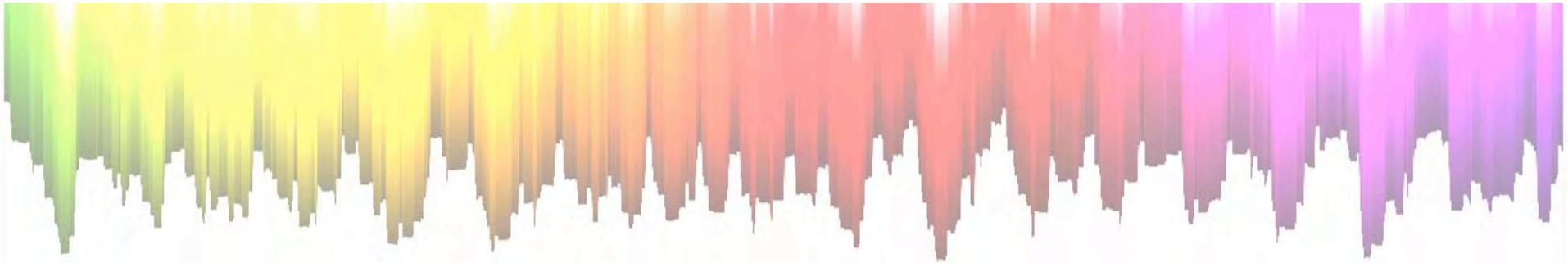


Polynomial Color Reproduction Device Model Term Significance

Davor Donevski, Diana Milcic & Dubravko Banic

Regression models

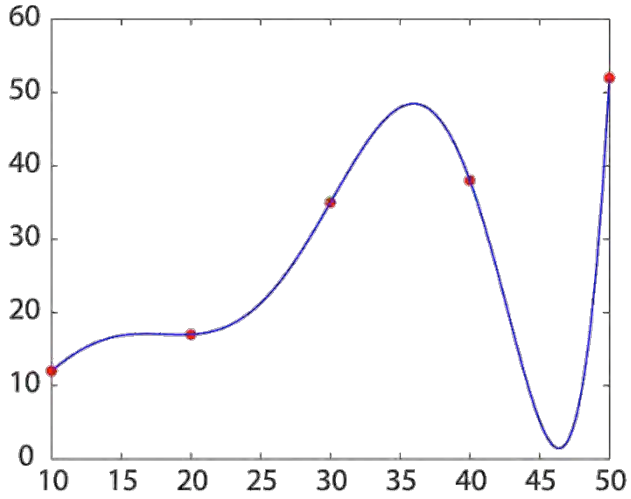
- Arbitrarily chosen function (polynomial)

Important properties:

- data fitting
- prediction power

Overfitting

- model fits the data used to train it (determine its parameters)
- model does not follow general trend of data (poor prediction power)
- overfitting occurs when model order is too high



$$y = a_0 + a_1x + a_2x^2 + a_3x^3 + a_4x^4 + \\ + a_5x^5 + a_6x^6 + a_7x^7$$

Introduction

- Different devices' channels have different curves
- Are all of the model terms significant predictors for any device?

