Assessing, exploring, and monitoring quality of offset colour prints

J. Lundström a,⇑, A. Verikas a,b, E. Tullander c, B. Larsson d

a Intelligent Systems Laboratory, Halmstad University, Box 823, SE-301 18 Halmstad, Sweden
b Department of Electrical & Control Equipment, Kaunas University of Technology, Studentu 50, LT-51368 Kaunas, Lithuania
c Quality Development Department, Hylte Mill, SE-314 81 Hyltebruk, Sweden
d V-TAB, Exportgatan 2-4, SE-422 46 Hisingsbacka, Sweden

ABSTRACT

Variations in offset print quality relate to numerous parameters of printing press and paper. To maintain a constant high print quality press operators need to assess, explore and monitor quality of prints. Today assessment is mainly done manually. This paper presents a novel system for assessing and predicting values of print quality attributes, where the adopted, random forests (RFs)-based, modeling approach also allows quantifying the influence of different paper and press parameters on print quality. In contrast to other print quality assessment systems the proposed system utilises common, simple print marks known as double grey-bars. Novel virtual sensors assessing print quality attributes using images of double grey-bars are presented. The inferred influence of paper and printing press parameters on quality of colour prints shows clear relation with known print quality conditions. Thorough analysis and categorisation of related work is also given in the paper.

1. Introduction

Offset lithographic printing is the most widely used technique for newspaper printing. Four primary inks, cyan (C), magenta (M), yellow (Y) and black (K), are used to produce colour images in lithographic offset four-colour printing. A roller clothed with a rubber blanket, a steel roller sprayed with fountain solution mainly consisting of water, a roller equipped with a printing plate and finally the paper, see Fig. 1, are the main elements in offset printing.

There are four printing plates, one for each primary colour. Printing plates are processed so that there appear areas with different surface properties, areas attracting water and areas repelling water. These surface properties allow the ink roller to transfer ink onto dry areas of the printing plates. The ink is then transferred to the rubber blanket which in turn transfers the ink onto the paper. A specific balance of the amount of ink and fountain solution has to be kept to maintain even quality throughout the whole print job. Other than this balance, parameters as paper properties, paper web tension, ink recipe, air humidity and temperature, ink temperature, wear degree of the printing plates also affect the print quality [1].

Print quality is usually related to a number of parameters (print quality attributes) and tolerances. In this work, we assume that ultimate assessors of print quality are humans. We use print quality attributes that can be objectively evaluated on a small printed area online and can be integrated into an overall print quality measure providing quality evaluations correlating with print quality assessments given by observers.

To maintain constant printing quality, the press operator samples the print manually throughout the job run and a great effort is made to compensate for colour deviations detected in the print. Usually, the initial assessment is visual. If deviations, noticeable to the operator eye, are detected, a densitometer or a spectrophotometer is then applied to get a numerical evaluation of the colour deviations and adjustments are performed to compensate for
the deviations. Each operator performs the necessary adjustments based on experience gained from working at that particular press. Typically, the perception of the printed result is different for different operators and consequently variations may appear in the print depending on the operator controlling the process. This problem and the difficulty to print within standards such as ISO 12647-3 are recognised in several studies [2–8]. Moreover, there are numerous parameters affecting the print quality. Today there exist commercial systems able to automatically control some specific print quality attributes. Such systems, however, are not capable of providing users with information about which process parameters influence a particular print quality attribute.

Image quality (IQ) is a widely discussed notion, regardless of the media used. One can distinguish sharpness, contrast, and colourfulness as important descriptors of IQ [9]. To assess image quality, we need a valid image quality model [10], i.e. the mapping from individual human perceptions, such as darkness or sharpness, to overall image quality. According to Pedersen et al. an image quality model is intended to evaluate the individual quality attributes and to explore their relationships [11].

Print quality assessment is closely related to assessment of IQ. Pedersen et al. identified and categorised image quality attributes important for assessing the overall print quality of colour images. In [12], it was demonstrated how offset print specific quality attributes, assessed on a small test area using virtual sensors, can be related to process parameters and integrated to obtain an overall print quality score. A virtual sensor assesses product properties using algorithms rather than physical sensors.

It is important to point out that some print quality attributes used in this work are intended for measuring the underlying factors affecting the IQ and lend themselves for varying by the press operator. Such print quality attributes can be denoted as domain specific IQ attributes. Other domain specific IQ attributes exist, for example in the field of image compression [13].

Nonetheless the broad attention to image quality, several communities dealing with various imaging systems, such as camera phones, scanners, colour printers, and copy machines, today lack unified standards for assessing image quality. Therefore, several standardisation initiatives have appeared [14,15]. Another example is the current development of ISO 15311-1 defining standard quality attributes for reproducing a single digital image on different types of digital printing systems [16].

