

3aSC2 substituted

Calibration and Fusion

Geoffrey Stewart Morrison

FORENSIC VOICE COMPARISON LABORATORY
SCHOOL OF ELECTRICAL ENGINEERING & TELECOMMUNICATIONS

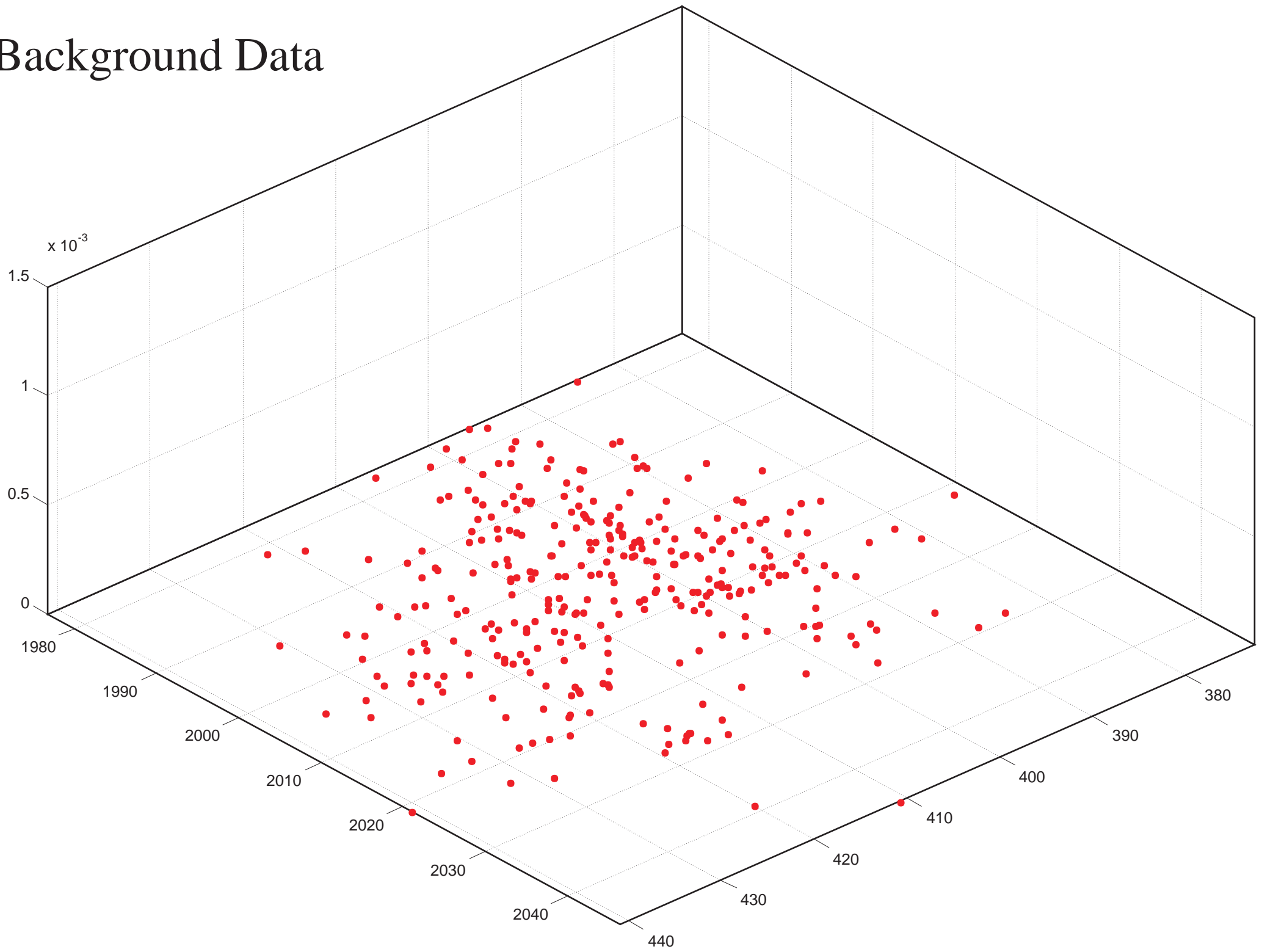


UNSW

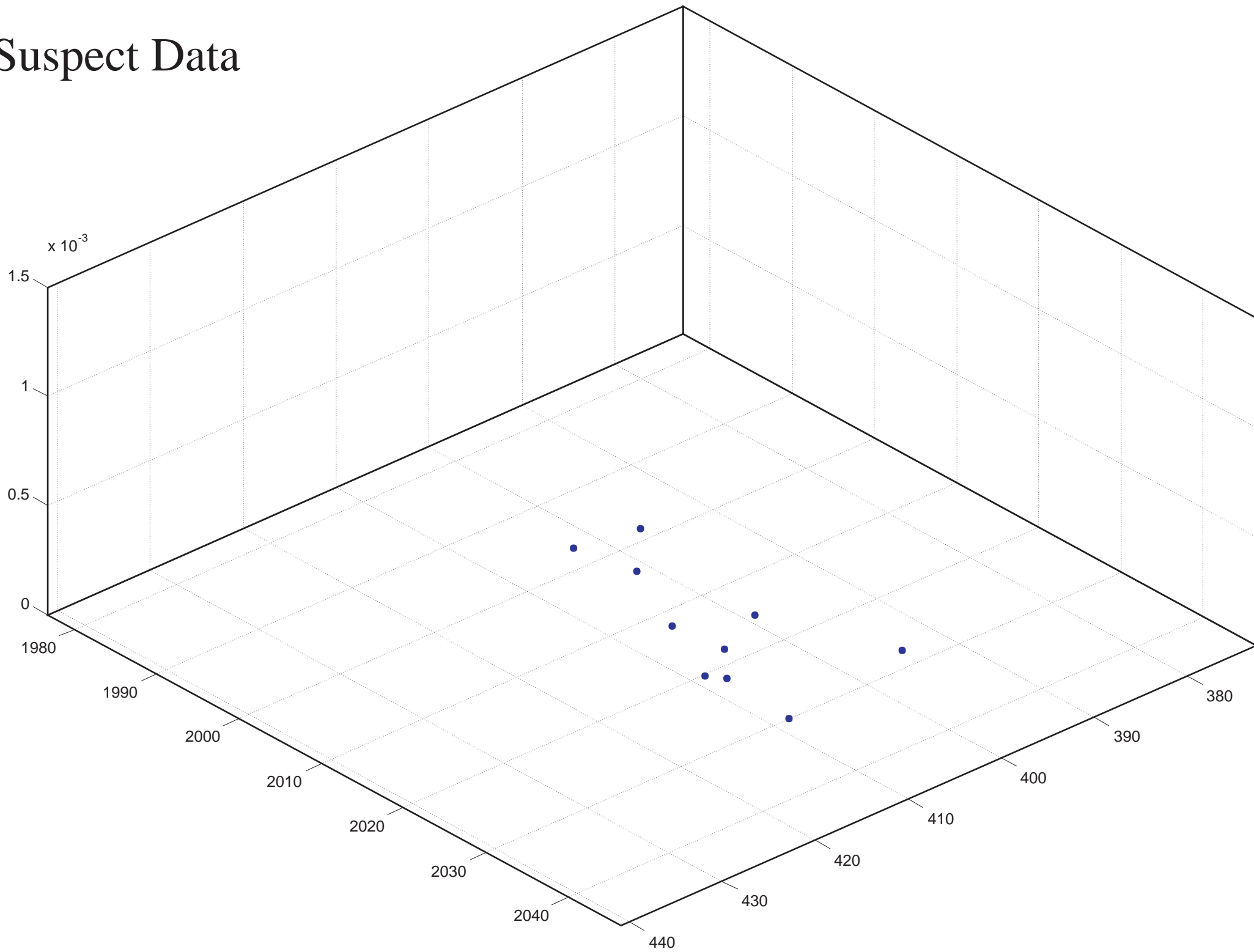
THE UNIVERSITY OF NEW SOUTH WALES
SYDNEY • AUSTRALIA

Gaussian Mixture Model Universal Background Model (GMM-UBM)

