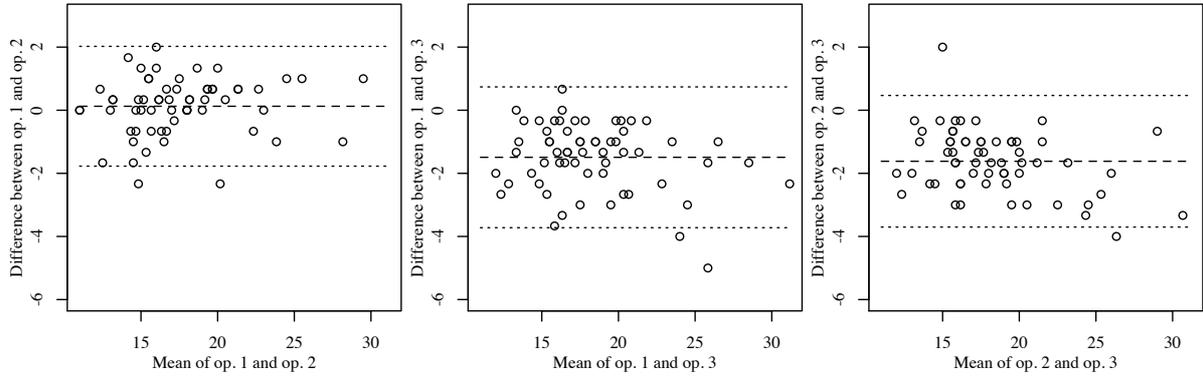
Figure 3: Plots of the difference between the mean of the replicate measurements of magnetic flux by distinct operators against the mean of those measurements for a total of 61 documents printed under controlled conditions. On each of the 61 documents, each operator has performed three measurements. In each plot, the dashed line (− −) represents the mean of the differences between averaged values for individual documents obtained by each operator, whereas the dotted lines (· · ·) represent the mean ± two times the standard deviations of the differences between single measurements (Equation (1)).

One general class of practical problems relates to page substitutions, focusing on – for example – analytical features of a questioned page when compared to the analytical features of the uncontested page(s).

Suppose a case involving a contract that consists of three pages. The pages one and three are not contested, but page two is disputed. The main issue in the case is the allegation that page two has been substituted. The forensic scientist may thus be interested in the extent to which the measurement of magnetic flux can help with the issue. The competing propositions can thus be defined as follows:

H_p: Page two has been printed by the same device as the one used for pages one and three;

H_d: Page two has been printed by a different device.

Denote by $\mathbf{x}_1 = (x_{11}, x_{12}, x_{13})$ and by $\mathbf{x}_2 = (x_{21}, x_{22}, x_{23})$ the measurements of magnetic flux obtained for the uncontested pages one and three, and by $\mathbf{y} = (y_1, y_2, y_3)$ the measurements of magnetic flux obtained for the questioned page 2. The value of the evidence (V) is the ratio of the marginal likelihood of the triplet $(\mathbf{x}_1, \mathbf{x}_2, \mathbf{y})$ under the competing propositions, that is:

$$\mathrm{V} = \frac{f(\mathbf{x}_1, \mathbf{x}_2, \mathbf{y} \mid H_p)}{f(\mathbf{x}_1, \mathbf{x}_2, \mathbf{y} \mid H_d)} = \frac{f(\mathbf{y} \mid \mathbf{x}_1, \mathbf{x}_2, H_p)}{f(\mathbf{y} \mid \mathbf{x}_1, \mathbf{x}_2, H_d)} \times \frac{f(\mathbf{x}_1, \mathbf{x}_2 \mid H_p)}{f(\mathbf{x}_1, \mathbf{x}_2 \mid H_d)} \tag{2}$$

$$= \frac{f(\mathbf{y} \mid \mathbf{x}_1, \mathbf{x}_2, H_p)}{f(\mathbf{y} \mid H_d)}. \tag{3}$$

From the case circumstances it appears reasonable to make the assumption that if all three pages have been printed by the same machine, that is page two is not substituted,[3] the toner present on the three pages will have corresponding analytical features (in some sense), such as magnetic flux. Therefore, in the numerator the question to be asked by the forensic scientist thus is: 'What is the probability density

---

[3]It is thus supposed in this case that substitution of a page printed with the same machine as used for pages one and two is not an issue.