A system, able to predict correct values of print quality attributes given a set of parameter values characterising papermaking and printing processes, would be of great value for manufacturers of paper and newspapers. Acceptable constant print quality of the newspapers results in fewer customer complaints. Such a system would make
an important link between two industries—papermaking and printing—and would contribute to further quality improvements of both paper and prints.

This paper presents an example of such a system developed to assess, explore, and monitor print quality in offset colour printing using data collected from the two industries. The core of the system is a set of virtual sensors operating on images acquired on-line from the printing press. The outputs from these sensors are soft measures— inferential calculations—of print quality. The developed software of the system allows predicting values of these measures, called quality attributes (QAs), using parameters characterising the paper manufacturing and printing processes as input variables to the prediction model. Random forests [17] is a core technique used for modelling. In addition to predicted values, random forests also provide an estimate of the variable importance—a measure of impact the variables have on model accuracy.

1.1. Virtual sensors in printing industry

There are numerous examples of virtual sensors applied in both papermaking and printing industries [18]. Trepanier et al. have demonstrated that image processing can be used to estimate paper surface characteristics such as unevenness of a paper sheet [19]. The sensor has the advantage of capturing spatial information in two dimensions which would be lost when using conventional methods. Surface unevenness affects ink transfer distribution and has impact on print quality [20]. Ink density has also been measured with sufficient accuracy using image processing and soft computing techniques [21]. A system able to detect regions suitable for measuring colour registry, i.e. misalignment of printing plates has been proposed [22]. Template matching systems able to detect regions and colours exposed for distortion in the printed page were developed [7,23]. A quality control system developed by Shankar et al. for flexo-gravure printing also is based on the template matching approach [24].

A good example of a virtual sensor for estimating a print quality attribute is the study of Sadovnikov et al. where a method to estimate motting in colour prints was proposed [25]. Virtual sensors aimed at assessing quality of halftone rasters have been studied in [26]. The authors showed how irregularity of a halftone dot pattern could be measured using three different types of soft measures based on coefficients of the 2D Fourier transform.

1.2. Soft computing in printing industry

Three generations of systems targeted for improving print quality or making the printing process more efficient can be distinguished. The categorisation is based on the ability of the system to monitor, explore and explain various quality attributes.

The first generation systems are able only to monitor some print quality attributes from images acquired online. These systems lack the ability of utilising the non-formalised knowledge that operators possess [22,24,27].

The second generation systems have the ability to exploit the operator knowledge to some extent. Perner presents an example of such a system [23], where image processing techniques and operator prior knowledge are used to handle a detected print quality defect. Evans and Fisher used decision trees to extract print operator knowledge in the form of explicit rules [28]. Surprisingly enough even knowledge not obvious for the operators was detected by the expert system. Perhaps the most obvious example of learning from press operators is the work by Almutawa and Moon, where the system learned to mimic operator actions when adjusting ink flow [2]. A neural network based controller was developed by Englund for controlling ink flow based on online analysis of double grey-bar images, see Fig. 2 [29].

The third generation systems are capable of utilising information not only from the press and operators but also information concerning properties of paper. It is believed that exploitation of such information may lead to higher quality of prints. To our knowledge, there have been only two attempts to use some paper parameters in print quality modelling [30,31]. Utilization of soft computing techniques such as artificial neural networks, fuzzy systems, case-based reasoning, and decision trees, is a distinctive feature of the second and third generation systems. It is worth mentioning that most of the earlier studies have focused on flexographic and ink jet printing.

2. Approach

The proposed system is capable of assessing print quality attributes from images acquired on-line in a printing press and connecting these quality attributes with parameters of both paper manufacturing and printing processes. In contrast to previous studies relying on detection of suitable spots, where an image is acquired and evaluated, the proposed system uses designated halftone areas known as double grey-bars, shown in Fig. 2.

Such grey-bars are used in ordinary production for the grey balance control—for balancing the CMY inks to reproduce a neutral grey. This is done by manual inspection by operators eyes or using densitometers often following a calibration methodology such as G7 [32]. Grey-bars are common in world lithographic newsprint and are typically printed at the edge of each page, and are, therefore, desirable to be used as measurement spots for automatic print.
quality assessment. A usual size of such grey-bar is 8 × 4 mm. Use of well defined measuring areas, such as the grey-bars, has an advantage when comparing different print jobs, detecting trends, and modelling print quality. Images of grey-bars are captured online by a colour camera and values of numerous quality attributes are measured on the images by virtual sensors. Print quality attributes reflect various common deficiencies appearing in offset colour prints and, therefore, are easy to understand for operators of a printing press.

To model print quality, random forests [17] are used in this work. Random forest is a general data mining tool capable of accomplishing various data analysis tasks. Ability to provide the importance of input variables for model accuracy is a very useful function of the random forest software.