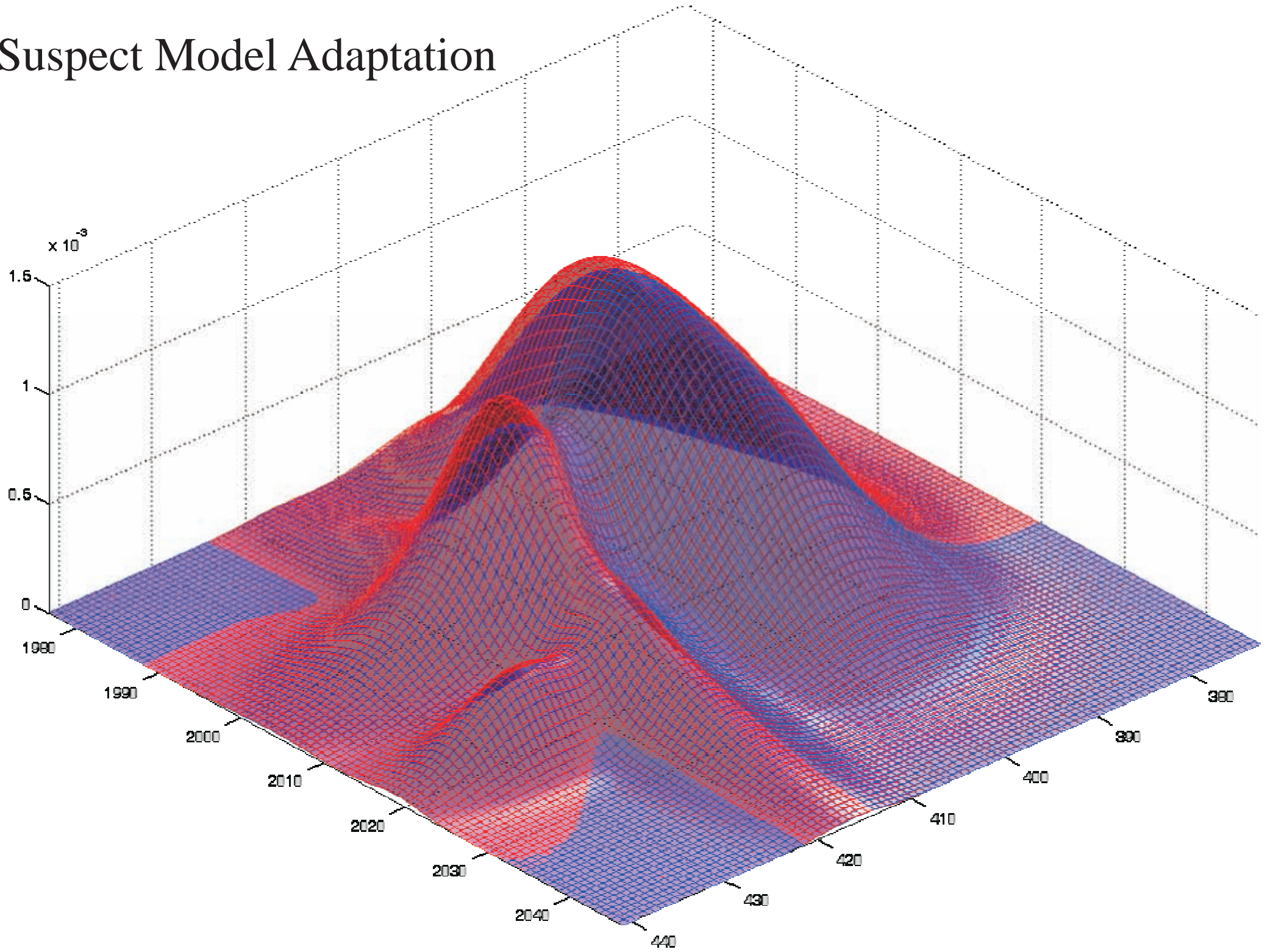
Background Data