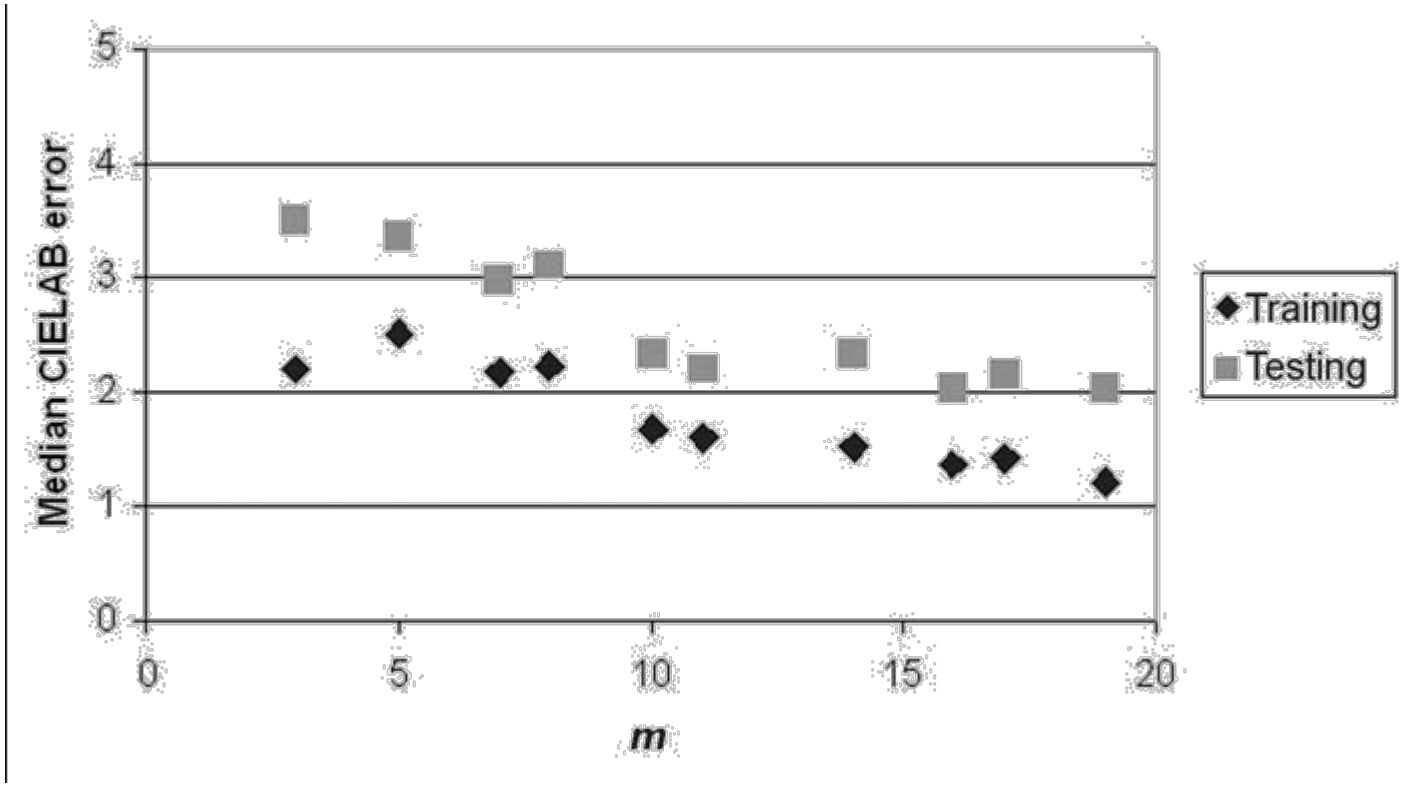
Earlier research

Hong, Luo & Rhodes

Matrix	Mean dE	Max dE
3 x 3	3.11	13.4
3 x 5	1.40	4.5
3 x 6	2.29	11.4
3 x 8	1.33	4.8
3 x 9	1.33	11.5
3 x 11	0.97	3.7

Earlier research

Cheung & Westland



Source: Westland, S., Ripamonti, C., Computational Colour Science Using Matlab, John Wiley & Sons, Chichester, 2004

Aims and hypotheses

Aims:

- determine significance of model terms
- forming optimal models by selecting their terms

Hypotheses:

- chosen terms may increase or reduce model precision
- device data characteristics can be used to select model terms

Devices and materials

Process	Tehnologija	No. inks	Substrate
Process A	Ink-jet piezo	6	Plain paper
Process B	Ink-jet piezo	6	Satinated paper
Process C	Ink-jet thermal	4	Plain paper
Process D	Ink-jet thermal	4	Satinated paper
Process E	Ink-jet thermal	6	Plain paper
Process F	Ink-jet thermal	6	Satinated paper
Process G	Laser	4	Plain paper
Process H	Laser	4	Satinated paper

Methodology

- 4 printing devices and 2 substrates = 8 processes
- characterization with 918 patch chart
- backward elimination (F-test) on maximum models
- evaluation on 918 values independent dataset + psychophysical evaluation

Methodology

Backward elimination (partial F-test)

- eliminating terms $1\ C\ M\ Y\ CM\ CY\ MY\ C^2$ …

- eliminating blocks of terms

  two ways of forming 4th order interaction terms:

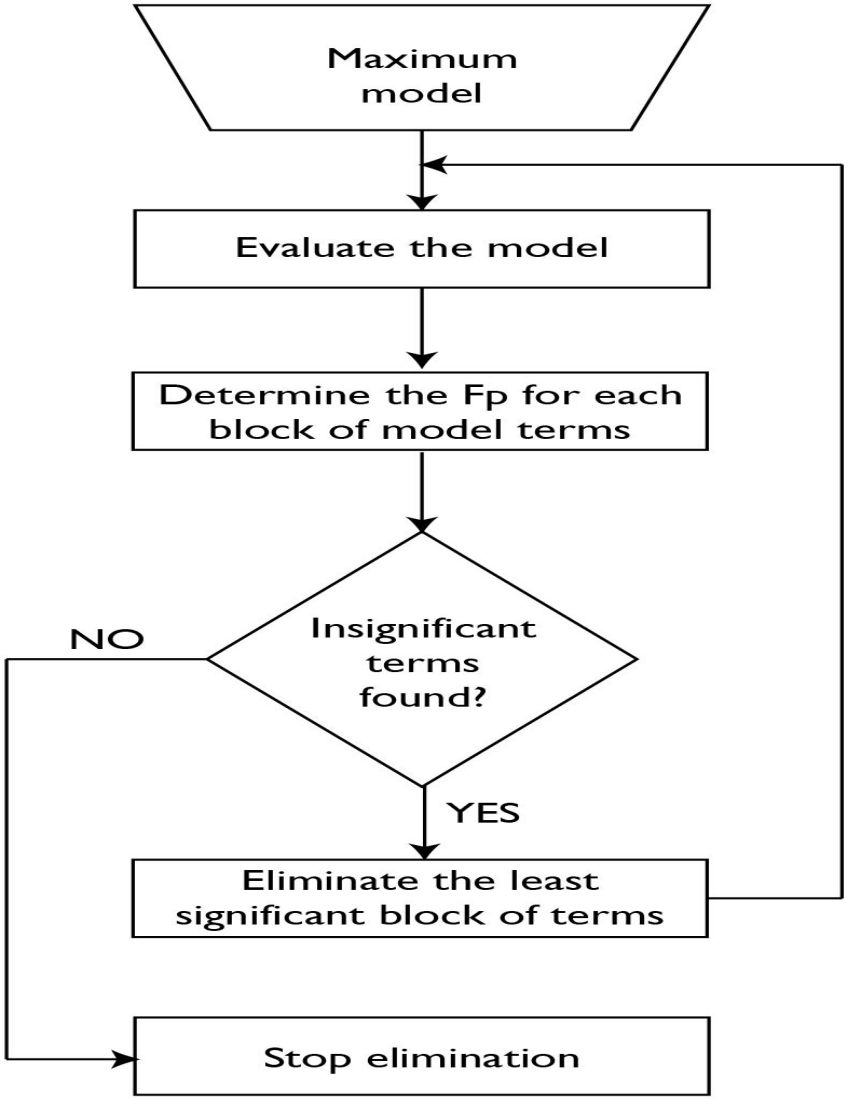
  $C^3M\ C^3Y\ M^3Y\ CM^3\ CY^3\ MY^3$

  $C^2M^2\ C^2Y^2\ M^2Y^2$

Methodology

Backward elimination:

1 C M Y CM CY…



Results

Blocks of terms eliminated throughout elimination procedure steps

Elimination step	Block of terms eliminated
2	C^4 M^4 Y^4
3	C^2 M^2 Y^2
4	C^3M^3 M^3Y^3 C^3Y^3
5	C^4M M^4Y Y^4C C^4Y M^4C Y^4M
6	CM MY CY
7	C^4MY CM^4Y CMY^4
8	C^3M^2Y C^3MY^2 C^2M^3Y CM^3Y^2 C^2MY^3 CM^2Y^3

Results

Statistical

	N – no. patches	dEab	Min	Median	Max	C.I. 95%
Maximum	918	2,45	0,21	1,96	11,58	0,21
Reduded	918	2,50	0,21	2,12	11,23	0,21

Psychophysical

Maximum model performance

Reduced model performance

Results

Other process:

	N – no. patches	dEab	Min	Median	Max	C.I. 95%
Maximum	918	2,14	0,19	1,94	8,19	0,15
Reduded	918	1,77	0,06	1,61	7,39	0,13

Conclusions

- no obvious regularity in eliminated blocks of terms
- slightly worse central tendency measures (elimination at 10% significance level)
- significantly reduced overfitting artefacts (psychophisical evaluation)
- high order polynomials can be used if appropriate terms are chosen (the order itself does not cause overfitting)