CGATS Recommended Industry Practice

Graphic technology — Data Set Development

CGATS

Association for PRINT Technologies

450 10th CIRCLE N, NASHVILLE, TN 37203 • TEL: (615) 432-5421• www.aptech.org

A CGATS Reccomended Industry Practice

APTech The Association for PRINT Technologies performs secretariat functions for the Committee for Graphic Arts Technologies Standards (CGATS) and its subcommittees, working groups and task forces. This function includes publishing its documents. APtech is not the author of this document, nor is it an entity that adopts any standard.

The information in this document was developed by members of CGATS Subcommittee 3 (Metrology & Process Control) in cooperation with Print Properties and Colorimetric Working Group of Idealliance. Although its contents were approved by ballot of the voting members of the subcommittee, this document has not been developed under the consensus process of the American National Standards Institute (ANSI) and is not an ANSI consensus document. Nor is it a mandatory standard.

At the Subcommittee's request, APTech is making this document publicly available at no cost for informational and reference purposes only.

This document does not relieve the reader of any burden imposed by law, and the reader should independently investigate and verify the performance, results, safety, and efficacy of his or her own product or process, even if it follows any recommendations contained in this document. CGATS, its subcommittees, and APTech expressly disclaim liability for any damages or injury arising out of or related to the use of information contained in this document.

This document may be updated as needed by the developing CGATS subcommittee. Such updates will be identified by version number and date and will be available on the APTech Standards Workroom.

Comments or questions on this document should be sent to the CGATS Secretariat, APTech The Association for PRINT Technology, 450 10th Circle N, Nashville, TN 37203.

Contents

Introduction

This document provides guidelines on how to create a standard printing condition dataset: from printing and recording initial measurement data, to creating and optimizing the final characterization data. Historically, different methodologies were used to create legacy-standard data sets such as SWOP, SNAP and GRACoL. These earlier methods are documented in CGATS 'Industry Practice Documents: Analysis-and-Reporting Press Run Guidelines' and can be referenced for historical purposes. This document provides a new and updated process that reflects the current best practices and recommendations for developing a characterization data set from the Idealliance Print Properties Committee.

The objective of this CGATS-recommended best-practices document is to provide a modern framework for planning and executing press runs for the purpose of deriving color characterization data that may be used as a reference. To that end, the documented practices define the key steps necessary to conduct press runs, collect the measurement data from printed Test Forms, and process the resulting characterization data to create a reference data set. This document is intended to serve as a checklist for standards and industry groups developing reference color characterization data, as well as for individual printing organizations who may be developing their own internal procedures. The sections that follow describe the key steps needed for the successful execution of a characterization press run and are organized into the chronological sections:

— Preparation
— Printing setup
— Specifications
— Target selection
— Sheet evaluation
— Measurement and collection
— Analysis
— Creation of data set

Graphic technology — Data Set Development

1 Scope

This document describes the methods used to produce a characterized print condition (CPC). A CPC describes the color behavior of a printing device. A CPC can be used in 2 ways: as a “color management reference”, allowing that specific printing condition to be simulated on another output device and; as a set of numerical “aim values” to be used as a target for determining compliance. CPC aims may be used to gauge overall compliance and may also be used as process control aims during the print production process. CPC data can further be compared to industry-standard reference data, such as GRACoL and SWOP, to understand how close a print condition is to an existing industry standard. In some cases, printers may be using print processes or substrates that cannot meet acceptable tolerances for any existing industry-standard references. In these cases, users and organizations may wish to nominate their CPC to become a CRPC (Charachterized Reference Print Condition). This document outlines the key steps to create a new CPC dataset, and includes the additional requirements for establishing the new CRPC dataset.

2 Normative References

The following standards contain provisions that, through reference in this text, constitute provisions of this standard. For dated references, only the edition cited applies. For undated references, the latest edition of the referenced document (including any amendments) applies.

3 Terms and Definitions

For the purposes of this document, the following terms and definitions apply.

3.1 CPC

(Characterized Printing Condition).
A printing condition and its color characterization data.

Note: A characterized print condition may also be described as a colorspace.