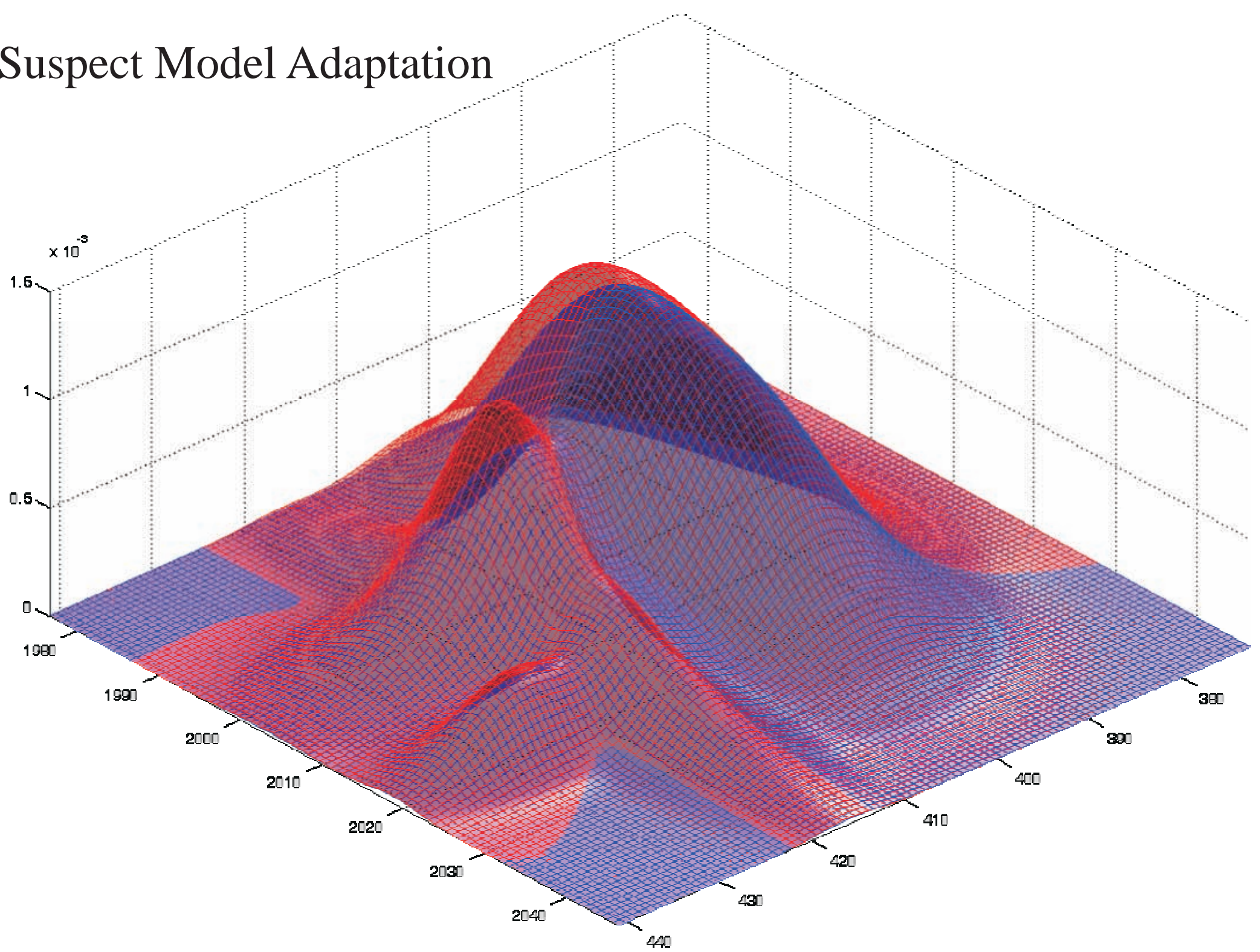
Suspect Data



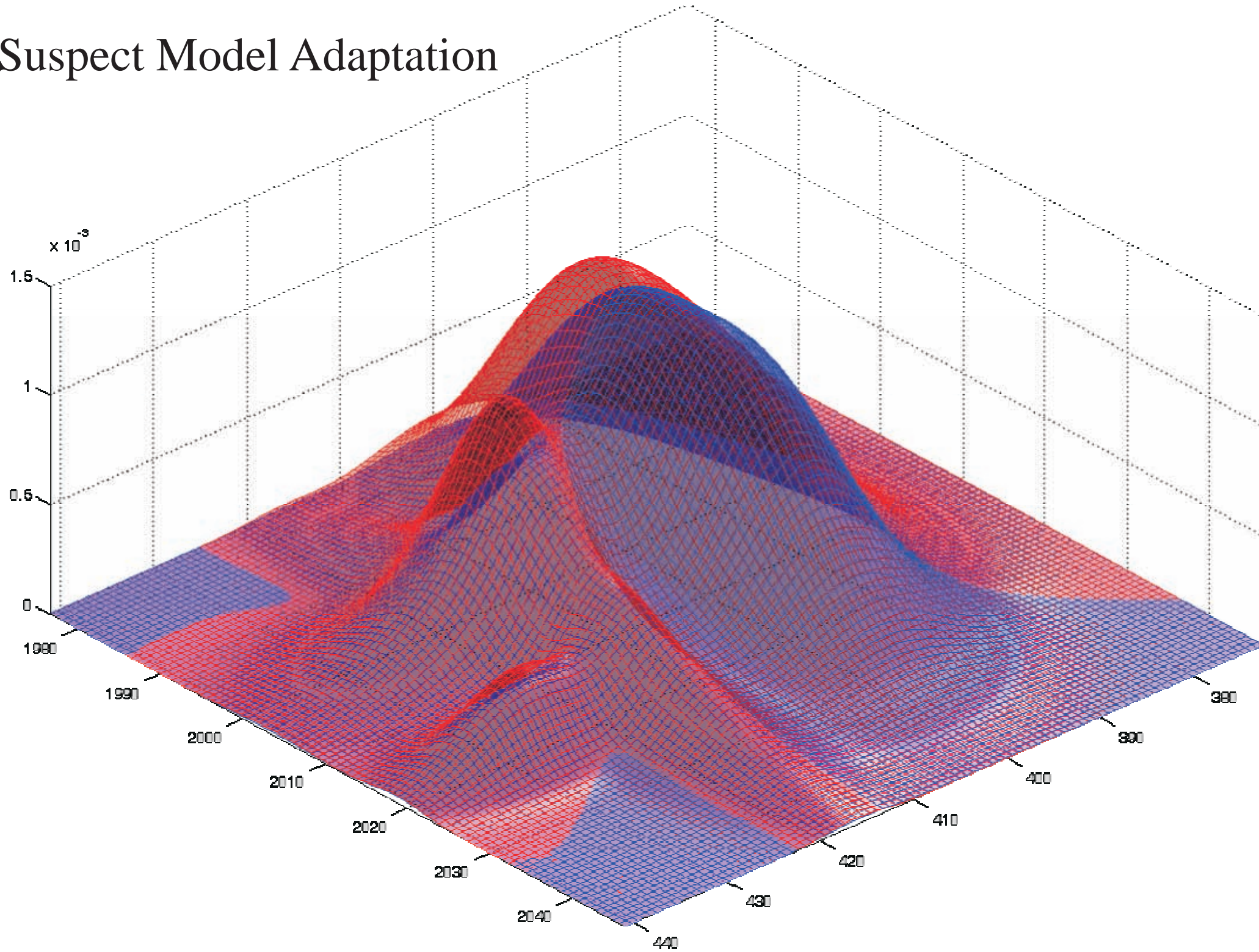