The rest of the paper is organised as follows. Section 3 explains the system architecture. Section 4 is about the parameters of paper and printing press used by the models. Section 5 concerns virtual sensors and print quality attributes used to assess print quality. Section 6 concerns print quality monitoring and exploration. Three experimental studies are described in Section 7. Discussion and conclusions are given in Section 8.

3. System design

A database containing parameters of printing press and paper, and quality attributes is a key-part of the developed system. A schematic view of the system is depicted in Fig. 3. An information flow leads from sub-systems at a paper mill and a press room towards the database. The print quality attributes and parameters are accessed by the developed software for data mining.

The virtual sensors for computing print quality attributes operate on images from an 8-bit CCD colour camera. The camera traverses over the web and acquires images of grey-bars of 1008 × 1007 pixels. One pixel corresponds approximately to 0.01 mm. Images of grey-bars are used to compute print quality attributes, discussed in detail in Section 5. In conjunction with the camera system a rotary encoder is mounted onto the black ink roller to synchronise grey-bar position with the camera and flash. Images and their identification numbers are stored in the centralised database. In total 20 grey-bars are scanned from edge to edge of the press cylinder, see Fig. 4, at web speeds up to 15 m/s. It takes approximately 2 min to make the scans. A xenon flashlamp with a continuous line spectrum was used, to have low illumination influence on colour measurements.

A reel identification virtual sensor was developed for the press room. This makes it possible to store each observation as grey-bar measurements with corresponding press and paper parameters. A reel is identified by its bar-code captured by the developed image processing software. A typical input image is seen in Fig. 5. Just before a reel is manually inserted into the press, a camera captures the front view of the reel from a 4 m distance. The system is able to analyse two images per second which is required due to the fast course of events when inserting a reel. Reading bar-codes in a press room is a non-trivial task. Illumination conditions for the bar-code sensor vary due to changing light from the nearby windows and occlusion of light sources. Also, bar-code occlusion occurs (e.g. by hands) when inserting into the printing press. Moreover, the convex shape of a reel can be a problem when reading bar-codes on reels. There are three identical bar-codes printed on a reel and information from all three bar-codes is processed for identifying the reel. The general concept of this novel bar-code reader is described by Algorithm 1. Data linked to a reel from the paper mill is received upon
request using a developed client–server protocol. Paper reel parameters are assembled into a file in the XML format.

**Algorithm 1.** Pseudo-code for the bar-code reader.

1. Transform image columns to compensate for convex reel shape.
2. Compute variance for each column vector to form a vector \( v_x \).
3. Detect the bar-code position in \( X \) direction using \( v_x \) and prior knowledge of bar-code height.
4. Detect the bar-code position in \( Y \) direction using prior knowledge of bar-code length.

for each Detected bar-code \( j \) in image \( BI \) do
   4.1 Compensate for illumination unevenness.
   4.2 Binarise the image using threshold \( b_j \) determined from \( BI \) by searching within \( b_j - k < x < b_j + k \) such that a bar-code initialisation sequence is found, where \( k \) determines the search range.

foreach Detected initialisation sequence \( h \) do
   Read rest of the bar-code, if possible, and save it in the list \( bl \).
end

5. Combine information from the list \( bl \) in such a way that a complete and likely bar-code is found.

Software was developed to interface and analyse data stored in the central database. The software comes with two modes of operation. Apart from the predictive modelling feature in the Exploring mode, the software lets users work in the Monitoring mode, where users can analyse quality attributes in time domain and over the full web dimensions. Interesting signals found automatically using the random forests algorithm can be further visually studied in Exploring mode.

- Monitoring mode enables visualisation of quality attributes over time and at different positions of a press cylinder (Fig. 6, right).
- Exploring mode (Fig. 6, left) allows building models of quality attributes found to be interesting for users (for example, by studying various relations in the Monitoring mode).
- Models built in the Exploring mode provide knowledge about which parameters of the printing press and paper could affect the print quality attributes.
- Exploring mode lets users interact with a tool for visualising signal relations for a specific span of time, position of a press cylinder, and press velocity (Fig. 7).

The \( R \) implementation of random forests ported by Liaw and Wiener [33] is used for modelling. \( R \) functions [34] are called from the developed software. The software was developed using \( C++ \) with dependence on the GUI library Qt [35]. Plot rendering is done by calling the Gnuplot and \( R \).