3.2 CRPC

(Characterized Reference Printing Condition).
Identified printing condition and its color characterization data *used as the aim* for a particular printing task. A characterized reference print condition may also be described as a colorspace.

3.3 dryback

change in color, gloss, or density of an ink film as it dries and penetrates the substrate. Occurs mostly in sheetfed offset litho where inks dry slowly. Is minimized with in-line AQ (water-based) coating.

3.4 finishing

a process such as lamination, coating or other final process after the printing that may change the measured color

3.5 G7®

calibration aims and methods based on gray balance and tonality

3.6 repeatability

The ability of a printing system to produce the same result over and over again.

3.7 reproducibility

The ability of different printing systems to create the same result.

3.8 Substrate Corrected Colorimetric Aims (SCCA)

A method for scaling/adapting a CRPC dataset so that the reference is more achievable on those print processes where the substrate whitepoint is significantly different than the original CRPC whitepoint

3.9 Spot Color Tone Value (SCTV/ISO 20654)

A color-based print curve tone-calibration method that is ideally suited for use with spot inks.

3.10 Synthetic calibration

A method of calibrating a dataset using a calibrating methodology.

3.11 Tonal Value Increase (TVI)

Another term for dot gain, and a way to quantify the growth of a dot during the printing process. TVI is a paper-relative ratio base on the density of a tint compared to the density of that same ink’s solid

4 Preparation

For the development of a CPC, it is important to plan and document the testing process. This documentation is important if testing will be performed in multiple locations and/or on multiple printing machines. Preparation required will vary depending on the type of printing process.

4.1 Target Characterization Parameters

Prior to performing any test runs, the following Target characterization parameters should be defined and documented, Parameters include, but are not limited to:

4.1.1 Printing process type (e.g. Flexo, Offset)

4.1.2 Substrate type (e.g Paper, Film)

4.1.3 Ink type (vehicle or cure type e.g. Water-Based, LED-UV cure)

4.1.4 Print Process Details

CMYK process, spot, or additional spot and process combinations, and sequence/laydown/order. Typically inks and overprints will also define the gamut of the test CPC.

4.1.5 Solids, overprints, and specific color solid ink aims

4.1.6 Resolution, quality modes, and screening methods

4.1.7 Coating, lamination or special finishing

4.1.8 Test elements

Includes graphics, images, colorbars, and color charts to be used on the Test Form.

4.1.9 Test Charts

Test Charts & Test Form layout should be based on test requirements and the measurement plan. The Test Form are recommended to include repeated Test Charts when characterizing a printer with that exihibts variability across the printable area. Orientation requirements based on printing process. Gripper or lead edge, button/operator or gear/drive side. If multiple up/on/out, die position number, location, position.

4.1.10 Test Form

The Test Form is a document, which includes the required Test Charts and any additional test elements or images, for use in charcaterizing the printing machine and evaluting print quality.

4.1.11 Printing Plan

Include speed, number of Test Form prints, sample selection methodology, and any relevant maintenance history for the printing machine.

4.1.12 Total area coverage (TAC) to be used with the specific print process

4.1.13 Ink coverage

On the press form (light coverage, medium coverage, heavy coverage) based on density of the artwork

4.1.14 Parameters to be measured before, during and after test run and to establish validation tolerances.

5 Printing Plan Guidelines

A printing plan should be documented in writing prior to performing any printing. The printing plan should:

- contain as much detail as possible about the test run procedure.
- enable the test run to be recreated by additional parties with similar machines, inks, substrate, and other consumables
- should define appropriate tolerances based on machine and process capabilities.

The plan should contain information on the target print conditions, methods of calibration, and recording of data. If there are objectives or specific matching requirements these should be noted in the plan, as well as any requirements such as compliance for a print quality program. The plan should be written as a recipe that others can follow to create the same results.

6 Printing Setup Guidelines

Prior to testing, the printing machine needs to be checked to confirm it has been well maintained and is in a stable and repeatable condition. The details of stabilizing the system will vary depending on printing method. The starting printing machine setup and configuration need to be documented in order to record the press conditions and other details, as well as for use in analysis and comparison of printed results.