for observing the triplet[4] $\mathbf{y} = (y_1, y_2, y_3)$ of magnetic flux obtained for page two, given the measurements $\mathbf{x}_1$ and $\mathbf{x}_2$ of magnetic flux from pages one and three, and given the proposition H_p that page two is not substituted?' Note that the conditioning information I is omitted for shortness of notation. Starting from the statistical model outlined in Section 3.4, the marginal probability density in the numerator can be derived as

$$f(\mathbf{y} \mid \mathbf{x}_1, \mathbf{x}_2, H_p) = \int_\Theta f(\mathbf{y} \mid \theta) f(\theta \mid \mathbf{x}_1, \mathbf{x}_2) \, d\theta, \tag{4}$$

where $f(\mathbf{y} \mid \theta) = N\left(\bar{y} \mid \theta, \frac{\sigma^2}{n}\right)$ and $f(\theta \mid \mathbf{x}_1, \mathbf{x}_2) = N\left(\theta \mid \mu_x, \tau_x^2\right)$ is the posterior distribution for θ with hyperparameters that have been calculated according to the well known updating rules

$$\mu_x = \frac{\mu \sigma^2/n + \bar{x}\tau^2}{\tau^2 + \sigma^2/n} \qquad \text{and} \qquad \tau_x^2 = \frac{\tau^2 \sigma^2/n}{\tau^2 + \sigma^2/n}, \tag{5}$$

with $\bar{x} = \frac{\bar{x}_1 + \bar{x}_2}{2}$. Note that Bayes' theorem enables a sequential update of the uncertainty about the parameter θ as more observations become available. The posterior distribution that has been obtained by a single application of Bayes' theorem to the entire set of data, is the same as the one obtained by a two stage process where the prior distribution is first updated by incorporating available measurements on page 1, $\mathbf{x}_1$, and then successively updated by incorporating available measurements on page 3, $\mathbf{x}_2$. Standard calculation allows one to solve the integral in (4) and to derive the marginal probability density that is still normal with mean equal to the prior mean and variance equal to the sum of the variances, that is $f(\mathbf{y} \mid \mathbf{x}_1, \mathbf{x}_2, H_p) = N\left(\bar{y} \mid \mu_x, \tau_x^2 + \sigma^2/n\right)$.

In turn, if page two has been substituted, that is printed with an unknown printing device (i.e., different from the machine used for printing pages one and three), the analytical features of the toner present on page two will be *independent* from those of the toner present on the other two pages. In other words, the question in the denominator to be asked by the forensic scientist is: 'What is the probability density for observing the triplet y_1, y_2 and y_3 of magnetic flux measurements obtained for page two, given the proposition H_d that page two has been substituted?' More formally, this probability density can be obtained as

$$f(\mathbf{y} \mid H_d) = \int_\Theta f(\mathbf{y} \mid \theta) f(\theta) \, d\theta. \tag{6}$$

Standard calculations allow one to solve the integral in (6), as it was done for the numerator, and to obtain the marginal distribution that is still normal with mean equal to the prior mean μ and variance equal to the sum of variances, $f(\mathbf{y} \mid H_d) = N(\bar{y} \mid \mu, \tau^2 + \sigma^2/n)$.

The probabilistic assumptions underlying the computations described in the above two steps can be visually summarised in terms of a Bayesian network [8] as shown in Figure 4. Combining the answers to the above two questions, that is values for $f(\mathbf{y} \mid \mathbf{x}_1, \mathbf{x}_2, H_p)$ and $f(\mathbf{y} \mid H_d)$, provides the numerator and denominator of the likelihood ratio V, the probative strength of the measurements of magnetic flux.

3.5.2. Numerical example

Suppose that in the case introduced earlier in Section 3.5.1 the scientist obtains the measurements $\mathbf{y} = (15, 15, 16)$ for the questioned page two, and measurements $\mathbf{x}_1 = (16, 15, 16)$ for the undisputed page one, and $\mathbf{x}_2 = (16, 15, 15)$ for the undisputed page three. The measurement means thus are $\bar{y} = 15.33$, $\bar{x}_1 = 15.67$ and $\bar{x}_2 = 15.33$. To obtain a value for the numerator of the likelihood ratio, the following assignments are made:

---

[4]It is supposed here that the scientist will make three measurements on each examined document, irrespective of the document being questioned or not.
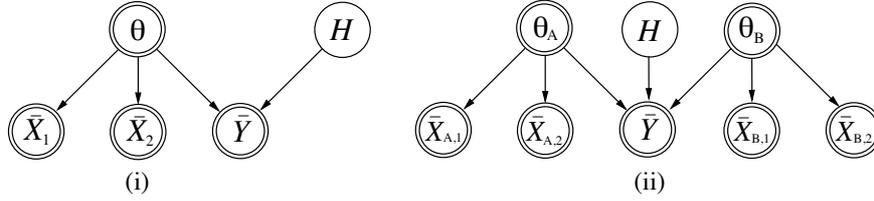
Figure 4: (i) Case example 1 (Section 3.5): Bayesian network for illustrating the probabilistic modelling assumptions made for evaluating measurements of magnetic flux in a case of suspected page substitution (i.e., a contract that consists of three pages, the first and the third pages being uncontested and the second page suspected to be substituted). The nodes with double border represent continuous variables, that is: node θ represents the mean magnetic flux of documents produced by the printing machine used for printing pages one and three (i.e., uncontested pages), nodes $\bar{X}_1$ and $\bar{X}_2$ represent the mean magnetic flux measured on the uncontested pages one and three, node $\bar{Y}$ represents the mean magnetic flux measured on page two (i.e., the questioned page). Node H is discrete and has two states H_p ('Page two has not been substituted') and H_d ('Page two has been substituted'). (ii) Case example 2 (Section 3.6): Bayesian network for illustrating the probabilistic modelling assumptions made for evaluating mean measured magnetic flux $\bar{Y}$ in a case involving a single questioned page, with uncertainty about whether the page has been printed by machine A (H_p) or machine B (H_d). Nodes $\bar{X}_{(\cdot),i}$ represent results for measurements on pages $i = \{1, 2\}$ printed under controlled conditions with machines A and B, whereas the nodes $\theta_{(\cdot)}$ represent the mean magnetic flux of documents produced by the two machines A and B.

- Prior distribution for θ, the mean magnetic flux of documents produced by the machine used for printing pages one and three: the prior mean μ and variance τ^2 are chosen on the basis of the population study conducted in this paper, which provides the values 17.60 and 3.91, respectively. A normal distribution $N(17.60, 3.91^2)$, show in Figure 5(i), translates the view that values below 10 and above 25 are very uncommon. Note that this choice requires the area of measurement to be the same as the one covered by the standard text (Figure 1), because the magnetic flux depends on the measured area (see also findings reported in Section 3.2).

- Standard deviation σ of the measurement method: based on a large number of previous experiments with the device for measuring magnetic flux, the standard deviation is assumed to be constant and equal to 0.23.

- Posterior distribution for θ: Using the assignments made above, and the Bayesian updating procedure described in Section 3.5.1, leads to the posterior distribution $\theta \sim N(15.50, 0.0088)$, shown in Figures 5(i) and (ii) (solid line).

Based on the posterior distribution for θ, one obtains the marginal distribution for the mean of measurements made on page two, that is $(\bar{Y} \mid \bar{x}_1, \bar{x}_2, H_p)$ which is given by $N(15.50, 0.0088 + (0.23^2/3))$ shown in Figure 5(iii). The density that corresponds to the measured mean $\bar{y} = 15.33$ is 1.44 (see dotted line Figure 5(iii)). The numerator of the likelihood ratio thus is $f(\bar{y} \mid \bar{x}_1, \bar{x}_2, H_p) = 1.44$.

For the denominator, the marginal distribution for the mean measurements made on page two, that is $(\bar{Y} \mid H_d)$ is given by $N(17.60, 3.91^2 + (0.23^2/3))$. The likelihood ratio V thus is:

$$\text{V} = \frac{f(\bar{y} \mid \bar{x}_1, \bar{x}_2, H_p)}{f(\bar{y} \mid H_d)} = \frac{f(15.33 \mid 15.67, 15.33, H_p)}{f(15.33 \mid H_d)} = \frac{1.44}{0.086} \approx 16.7.$$

This result can readily be obtained with a computerized implementation of the Bayesian network defined in Figure 4(i). Figure 6 illustrates that the same likelihood ratio (16.7) is found.