### 4. Characterising paper and printing press

#### 4.1. Paper parameters

Paper manufacturing is today a highly evolved and complex process. The general principle is: wood fibres are processed and mixed with water and other compounds to form pulp. A typical concentration is 99% water and 1% fibres. The pulp is then sprayed through the head-box onto a woven fabric that has been used to cloth the first part of the paper machine. The idea is to drain the water from the pulp as effectively as possible in the forming and press sections (see Fig. 8), leaving only the drained fibres in a homogeneous layer on a surface of the clothing. The paper is then created by drying the pulp in a heated area of the paper machine. Last step of the process involves rolling the paper into a jumbo-reel, the right-most section in Fig. 8. A typical jumbo-reel is eight to eleven meters wide. This reel is split into several smaller reels which are used in the printing press.

Paper parameters used in this project can be divided into four groups:

- **Lab tests** are 13 different measurements taken from a paper sample strip from one of the reels comprising a jumbo-reel. Typical measurements are optical, mechanical and surface properties such as fibre orientation, surface roughness and density. These measurements are valid for a set of reels that has been cut from the same jumbo-reel.
- **Online scanning** with the Quality Control System (QCS) is located just before the paper is rolled into a jumbo-reel. A measurement head is traversed over the paper width, where each edge-to-edge scan takes 30 s. The head is equipped with sensors to measure moisture content, reflected light, paper thickness and dry weight, see Fig. 9. The measurements are used for controlling different actuators in the paper machine in a closed-loop approach. Output data are stored as three matrices with approximate dimensions of \( 360 \times 70 \) values in cross and machine directions (CDs and MDs), respectively. The data is cut along CD and MD to fit single reels with dimensions of \( 60 \times 28 \) values per reel.
- **Pulp recipe** is acquired from the mill recipe system and contains three parameters about proportions of different pulps such as deinked pulp and mechanical pulp.
- **Paper machine parameters** are registered such as press section speed and amount of additives. In total values of 20 machine parameters are stored for each reel.

**Fourier transform of QCS profiles**. For a profile vector \( x \) with Fourier coefficients \( k \) the sum of squared Fourier coefficients was computed:

\[
S\hat{F}_{mc} = \sum_{p=q}^{i-q} k^2 
\]

where \( k_i \) is a single Fourier coefficient related to frequency \( i \) Hz. For the moisture content signal interesting frequencies range from 5 to 8 Hz. Other frequency intervals are used for the thickness and dry weight profiles.
4.2. Press parameters

Several sub-systems are used to acquire relevant printing press parameters, see Fig. 10. Amount of ink (ink density), press speed, ink temperature, air temperature and humidity, amount of dampening solution, and all adjustments performed by the press operator constitute a set of recorded press parameters. An ink temperature system [29] is used to monitor the temperature for the four ink chambers. The amount of dampening solution is stored for eight nozzles of each ink, see Fig. 4. The colour register is controlled by moving the ink rollers in machine-direction (MD) and cross direction (CD). The sensitivity is 1/10 mm.

The Grey-bar camera system is located just after the black ink roller. For each scanned grey-bar, 23 print quality attributes are computed using virtual sensors, see Section 5.

4.3. Linking paper and press parameters

An image of each grey-bar is mapped to 130 parameters characterising paper and the printing press. However, when using models in the Exploring mode (Fig. 6, left), each observation vector corresponds to average and variance descriptors of a set of parameters characterising one reel. Nonetheless the software determines an optimal subset of parameters to build a specific model, there is also a possibility for a user to chose any subset of the 130 parameters available for modelling.

5. Assessing print quality

Various attributes are used to characterise print quality [18]. The following print quality attributes are used in this work:

1. Colour deviation—usually measured as $\Delta E$ in the $L^*a^*b^*$ colour space:

$$\Delta E = \sqrt{\left(\Delta L^*\right)^2 + \left(\Delta a^*\right)^2 + \left(\Delta b^*\right)^2}$$

where the difference is computed between “coloured” and “black” parts of the double grey-bar. We used $\Delta E$, $\Delta L^*$, $\Delta C = [(\Delta a^*)^2 + (\Delta b^*)^2]^{1/2}$, $a^*$, and $b^*$ to characterise colour deviation. In this work, we computed $L^*a^*b^*$ values using D50 reference white.

2. Mis-registration—alignment of ink planes in both $X$ (CD) and $Y$ (MD) directions for $C$, $M$, and $Y$ inks with respect to $K$. The virtual sensor developed for assessing mis-registration relies on a current image as well as the prior knowledge, i.e. the unique raster angle in halftone patterns for each ink and the dimensions of the grey-bar, see Fig. 11. A pseudo-code of the algorithm is shown in Algorithm 2.

3. Dot deformation. Two measures were applied: standard deviation of the distance to the dot centre $s_d$ and dot shape factor $f_s$:

$$f_s = \frac{p^2}{4\pi A}$$
where $p$ is a dot perimeter and $A$ stands for dot area.