The details of printing setup may include the following:

- Mechanical condition check
- Preventative maintenance
- Comparison of mechanical attributes to manufacturer/factory settings
- Documentation of all relevant mechanical attributes, settings, and consumables such as inks, plates, screening, resolution, etc. When available use existing prequalification kits.
- Instrument certification, evaluation, and compliance

7 Selection of Printing Aims

The printing target may be selected based on type of printing, substrate, and relevant needs. If the target is based on prior characterization data or is an adaptation of existing characterization data, then the target should be noted as an updated version or adaptation. If based on an existing set of characterization data having a significantly different substrate white point, the colorimetric aims should be adapted using SCCA or other acceptable adaptation methods. If there is no existing or relevant characterization data then the characterization data may be new custom data, and not based on existing characterization data.

7.1 Creating a dataset based on an existing characterized print condition.

- **Evaluate** existing characterized print conditions to determine appropriate source, typically closest to process, gamut and substrate
- **Select** substrates, inks, processes/screening (gamut to be achieved)
- **Define** aims for characterization data if any, including use of SCCA or other acceptable methods used to create the aim values

— **Determine** process calibration/control philosophy to be used

— **Create** press tests to provide candidate reference data

— **Record** test run data

— **Correct** and modify data as needed

— **Evaluate** reproducibility of new aims

7.2 Creating a new reference dataset

— **Define** color matching requirements (such as if it needs to match existing CPCs, gray balance, etc.)

— **Define** differences or needs not addressed by current CPCs

— **Select** substrates, inks, processes/screening (gamut to be achieved)

— **Define** process setup – define new aims for characterization data

— **Develop** strategy to accommodate press/printing variability

— **Define** scope of use for the RPC: Process specific vs multi process

— **Define** Process calibration/control method to be used

— **Create** press tests to provide candidate reference data

— **Define** number of press runs and machines to be tested

— **Record** data and harmonize

— **Correct** and modify as needed

— **Evaluate** viability of new aims based on color matching requirements

8 Ink and Substrate Selection Guidelines

Ink and substrate should be selected based on the target print condition and desired use. Substrate should be defined using colorimetric values to determine closest white point. This can be done by averaging measurements of appropriate substrates. Surface characteristics, coating or lamination, name and manufacturer, and OBA level should be recorded as well. Be aware that not all regions or testing sites may be able to obtain and use the exact same substrate or ink.

9 Test Chart and Test Form guidelines

Test Forms should use industry standard targets whenever possible. Test Charts should be created in a format that can be measured by the agreed upon measurement devices. For process characterization it is recommended that the IT8.7/5 Test Chart be used as part of data collection. Smaller Test Charts may

also be used for calibration, or for during run process control, quality control and evaluation. Depending on the printing process, Test Charts may also be designed with the intent of targets repeating in multiple locations on the sheet to capture and record variation.

10 Calibration Methodology

Calibration methodology should be clearly defined, and whenever possible industry standard calibration methods should be followed. These include but are not limited to G7, SCTV, TVI, ICC and other calibration methods. If a non-standard calibration methodology is used, or standard calibration methodology is modified, the specific procedure must be documented.

11 Sampling Strategies

When collecting and measuring prints the following sampling strategies can be used:

— Random sampling

— Uniform sampling

— Sequential sampling

11.1 Random sampling

Random sampling is suggested in cases where the printing conditions are believed to be consistent and stable within the press run and no efforts were made to tune a specific portion of the press run. Random samples are selected by generating scaled random numbers from random number generators (such as those found in common office suite software, or from tables of random numbers found in the appendices of most statistics textbooks). If a random sampling strategy is chosen, CGATS recommends collecting a larger number of sheets than when sheets are pulled from a specific portion of the press run.

11.2 Uniform sampling

Uniform sampling addresses long term drift in printing conditions within a press run. It is typically taken from a press run at uniform intervals. The uniform sampling method is suggested in cases where the long-term variation of the press is deemed unacceptably large. For example, in a press run of 5,000 sheets, samples might be taken at intervals of 100 impressions. Alternatively, samples might be taken in uniform measures of time (such as every 30 seconds).