9
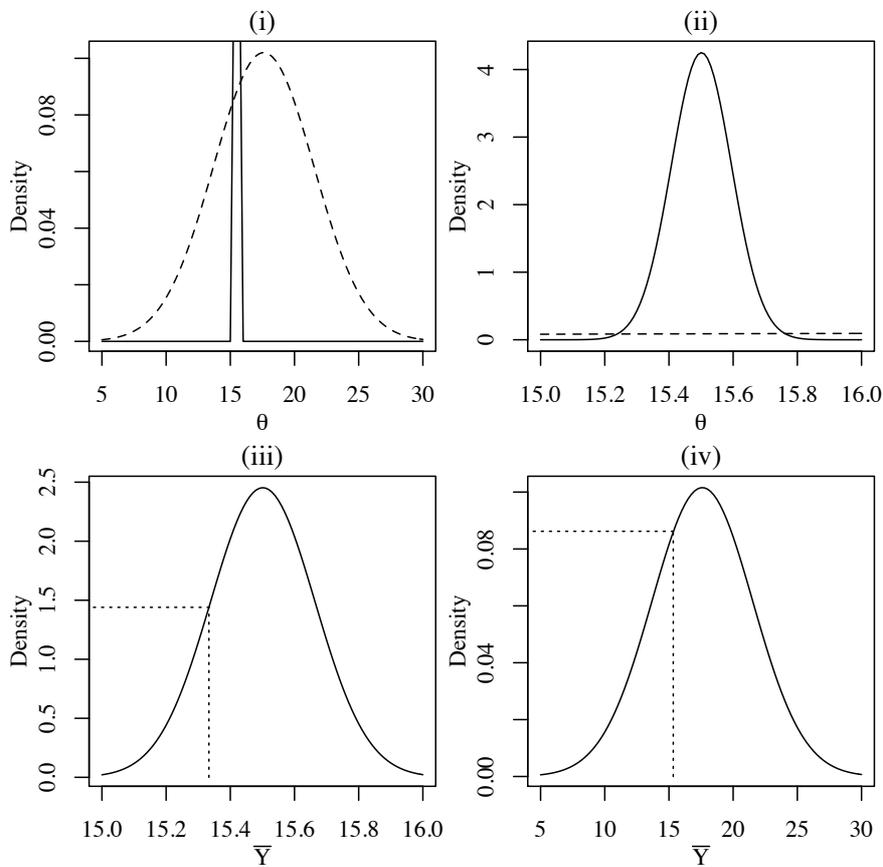
Figure 5: (i) and (ii): Different representations of the Normal prior (dashed line) and posterior (solid line) distributions for the mean magnetic flux θ of documents printed by the machine used for generating documents one and three in the numerical example discussed in Section 3.5.2. (iii) and (iv) Marginal distributions at the numerator (iii) and denominator (iv) for the mean magnetic flux $\bar{Y}$ on the questioned page two (example from Section 3.5.2) with the dotted line indicating the density corresponding to the measured value $\bar{y} = 15.33$.
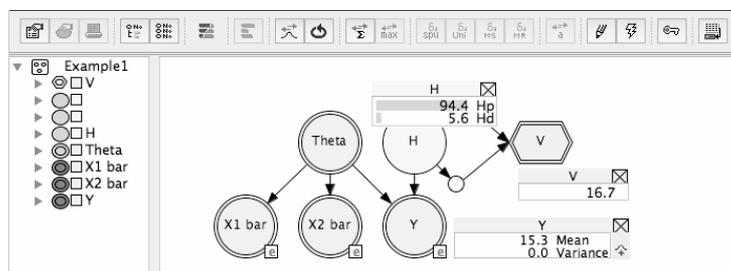
Figure 6: Illustration of a computerized implementation of the Bayesian network described in Figure 4(i), using the software package Hugin (www.hugin.com), and the propagation of the findings of the numerical example presented in Section 3.5.2, including the likelihood ratio of 16.7.

10

3.6. Case example 2: Discrimination in a closet set

Suppose a case involving a single document, such as a contract, and the issue of interest is which of two[5] machines has been used to print the questioned document. Measurement of magnetic flux on the questioned document leads to the following results: $y_1 = 21$, $y_2 = 20$ and $y_3 = 20$. The two potential sources, printer A and printer B, are available for examination. The propositions of interest can be specified as:

H_p: The questioned document has been printed with machine A;

H_d: The questioned document has been printed with machine B.