4. **Ink density estimate.** The average and variance of black ink density $d(x,y)$ estimated at each pixel $(x,y)$ in a sample area, shown by a square in Fig. 11:

$$d(x,y) = -\log \frac{p(x,y) - n_{\text{cam}}}{\overline{\text{pref}} - n_{\text{cam}}}$$

where $n_{\text{cam}}$ is average camera noise, $p(x,y)$ is intensity of pixel $(x,y)$, and $\overline{\text{pref}}$ is paper reference average intensity.

5. **Tone value**—the apparent dot area. A tone value was measured in the “black” area of a double grey-bar.

6. **Noise level assessed for C, M, Y, and K areas.** The noise level was assessed by applying the following measure:

$$\sigma^2_n = \frac{1}{36(W-2)(H-2)} \sum_{x,y} (I(x,y) + N)^2$$

where $W$ and $H$ is the image width and height, respectively, $N$ is a filter making the estimate less sensitive to structure in the image, and * stands for convolution. The measure assesses the standard deviation of the noise and was originally suggested by Immerkaer [36]. The technique is suitable for our task due to the low sensitivity to structure.
Assessing mis-registration.


2. Compute binary images $B_C$, $B_M$, and $B_Y$.
3. Compute $B_Y = C < t_C$ & $M < t_M$ & $C < t_Y$, where $t_C$, $t_M$, and $t_Y$ are the binarization thresholds.

```plaintext
foreach C, M, and Y ink i do

4.1 Compute a 2D FFT of image $B_i \rightarrow F_i$.
4.2 Remove frequencies from $F_i$ not expected in a halftone raster for ink $i \rightarrow F_l_i$.
4.3 Inverse transform $F_l_i$ and multiply pixel-wise by $B_k$ and $B_l \rightarrow B_h_i$.
4.4 Build average vectors of rows and columns ($a_r$ and $a_c$) from $B_h_i$.
4.5 Apply low-pass and derivative filters on $a_r$ and $a_c$ and obtain edges by thresholding $\rightarrow a_f$ and $a_f$.
4.6 Use prior knowledge of grey-bar height and width in conjunction with candidate edges in $a_f_c$ and $a_f$, to build membership functions of “plausible edges” fuzzy sets.
4.7 Multiply values of membership functions to extract the most “plausible” edge candidates from $a_f_c$ and $a_f$.
4.8 Using information on edge positions compute X and Y coordinates of the upper-left corner (see Fig. 11) of ink $i \rightarrow Cord_i$.
end
5. Compute X and Y coordinates of the upper-right corner of black ink (using Step 4 without the “fuzzy part”) $\rightarrow Cord_{black}$.
6. Mis-registration is given by: $Cord_{C/M/Y} - Cord_{black}$.
```

In total 23 attributes, listed in Table 1, were used. These attributes are targeted to capture known print defects and aid the quality interpretation from the printing industry perspective. A unified measure of print quality has been studied earlier using similar set of quality attributes in [12].

6. Exploring and monitoring print quality

Nonlinear modelling was used to predict quality attribute values using paper and press parameters as independent variables. For this purpose a data mining technique called random forests (RFs) [17] has been used. Due to its low computational complexity, RF can handle thousands of variables of different types with many missing values. For an RF tree grown on a bootstrap sample, the out-of-bag (OOB) data can be used as a test set for that tree. As the number of trees increases, RF provides an OOB data-based unbiased estimate of the test set error. An estimate of variable importance and the data proximity matrix can also be obtained from RF.

We use the regression accuracy-based estimator of variable importance [38] in this work. The importance measure $D_j$ for variable $x_j$ is given by

$$D_j = \frac{1}{B} \sum_{b=1}^{B} \left( R_{ob}^{bj} - R_{ob}^{iob} \right)$$

where $B$ is the number of bootstrap samples (trees in the forest), $R_{ob}^{ij}$ is the regression error for the OOB data by the tree $T_b$, and $R_{ob}^{iob}$ is the regression error for the OOB data by the tree $T_b$ when values of $x_j$ in the OOB set were randomly permuted.

We used RF variable importance estimates to reduce the dimensionality by eliminating the least important variables. One can use principal component analysis (PCA) to reduce dimensionality. However, PCA requires measuring all input variables and does not provide importance of them. In [12], the superiority of RF over linear as well as support vector machine regression in the print quality modelling task was demonstrated.

In the Monitoring mode, users can analyse quality attributes as matrices or images, where each row corresponds...
to a scan of grey-bars in CD and columns corresponds to the time dimension. We use colour to depict values of quality attributes in such type of data representation, see Fig. 6 (right). Each scan is tagged with date and time. Measurement errors, e.g. black ink is missing, are presented in the matrix as “invalid measurement”. Deviations in quality attribute values are easily detected by the operator as colour of the matrix elements change.