11.3 Sequential sampling

Sequential sampling addresses sheet-to-sheet variation in printing conditions within a press run. For example, selecting sheets 101, 102, 103...n would provide the ability to determine press performance at certain points in the press run. Sequential sampling is suggested in cases where it is felt to be important to apply specific controls or adjustment of printing conditions, within a press run, to ensure a close match to the specified aims and tolerances.

12 Test Runs

Test runs should be performed using the recommended number of sheets to assure the printing machine is stable and in a production like environment. This means enough substrate and time should be available to avoid having the test run take place in a cold state. The test run must use an appropriate number of sheets and time for the specific print process. This allows for greater confidence that the press run can be used successfully.

13 Sheet Evaluation Guidelines

Sheets should be evaluated to ensure the machine has printed correctly. Depending on printing process this may mean print is in registration, fine type and details are printing correctly, there is no banding, mottling, or other defects in print. In addition to looking at the overall sheet, prior to being taken for measurement test targets need to be evaluated to make sure there are no defects such as scuff, over inking, mottling, or scratches, banding or other defects within the test targets to be used for measurement in calibration and data collection.

14 Press Sheet Collection Guidelines

Sheets should be collected using one of the sampling methods described above and clearly marked with the date, position in the run, and relevant print characteristics based on method such as screening, rotation, resolution, numbers of passes, coating, etc. Sheets must be measured within a process appropriate amount of time to not be changed by dryback or fading when applicable.

The objective of the color characterization data is to record press sheets that match as closely as possible to a set of predefined printing criteria, therefore an additional selection refinement step is often appropriate. This can be used to further refine a larger data set, selected using either the random or sequential sampling techniques.

Sub-sampling requires that specific criteria associated with a particular printing specification (such as solid lab values), gray balance, color, overprints, etc.) have been selected as the factors to be used

15 Finished Print and Post Process Characteristics

Sheets used for characterization data should contain finished effects, such as coating, lamination or other final finishing processes. In most cases characterization data should be based on the final print characteristics. (In some cases it may be desirable to measure both both unfinished and finished data.)

16 Use of Imperfect Sheets

In many cases press sheets do not need to be from a perfect press run as long as the gamut is matched to the target print condition. For example, samples can be synthesized to adapt to the target white point, as well as to perform synthetic G7, TVI and SCTV calibration to the sheet for use in analysis. (Add more on modeling, how to create a dataset without a press run or perfect press sheets.)

17 Measurement Guidelines

Once samples are collected, 6-12 sheets from each press run that align with stable production should be measured. Measurements should be performed using the desired M condition, and if possible, performed using all available measurement conditions (typically M0, M1, M2).

Samples should also be collected from a number of separate presses for which the data set should be considered valid. For data sets to be used across many locations and machines, there should be 6-12 press runs conducted to achieve the desired level of accuracy, using different sites and machinery applicable to the printing process and substrate. For custom machine use, data sets can be based on a smaller sample set.

The measurement devices used to record press runs should be checked to make sure they have recently passed a performance certification. If multiple measuring devices are being used, they should be checked and compared to make sure they are measuring in a similar manner.

Where multiple laboratories are participating in the development of data sets, CGATS recommends that one or more samples be measured by a coordinating lab and that those measured samples, along with additional samples, be supplied to each of the other participating laboratories to provide measurement system crossover data.

NOTE 1: When multiple measurement laboratories are being used, it is very important that they establish measurement system agreement and control prior to making measurements to be used in generating the color characterization data. See the CGATS Recommended Industry Practice: *Color characterization data set development — Procedures for color measurement system process control and inter-lab coordination*[9] for further information.

Press sheets and proofs may degrade relatively quickly. It is important to complete all measurements before discoloration has occurred. However, press sheets must be allowed to thoroughly dry before being measured to ensure that any effects of ink dry-back are excluded.

It is highly recommended that spectral data be collected. If only colorimetric data is available, CIEXYZ data is preferred. If CIEXYZ or CIELAB data is collected directly from measurement instrumentation the user should consult with the instrument manufacturer to ensure that the data is computed according to the applicable standards. The data files corresponding to each target measured should be separately identifiable.