Suppose the two machines A and B are used to print documents under controlled conditions, and that the forensic scientist selects two documents printed by each machine and performs three measurements of magnetic flux. The results are denoted $x_{.,i,j}$ with $i = \{1, 2\}$ denoting the page number for each printer and $j = \{1, 2, 3\}$ denoting the measurement number. Assume the following results:

Printer A: $x_{A,1,1} = 20$ $x_{A,1,2} = 19$ $x_{A,1,3} = 20$ $x_{A,2,1} = 20$ $x_{A,2,2} = 20$ $x_{A,2,3} = 21$
Printer B: $x_{B,1,1} = 22$ $x_{B,1,2} = 20$ $x_{B,1,3} = 21$ $x_{B,2,1} = 20$ $x_{B,2,2} = 21$ $x_{B,2,3} = 22$

Based on the probabilistic model defined above, the likelihood ratio can be defined as follows:

$$V = \frac{f(\mathbf{y} \mid \mathbf{x}_A, H_p)}{f(\mathbf{y} \mid \mathbf{x}_B, H_d)}.$$

The marginal probability density in the numerator (i.e., if the questioned page has been printed with printer A, that is hypothesis H_p is true) can be obtained as

$$f(\mathbf{y} \mid \mathbf{x}_A, H_p) = \int_\Theta f(\mathbf{y} \mid \theta_A) f(\theta \mid \mu_A, \tau_A^2) d\theta,$$

where $f(\theta \mid \mu_A, \tau_A^2) = N(\mu_A, \tau_A^2)$ is the posterior distribution for θ_A, with hyperparameters μ_A and τ_A^2 obtained according to the updating rules in (5) using the results obtained for two pages printed with machine A under controlled conditions. The marginal probability density is $f(\mathbf{y} \mid \mathbf{x}_A, H_p) = N(\bar{y} \mid \mu_A, \tau_A^2 + \sigma^2/n_y)$, where n_y is the number of measurements made on questioned page (i.e., 3 in case here), σ^2 is the variance of the measuring method as defined in the previous example. The means for the two pages printed under controlled conditions are $\bar{x}_{A,1} = 19.67$ and $\bar{x}_{A,2} = 20.33$. The posterior parameters μ_A and τ_A^2 are 20.00 and 0.0088 (values rounded). The numerator can thus be found as $f(\mathbf{y} \mid \mathbf{x}_{A,1}, \mathbf{x}_{A,2}, H_p)$, that is the density for $\bar{y} = 20.33$ of the Normal distribution $N(20.00, (0.23^2/3) + 0.0088)$, which is 0.2950. In the same way, one can obtain the marginal probability density in the denominator $f(\mathbf{y} \mid \mathbf{x}_B, H_d) = N(\bar{y} \mid \mu_B, \tau_B^2 + \sigma^2/n_y)$. The means for the two pages printed under controlled conditions are $\bar{x}_{B,1} = 21$ and $\bar{x}_{B,2} = 21$. The posterior parameters μ_B and τ_B^2 are 21 and 0.0088 (values rounded). The denominator can thus be found as $f(\mathbf{y} \mid \mathbf{x}_{B,1}, \mathbf{x}_{B,2}, H_d)$, that is the density for $\bar{y} = 20.33$ of the Normal distribution $N(21, (0.23^2/3) + 0.0088)$, which is 0.00057.

Combining the above results thus leads to a likelihood ratio V on the order of 500 in favour of the proposition H_p ('The questioned document has been printed with machine A'), rather than H_d ('The document was printed with machine B'). This result can also be tracked in a Bayesian network, shown in Figure 7.

Figure 8 shows the marginal distributions for $\bar{Y}$ given H_p and H_d. As may be seen, the distributions are rather neatly separated, which means that the likelihood ratio may be very sensitive to changes in observed result $\bar{y}$. To illustrate this, suppose that instead of the findings $y_1 = 21$, $y_2 = 20$ and $y_3 = 20$ (i.e.,

---

[5]The approach here is general and can be extended to situations in which there are more than two potential sources.

$\bar{y} = 20.33$), the findings would have been $y_1 = 21$, $y_2 = 20$ and $y_3 = 21$, leading to a mean of $\bar{y} = 20.67$. This result would lead to a numerator of 0.00053 and a denominator of 0.3077, which gives a likelihood ratio of 0.0017, that is a likelihood ratio on the order of 580 in favour of the alternative proposition H_d over H_p (result not shown here in terms of Bayesian network). So, there is not only a change in the magnitude of the likelihood ratio, but even a directional change of evidential support (i.e., another proposition is being supported). The likelihood ratio thus crucially depends on the properties of the distributions for $\bar{Y}$ constructed on the basis of measurements of magnetic flux performed on the pages printed under controlled conditions with machines A and B. In view of this result, it is advisable that such sensitivity analyses be conducted in actual cases because, potentially, the likelihood ratios obtained with this probabilistic model may be large, in particular when the distributions for $\bar{Y}$ are well separated.