7. Experimental studies

7.1. Exploring ink density variations

To explore the influence of the paper and press parameters on black ink density (attribute number four in Table 1) variance, the Exploring mode was used. The data used for this study are values of 77 parameters collected from about 30,000 grey-bars and 107 reels. A random forest of 2000 trees was created using eight randomly selected features to split a node.

Two data sets were used. First, all grey-bars were used—even those sampled at low press speeds. Then, to create the second data set, the start-up sequence was filtered out by leaving aside grey-bars corresponding to press velocities below 5000 prints/h. The variable importance values computed using the full and filtered data sets are shown in Figs. 12 and 13, respectively. In red¹ shown are press parameters and in green–paper parameters.

The figures indicate that press parameters are more important than the paper parameters when press start-

¹ For interpretation of color in Figs. 12 and 13, the reader is referred to the web version of this article.
ups and shut-down sequences are included. However, paper parameters are more important than the press parameters when the printing process stabilizes—when press start-ups and shut-downs are excluded. The five most important parameters for each model identified in the experiment are shown in Table 2, where PR stands for printing press and P means paper. The first column of Table 2 lists parameters that are significant for the press startup. When the low press speeds where filtered out, paper parameters appeared, see the second column of Table 2.

<table>
<thead>
<tr>
<th>Table 1</th>
<th>Print quality attributes.</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>Dot deformation $s_d$</td>
</tr>
<tr>
<td>2</td>
<td>Dot deformation $f_t$</td>
</tr>
<tr>
<td>3</td>
<td>Tone value</td>
</tr>
<tr>
<td>4</td>
<td>Estimate of average black ink density</td>
</tr>
<tr>
<td>5</td>
<td>Mis-registration of C in X direction</td>
</tr>
<tr>
<td>6</td>
<td>Mis-registration of C in Y direction</td>
</tr>
<tr>
<td>7</td>
<td>Mis-registration of M in X direction</td>
</tr>
<tr>
<td>8</td>
<td>Mis-registration of M in Y direction</td>
</tr>
<tr>
<td>9</td>
<td>Mis-registration of Y in X direction</td>
</tr>
<tr>
<td>10</td>
<td>Mis-registration of Y in Y direction</td>
</tr>
<tr>
<td>11</td>
<td>$\Delta E$</td>
</tr>
<tr>
<td>12</td>
<td>$\Delta L^*$</td>
</tr>
<tr>
<td>13</td>
<td>$a^*$ component of the black grey-bar part</td>
</tr>
<tr>
<td>14</td>
<td>$b^*$ component of the black grey-bar part</td>
</tr>
<tr>
<td>15</td>
<td>$a^*$ component of the coloured grey-bar part</td>
</tr>
<tr>
<td>16</td>
<td>$b^*$ component of the coloured grey-bar part</td>
</tr>
<tr>
<td>17</td>
<td>Contrast measured in the black grey-bar part</td>
</tr>
<tr>
<td>18</td>
<td>Contrast measured in the coloured grey-bar part</td>
</tr>
<tr>
<td>19</td>
<td>Noise attributed to black ink in a grey-bar</td>
</tr>
<tr>
<td>20</td>
<td>Noise attributed to cyan ink in a grey-bar</td>
</tr>
<tr>
<td>21</td>
<td>Noise attributed to magenta ink in a grey-bar</td>
</tr>
<tr>
<td>22</td>
<td>Noise attributed to yellow ink in a grey-bar</td>
</tr>
<tr>
<td>23</td>
<td>$\sqrt{(\Delta C_k - \Delta C_c)^2}$, where $\Delta C_k$ and $\Delta C_c$ is $\Delta C$ (see Eq. (2)) for the black and coloured grey-bar part, respectively</td>
</tr>
</tbody>
</table>

Surface roughness is a parameter, known for affecting ink transfer from blanket to paper [1,39].

### 7.2. Monitoring and predicting mis-registration

To monitor and revisit a press run, users can draw benefit from using the “Monitoring mode”. Data from two reels, A and B, acquired online in ordinary production were explored by analysing each grey-bar and applying the virtual sensor to assess the mis-registration degree. To obtain the data, the grey-bar camera, see Fig. 10, was traversed over the paper web and images of grey-bars were recorded at each CD position. Fig. 14 illustrates the mis-registration degree assessed for C, M, and Y inks at different positions across the printing cylinder. The Y axis in Fig. 14 represents time. We use colour to depict values of C, M and Y mis-registrations in CD. Warm colours represent mis-registration deviating to the right-hand side and cold colours to the opposite direction.

One can clearly see similar mis-registration patterns for C and M inks, but not Y. This is due to an effect known as fan-out (related to paper swelling). The effect can now be studied intuitively and graphically.