Suspect Model Adaptation



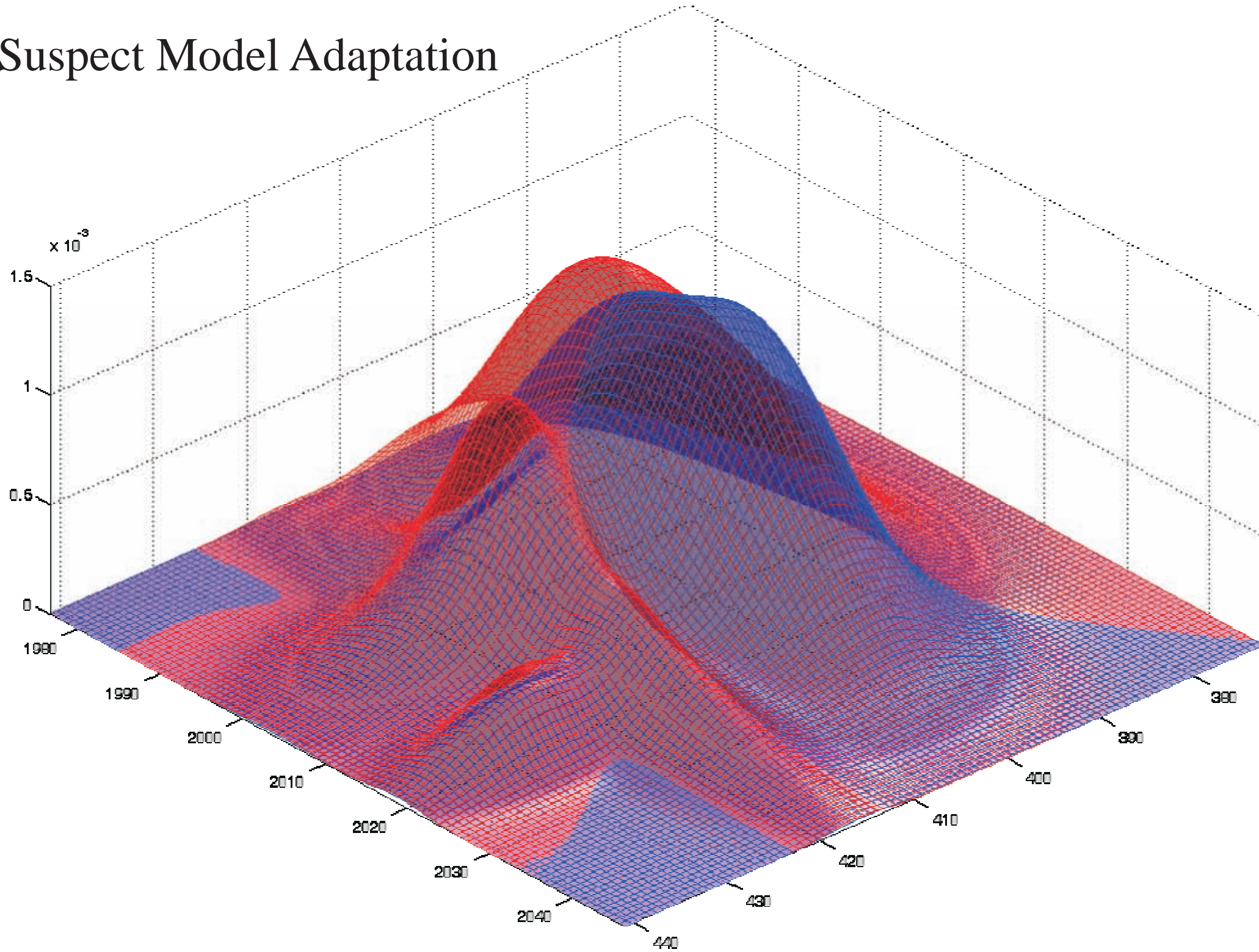