4. Discussion and conclusions

The results obtained in this paper point out the potential of analyses of magnetic flux as an additional technique for forensic document examination. Our findings agree with the results obtained by Herlaar et al. [3] and are complementary to those obtained by FTIR. For example, the 15 reference documents from distinct Canon iR 2230 models, using the same type of toner (i.e., Canon C-EXV 11), were classified into the same group 3 according to the FTIR analysis but could be shown to exhibit considerable variation regarding magnetic flux.

A limitation of this technique is its dependence on software, hardware and time. As indicated by Flynn [12, at p. 194], "[t]he modern FDE must be aware of such things as the operating system, the word processing program, the version of the word processing program, the digital font and version of the font file, the printer driver, (...). All can have an impact on the appearance of the printed text, as can the many variables associated with the printer itself". The preliminary study reported in this paper focused only on the printer variable, that is on the different output documents that can be generated by different laser printers.

However, it is important to keep in mind that the printer cartridge and/or toner type can lead to considerable measurable differences. For example, Tse et al. [13] were able to show that character width is strongly related to electrophotographic properties of the OPC drums and to the type of toner used. Two brand new identical OPC drums may have different photosensitivity properties that can lead to stroke width differences. Similarly, a different batch of toner can drastically change the output of a given laser printer. Such differences in output should lead to detectable differences in magnetic flux. Care is thus needed when evaluating cases in which it is alleged that one or more pages of, for example, a contract were prepared in distinct ways (i.e., using different printing devices) or when it is supposed that part of a text has been altered.

It is also worth noting that toner present on printed documents may undergo detectable changes due to storage and other exposure conditions, such as heat and humidity, in analogy to what is known from forensic ink analyses, though toner resins tend to be more stable than volatile compounds such as inks. In the research here, all toner printed specimens have been preserved under the same conditions (i.e., in a binder stored in the laboratory). We take it that before extending empirical investigations to further factors, it is important to conduct a study under controlled and stable laboratory conditions. There is much room, however, for further research on influencing factors. The latter may be chosen according to the needs encountered in practical cases, for example when it is alleged that questioned and known items have been exposed to different conditions.

The standard probabilistic approach, using parametric distributions, for evaluating measurements is readily applicable in cases where the aim is to discriminate between potential sources (i.e., printing devices), because the marginal distributions under the two competing propositions can be worked out on the basis
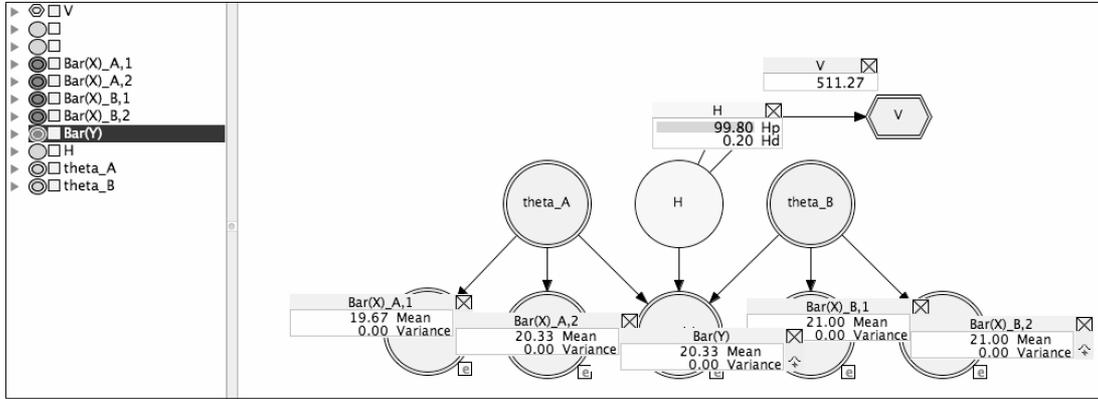
Figure 7: Illustration the computations for case example 2, using the Bayesian network defined in Figure 4(ii), implemented with the software package Hugin (www.hugin.com). Observations $\mathbf{x}_{A,1}$, $\mathbf{x}_{A,2}$, $\bar{y}$, $\mathbf{x}_{B,1}$ and $\mathbf{x}_{B,2}$ are entered in the bottom layer of nodes (from left to right). The node H displays the posterior probabilities for the main propositions, whereas the function node V provides the likelihood ratio (on the order of 500).
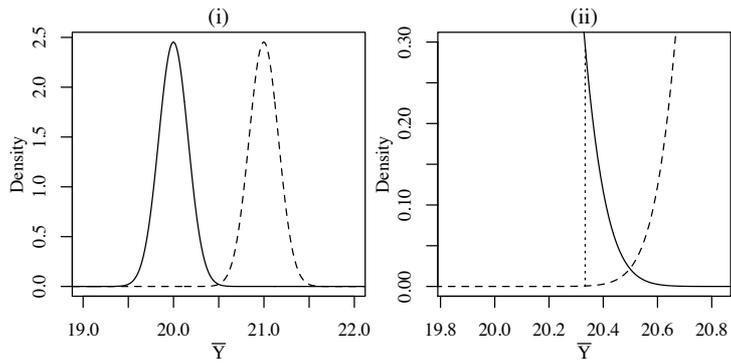
Figure 8: Marginal probability distributions for the mean magnetic flux on a questioned document given H_p (solid line) and given H_d (dashed line) as discussed in case example 2 (Section 3.6). The dotted line in Figure (ii) shows the density value corresponding to result $\bar{y} = 20.33$.