Calculation of the CIE tristimulus values, XYZ, and any subsequent CIELAB values should use the weighting functions and equations of ANSI CGATS.5, which are based on the 1931 CIE 2˚ observer and the D50 illuminant.

18 Analysis

18.1 Objectives of analysis

The objective is to have smooth and consistent data. In the past this was done by collecting and avergaing large numbers of samples. The result should be a subset of data to be used for data analysis and data set creation. In more contemporary applications software, spreadsheets and tools are used to model and optimize the data.

18.2 Modeling

Modeling is is using a mathmatical approach to represent the behavior of measured data. Tools for modeling data can be used to optimize a measured data set to effectively average multiple samples, remove outliers and/or defects, and to smooth or idealise the characterization data to meet a desired calibration requirement. Care has to be used in modeling to make sure it acurately characterizes the behavior.

- The data set should be compared with with original print samples to compare the relationship between the dataset and original prints.

18.3 Data evaluation

- Patch specific evaluation
- Composite data set evaluation
- Data set preparation

After the selected press sheets have been measured, the resulting pool of data must be evaluated to look for outlier data points and/or extreme (suspect) data. This ensures that the final data set is as statistically sound as possible, and that sufficient information is included with the data set to allow user evaluation.

18.4 Patch specific evaluation

Before multiple measurements of individual patch samples (either repeat readings of the same sample or readings from different samples) can be combined, they must be evaluated to determine consistency of the data. The method currently used by CGATS is to compute CIELAB data for each sample, and the CIELAB ΔE between each sample and the average CIELAB for that patch. The statistics of the CIELAB ΔE population can then be used to identify potentially bad data. Samples with significantly larger CIELAB ΔE values can indicate either measurement errors or printing anomalies that were missed in earlier evaluations.

Irregular patch measurements should be identified and corrected or removed from the data before further evaluation.

18.5 Composite data set evaluation Guidelines

Once outlier points have been removed, new average values and CIELAB ΔE values can be computed. The total population of CIELAB ΔE values can then be pooled and histograms and cumulative probability plots prepared using any of the typical spreadsheet tools.

A simplified way to create a cumulative probability plot is as follows:

- Sort all the CIELAB ΔE values from smallest to largest and number the resulting list from 1 to N (where N is the number of samples).
- Plot the resulting data (rank number vs. CIELAB ΔE values) using CIELAB ΔE values as the x axis (horizontal) and the sample number divided by N as the y axis (vertical).
- Compare the shape and values of the resultant histogram and cumulative probability plots with those of other data sets can help evaluate the consistency of the composite data set.

While the two types of plots show similar characteristics of the data, each has advantages and disadvantages. The histogram shows the distribution of the ΔE values and can be useful in understanding what realistic variations are.

The shape of the cumulative probability plot helps provide an insight into how tightly grouped the data is. For example the 50th percentile point for the data set of the tightly distributed data set correlates well with the peak of the histogram. On the other hand the wider distribution has a peak that occurs before the 50th percentile point. In general, the steeper the cumulative probability curve, the tighter the data distribution. However, any evaluation of the distribution of the ΔE values of a data set must also take into account the size of the data set.

The use of data from multiple press runs further improves accuracy and precision; results converge closest to the process specification when data from more than one press run are combined. The inclusion of more than one press run allows further increase in the number of variables sampled. The effects of different lots of ink and substrate, press crew and environmental conditions as well as random fluctuations in press conditions can be accounted for in the resultant characterization data. With more replication and more variables sampled, accuracy and precision are improved. Results from a single press run are biased towards the conditions that existed during that press run only and gauging the reproducibility of those results is impossible. The criticality of the color output should govern the decision to use more than one press run.