Suspect Model Adaptation



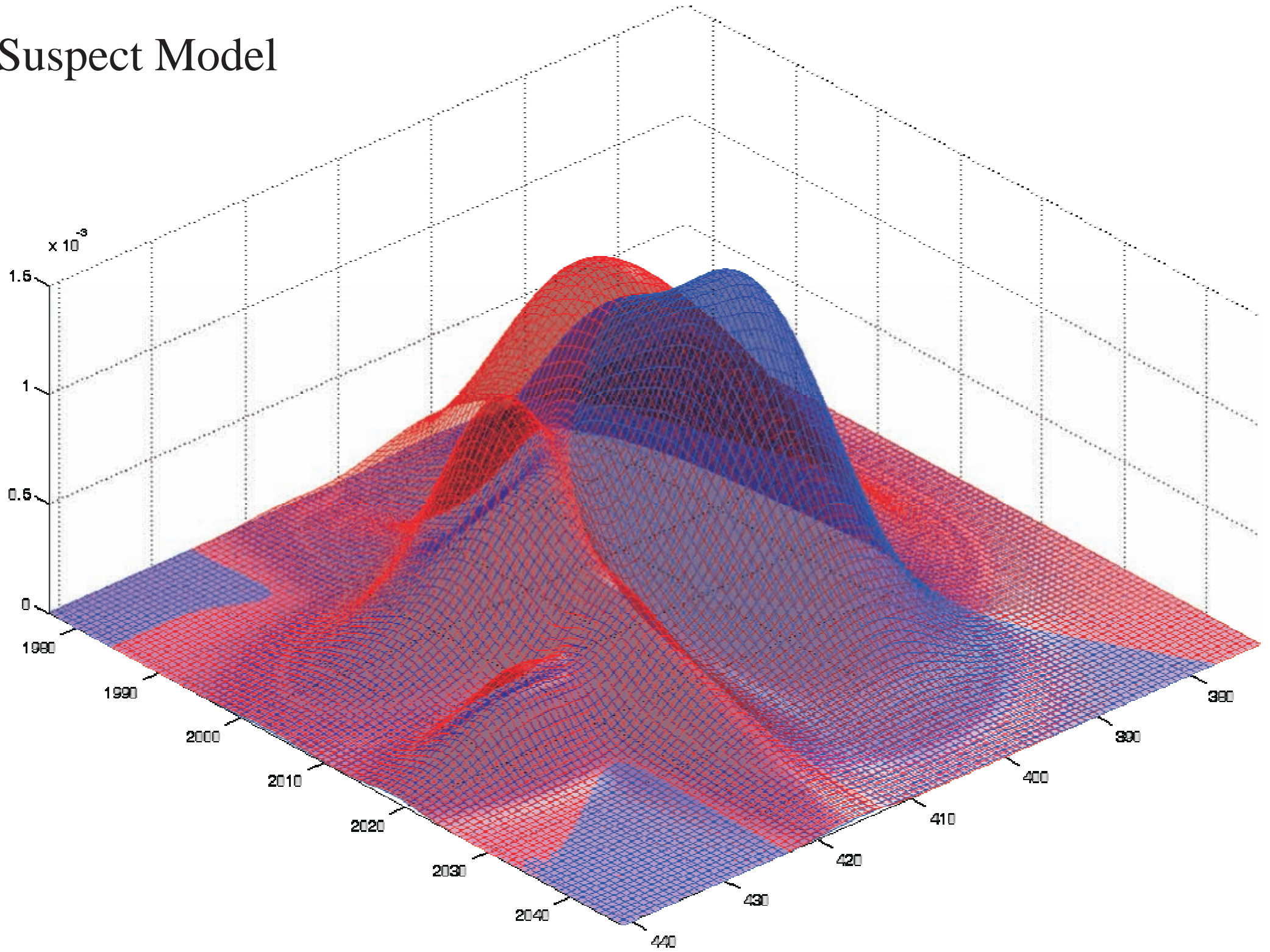