of measurements made on exemplar printing outputs made with the known potential sources. In cases requiring inference of source, when the alternative proposition is that an unknown printing device was used to produce the questioned item, the marginal distributions under the competing propositions may not be immediately available. The main reason for this is that it must be ensured that the area of measurement on the questioned item is comparable to the area measured on the documents used to build the Bayesian statistical model that was proposed in Section 3.4. More generally, however, it is important to keep in mind that evaluation does not reduce to a consideration of measurements alone, but requires a broader view of all relevant factors, as noted also in the above paragraph.

Acknowledgements

References

[1] A Biedermann, S Bozza, F Taroni, and WD Mazzella. Implementing statistical learning methods through Bayesian networks (Part II): Bayesian evaluations for results of black toner analyses in forensic document examination. *Forensic Science International*, 204:58–66, 2010.

[2] R A Merrill, E Bartick, and W D Mazzella. Studies of techniques for analyzing photocopy toners by IR. *Journal of Forensic Sciences*, 41:264–271, 1996.

[3] K Herlaar, M Mieremet, and M Fakkel. Measuring magnetic properties to discriminate between different laser printers. *Journal of the American Society of Questioned Document Examiners*, 18:51–66, 2016.

[4] C G G Aitken and F Taroni. *Statistics and the Evaluation of Evidence for Forensic Scientists*. John Wiley & Sons, Chichester, second edition, 2004.

[5] N Meyer and W D Mazzella. Geo-forensic analysis of photocopy toners. In *68th Annual Meeting of the American Society of Questioned Document Examiners*, Victoria, BC, Canada, 2010.

[6] N Meyer and W D Mazzella. Geo-forensic analysis of photocopy toners. In *6th Conference of the European Document Experts Working Group (EDEWG)*, Dubrovnik, Croatia, 2010.

[7] F Taroni, S Bozza, A Biedermann, G Garbolino, and C G G Aitken. *Data Analysis in Forensic Science: a Bayesian Decision Perspective*. Statistics in Practice. John Wiley & Sons, Chichester, 2010.

[8] F Taroni, A Biedermann, S Bozza, G Garbolino, and C G G Aitken. *Bayesian Networks for Probabilistic Inference and Decision Analysis in Forensic Science*. Statistics in Practice. John Wiley & Sons, Chichester, second edition, 2014.

[9] D G Altman and J M Bland. Measurement in medicine: The analysis of method comparison studies. *Journal of the Royal Statistical Society. Series D (The Statistician)*, 32:307–317, 1983.

[10] J M Bland and D G Altman. Statistical methods for assessing agreement between two methods of clinical measurement. *The Lancet*, 327:307–310, 1986.

[11] J M Bland and D G Altman. Measuring agreement in method comparison studies. *Statistical Methods in Medical Research*, 8:135–160, 1999.

[12] W J Flynn. The examination of computer-generated documents. In J S Kelly and B S Lindblom, editors, *Scientific Examination of Questioned Documents*, pages 191–216. CRC Press, Boca Raton, second edition, 2006.

[13] MK Tse, DJ Forrest, and KY She. Use of an automated print quality evaluation system as a failure analysis tool in electrophotography. In *Eleventh International Congress on Advances in Non-Impact Printing Technology*, Hilton Head, South Carolina, 1995.