Based on the quality objective of the press run for characterization, accuracy and precision, an exhaustive sampling plan ensures the most accurate and precise data but may be economically prohibitive due to the substantial amount of measurement and analysis resources required. The compromise between accuracy and precision and resources is found by focusing the sampling plan on the particular needs of the press run. In the case of press characterization, the need is accuracy relative to intended aim values, which is found most efficiently through the use of directly measured total statistics so long as the press run is compliant with process specifications on average. The direct measurement of total statistics requires that measurements be taken from several within-sheet locations and from several sheets. The measurement of all test elements on all sampled sheets is not necessary. Total statistics can be adequately determined by measuring a few different elements from a few different sheets.

19 Dataset Creation

The following are steps recommended for use in creation of a new reference dataset once press sheets have collected:

- **Validate** instruments.
- **Define** laboratory standard operating procedures and define SOP for measurement.
- **Measure**. The quantity and frequency of measurements are to be based on project requirements.
- **Identify and eliminate** of outliers and undesirable measurements and samples.*
- **Average** or smooth measurement data as required*
- **Adapt** target charcaterization data white point as required*
- **Correct** and/or optimize data to G7/TVI/SCTV as needed to create idealized characterization data*
- **Re-evaluate** any data corrections
- **Save** in Data format ISO 17972-3 CxF/X-3 for XML or ISO 28178/CGATS.17 for ASCII text
- **Prepare** data for sharing and review
- * If working with data from different printing machines data correction should occur prior to the averaging and smoothing of data. The sequence can be varied as desired.

20 Data Reporting and Sharing

Characterization data typically reports average CIELAB data as a function of patch ID and a cumulative distribution function of pooled CIELAB ΔE. It is recommended that the average spectral data be reported for further analysis of the press run and various reproduction objectives. The relationship between patch ID and CMYK tone values is contained in the appropriate target standard (e.g. ANSI IT8.7/3, ANSI IT8.7/5, etc.) and may be copyrighted information. It is not a necessary part of a characterization data set.

The data files should conform to one of the data reporting formats:ISO 17972-3 CxF/X-3 for XML or ISO 28178/CGATS.17 for ASCII text.

21 Peer Review and Testing Guidelines

Upon completion of collection and adjustment, data should be submitted for peer review and testing. Review should consist of analysis by other experts, as well as verification press runs to validate that the data is representative and can be achieved on a majority of machines.

The review process should be conducted for a specific time period and/or number of verification runs. Once the characterization data has been reviewed, needed adjustments or changes can be made, as well as retesting and reviewing the proposed final data. Findings may take the process back to the analysis step for further refinement. When the involved individuals and team have agreed the characterization data is valid and complete it may be released for distribution.

22 Using a Dataset as a Reference

A document should be created explaining how users can best use the data as a reference. This should describe the workflow for using the data and profiles. If the dataset is based on an existing dataset or method the document should describe the relationship, and which datasets and profiles to use, and how to use each properly within the workflow. If this a new dataset and is unrelated to prior datasets then the document should explain what is different about this dataset, use cases, and the recommended workflow for using the dataset and profiles.

22.1 Distribution of data

Upon completion of the project the characterization data is to be distributed for use. This should include characterization data, ICC profiles, and information on the characterization data as well as use cases. In some cases, the distribution may not contain an ICC profile or a dataset.

Data is usually packaged and distributed as the following:

- CGATS file containing IT8.7/5 data
- ICC profile created from the above data
- Readme file containing relevant information on the characterization data and use cases

Editors: RE, ML, RC, SS, PP, MU

Appendix: Data Set Correction and Creation of Idealized Characterization Data

Companies that wish to create a G7 or near-neutral-based printing reference for use across a several printers may wish to idealize the gray balance of the final characterization to precisely align to G7 gray balance aims.

When a press is characterized, and then optimized to meet G7 or near-neutral requirements, the resulting dataset will meet G7 targets within required tolerances but will not precisely match the G7 or near-neutral gray-balance targets. In other words, any given characterization, from real press data, that meets G7 or near-neutral requirements will include small deviations or biases from perfect G7 or near-neutral gray balance values.

When using the characterization as reference, users may wish to remove these biases by optimizing the data set to achieve a synthetically perfect G7 or near-neutral gray balance behavior. Using the idealized characterization data would then give other printers the best chance to achieve compliance to the reference characterization while also meeting G7 or near-neutral requirements.