The fan-out effect is less for the yellow ink because the paper has a smaller distance to travel from the yellow ink chamber to the black ink chamber and therefore less swelling of paper occurs. This is further explored in Fig. 15 confirming the effect. Blue dots in the plots show the average mis-registration for each ink computed using information from 100 reels. One measurement point in the plots show the average CD mis-registration in one reel for a specific ink. There are 20 measurement positions across the web. The figure presents measurements made at positions 1, 10, and 20. As can be seen, the average mis-registrations for C and M inks are larger in the outer-

![Fig. 12](image-url)  
**Fig. 12.** Variable importance when modeling black ink density variation. Full range press velocity was used.
most positions of the web (positions 1 and 20) than in the middle, see Fig. 15 and Table 3. The positive correlation seen in the top three plots of Fig. 15 reinforces the observation, since mis-registration for C and M inks have a tendency to occur in the same direction.

To explore and identify influential parameters (variables) for the fan-out effect, a model of C mis-registration in CD was built using RF. Printing cylinder positions 1–5 were used to collect data for the modelling. Concerning plots shown in Fig. 14, printing cylinder positions 1–5 are on the left-hand side of each plot. Parameters of printing press and paper linked to these positions were collected from 85 reels. Press start-ups were excluded from the data-set. The variable importance obtained from the RF when using all 55 variables for the modelling is shown in Fig. 16. Iterative variable elimination was then performed aiming to obtain a model with the lowest prediction error–mean squared error (MSE). In each iteration a new RF was built using \( \sqrt{D} \) variables to split a node, where \( D \) is the number of remaining variables. Fig. 17 illustrates the variable elimination process. The arrows in Fig. 16 marks variables that survived the variable elimination. As can be seen from Fig. 16, all the five remaining variables are amongst the eight determined as the most important by the initial RF without any variable elimination. This fact substantiates the ability of the variable importance measure to correctly determine relevant variables even in the case of large redundancy.

Table 4 shows the five remaining features after the elimination. According to the RF variable importance measure, Tensile Stiffness Orientation (TSO) is the most important variable. This is reasonable, since the feature essentially represents the fibre orientation in the paper, which is known by paper experts and press operators to affect dimension stability of the paper. It is also not surprising that tension and amount of fountain solution the paper web is exposed to in the press is exposed to in the same direction.

Predicting mis-registration is a difficult task, since it is difficult to build a comprehensive model of the phenomenon. Even moderate agreement between measured and predicted values can be considered as a good result. Fig. 18 plots measured versus predicted values of C mis-registration in CD. The correlation coefficient of 0.48 was computed for the measured and predicted values. Bearing in mind that only five variables were used to model the phenomenon, the obtained correlation is rather encouraging. The experimental results show that the system allows modelling different print quality attributes with reasonable accuracy and exploring paper and press parameters affecting the attributes. This is of great value for both papermaking and printing industries.

7.3. Comparing the suggested technique to assess print quality attributes to conventional methods

CMY mis-registration (rows five to ten in Table 1) and tone value (row three in Table 1) are the quality attributes

Table 2
Top five parameters.

<table>
<thead>
<tr>
<th>Rank</th>
<th>Model with start-ups</th>
<th>Model without start-ups</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>(PR) Web tension variance (sensor 3)</td>
<td>(P) Surface roughness</td>
</tr>
<tr>
<td>2</td>
<td>(PR) Press speed variance</td>
<td>(PR) Fountain solution nozzle setting</td>
</tr>
<tr>
<td>3</td>
<td>(PR) Fountain solution setting variance (sensor 1)</td>
<td>(PR) Mean humidity in press room</td>
</tr>
<tr>
<td>4</td>
<td>(PR) Web tension variance (sensor 2)</td>
<td>(PR) Web tension mean (sensor 3)</td>
</tr>
<tr>
<td>5</td>
<td>(PR) Web tension variance (sensor 2)</td>
<td>(P) Paper permeability</td>
</tr>
</tbody>
</table>

Fig. 13. Variable importance when modeling black ink density variation. Low press velocities excluded.
we used for the comparison in this study. In a modern printing press, some on-line camera-based system is usually used to assess CMY mis-registration, while tone value is usually checked using a densitometer or spectrophotometer. When an automated on-line camera-based system is not available, press operators often use visual inspection with or without a magnifying glass, to assess mis-registration.

The CMY mis-registrations for the data set consisting of 80 samples were manually annotated by inspection of each grey-bar image. This ground truth was used to assess the measurement accuracy of the proposed technique as well as the online camera-based system used for the comparison. To assess CMY mis-registrations online, commercial camera-based systems usually require some “register marks–register imprint” to be printed. Fig. 19 presents an example of register imprint used for measuring mis-registration by the commercial OPM system. This type of “register mark” was also used in this study for the comparison. Therefore, each of the 80 samples also contained register imprints in addition to grey-bars. The imprints-based assessment of the CMY mis-registrations was obtained using an off-line version of the OPM system. The same RGB camera of 1280 × 960 pixels was used to capture both grey-bar and register imprint images.