No.	Operator 1 Mean	SD	Operator 2 Mean	SD	Operator 3 Mean	SD	FTIR type
1	13.7	0.6	16.0	0.0	14.0	0.0	3
2	19.0	0.0	21.3	0.6	21.7	0.6	3
3	22.0	0.0	22.7	0.6	26.0	0.0	3
4	16.7	0.6	15.3	0.6	16.0	0.0	3
5	26.0	0.0	25.0	0.0	27.0	0.0	2
6	15.7	0.6	15.7	0.6	16.0	0.0	3
7	16.3	0.6	16.0	0.0	17.0	0.0	3
8	27.7	0.6	28.7	0.6	29.3	0.6	2
9	17.0	0.0	16.7	0.6	18.0	0.0	3
10	12.7	0.6	12.0	0.0	14.0	0.0	3
11	16.0	0.0	15.0	0.0	16.3	0.6	3
12	15.0	0.0	14.7	0.6	16.0	0.0	3
13	15.3	0.6	16.0	0.0	17.0	0.0	4
14	18.0	0.0	18.0	0.0	19.0	0.0	4
15	17.0	0.0	17.3	0.6	19.0	0.0	3
16	20.0	0.0	19.3	0.6	20.3	0.6	3
17	17.7	0.6	17.0	0.0	18.0	0.0	3
18	15.0	0.0	15.0	0.0	16.0	0.0	3
19	14.0	0.0	15.0	0.0	16.7	0.6	3
20	30.0	0.0	29.0	0.0	32.3	0.6	2
21	19.7	0.6	19.0	0.0	21.0	0.0	3
22	21.7	0.6	21.0	0.0	22.0	0.0	3
23	23.0	0.0	23.0	0.0	26.0	0.0	2
24	20.7	0.6	20.3	0.6	22.0	0.0	2
25	21.7	0.6	21.0	0.0	24.0	0.0	2
26	13.3	0.6	13.0	0.0	13.3	0.6	4
27	16.3	0.6	17.0	0.0	18.0	0.0	4
28	14.7	0.6	14.7	0.6	15.0	0.0	3
29	18.0	0.0	18.0	0.0	20.0	0.0	3
30	18.0	0.0	17.0	0.0	19.0	0.0	3
31	13.0	0.0	13.0	0.0	14.0	0.0	3
32	13.7	0.6	15.3	0.6	16.0	0.0	4
33	17.0	0.0	16.3	0.6	18.0	0.0	3
34	25.0	0.0	24.0	1.0	26.7	0.6	2
35	18.3	0.6	18.0	0.0	20.0	0.0	3
36	19.0	0.0	19.0	1.0	20.0	0.0	3
37	13.3	0.6	13.0	0.0	15.3	0.6	3
38	17.0	0.0	15.0	0.0	17.3	0.6	3
39	16.0	0.0	16.7	0.6	19.0	0.0	3
40	11.0	0.0	11.0	0.0	13.0	0.0	3
41	16.3	0.6	16.0	0.0	16.3	0.6	3
42	15.0	0.0	13.3	0.6	15.7	0.6	3
43	19.7	0.6	19.0	1.0	20.0	0.0	3
44	23.3	0.6	24.3	0.6	28.3	0.6	2
45	19.3	0.6	19.0	0.0	22.0	0.0	2
46	15.7	0.6	14.3	0.6	17.3	0.6	3
47	14.3	0.6	15.0	0.0	16.0	0.0	3
48	20.0	0.0	19.3	0.6	20.7	0.6	3
49	11.7	0.6	13.3	0.6	14.0	0.0	3
50	14.7	0.6	16.0	0.0	18.0	0.0	3
51	14.0	0.0	14.7	0.6	17.7	0.6	3
52	16.0	0.0	17.0	0.0	17.3	0.6	3
53	18.3	0.6	18.0	0.0	19.7	0.6	3
54	18.0	0.0	18.0	0.0	21.0	0.0	3
55	19.3	0.6	18.0	0.0	20.3	0.6	1
56	15.3	0.6	15.0	0.0	16.7	0.6	3
57	23.0	0.0	22.3	0.6	24.0	0.0	2
58	11.0	0.0	11.0	0.0	13.7	0.6	3
59	16.0	0.0	15.0	0.0	17.3	0.6	3
60	20.7	0.6	19.3	0.6	21.0	1.0	3
61	17.0	0.0	17.0	0.0	18.3	0.6	3

Table 3: Mean and standard deviation (SD) of measured magnetic flux by three different operators on each of 61 documents printed under controlled conditions. Each operator performed three measurements per document. Note that the measuring device only gives integer numbers as an output reading. The column on the far right-hand side indicates the toner type as determined in previous research [5, 6].