Table 5 presents results of the test, where given are the 95% confidence intervals for the C, M, and Y mis-registration (in millimeters) measuring accuracy in the x and y directions for the two systems. As can be seen from Table 5, both systems provided approximately the same measuring accuracy. Bearing in mind that the proposed system does not require printing specific register imprints and that grey-bars are usually printed anyway for visual inspection, densitometer measurements, and other types of printing quality inspection, the advantage of the proposed system is obvious.

To assess tone value (dot coverage area), we binarize the “black” grey-bar image part and compute the percentage of the area covered by ink. Fig. 20 presents an example of an image of a black ink halftone area from a grey-bar along with the binary counterpart of the image. In total, 80 aforementioned grey-bar samples were processed and

![Fig. 14. Screen-shots from software: matrices of mis-registration degree for C, M, and Y inks measured by the virtual sensor.](image-url)
the mean tone value and the standard deviation of tone value were computed.

In a printing shop, a densitometer is usually used to assess dot coverage area indirectly. A densitometer uses spectral information to estimate the optical dot area coverage according to the Murray–Davies formula [40]. We measured the dot coverage area of black ink in the same 80 grey-bars using the Techkon densitometer.

The mean value and the standard deviation computed from the 80 samples were: 53 ± 3.2% and 42 ± 2.1% for the densitometer and the camera-based measurements, respectively. The larger mean value of the densitometer-based measurements could be explained by the optical dot gain effect (the Yule–Nielsen effect [41]) that makes the dots appear larger than they actually are, which is taken into account in the Murray–Davies formula. Visual inspection of the images presented in Fig. 20 confirms the validity of the results obtained from the proposed technique.

<table>
<thead>
<tr>
<th>Ink/position</th>
<th>1</th>
<th>10</th>
<th>20</th>
</tr>
</thead>
<tbody>
<tr>
<td>Cx</td>
<td>-28</td>
<td>-7</td>
<td>19</td>
</tr>
<tr>
<td>Mx</td>
<td>-14</td>
<td>-9</td>
<td>5</td>
</tr>
<tr>
<td>Yx</td>
<td>4</td>
<td>-6</td>
<td>0</td>
</tr>
</tbody>
</table>

### Table 3
Mean mis-registration (average of 100 reels) in CD (in pixels) for each ink. One pixel corresponds to approximately 0.01 mm.

The mean value and the standard deviation computed from the 80 samples were: 53 ± 3.2% and 42 ± 2.1% for the densitometer and the camera-based measurements, respectively. The larger mean value of the densitometer-based measurements could be explained by the optical dot gain effect (the Yule–Nielsen effect [41]) that makes the dots appear larger than they actually are, which is taken into account in the Murray–Davies formula. Visual inspection of the images presented in Fig. 20 confirms the validity of the results obtained from the proposed technique.

**Fig. 15.** Mean (per reel) mis-registration (in pixels) in CD. Cx, Mx and Yx stand for mis-registration in CD of C, M, and Y inks, respectively. Blue dots represent means for 100 reels. (For interpretation of the references to colour in this figure legend, the reader is referred to the web version of this article.)

**Fig. 16.** Variable importance for predicting the C mis-registration in CD.

**Fig. 17.** MSE for models with different number of features.
8. Discussion and conclusions

This work has resulted in a novel system for assessing, monitoring and exploring print quality. An important part of the system is the set of virtual sensors used to assess print quality attributes. Quality attributes are linked to paper and press parameters and allow users, in intuitive fashion, to assess, model, and monitor print quality attributes using the Monitoring and Exploring modes of the software.

It was demonstrated experimentally that models are able to capture various effects of paper and print interaction such as influence of surface roughness on ink density, fan-out, and others. The experimental results show that the system allows modelling different print quality attributes with reasonable accuracy and exploring paper and press parameters affecting the attributes. This is of great value for both papermaking and printing industries.

Future work involves acquiring a large data set that enables studies of various "low frequency" effects, seasonal variations for example, on print quality. Hundreds or even thousands of reels are required to obtain a statistically significant result. Also, the Monitoring mode can be extended to an online version, where operators in realtime can follow the quality matrices emerge as the print run continues. The work can be extended to obtain a proactive decision support system, able to infer a degree of certainty about various relationships. Future work also includes investigations towards a system capable of suggesting reel-specific press settings to enhance print quality.

Acknowledgments

This work was supported by the Knowledge Foundation of Sweden, Grant Number 2007/0279. Part of this work was published in [42].

References