Suspect Model Adaptation



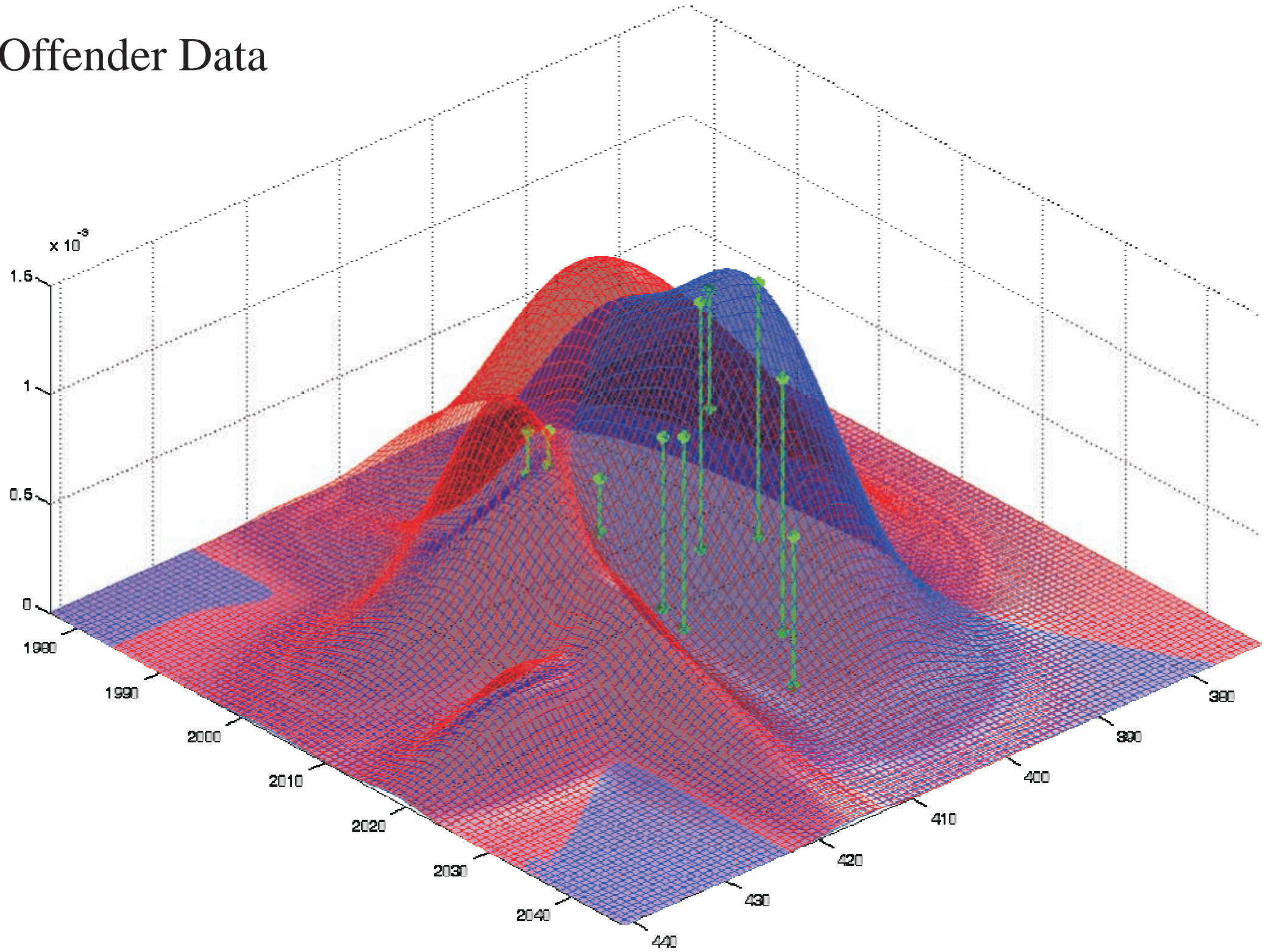
Suspect Model Adaptation



Suspect Model



Offender Data



Calibration and Fusion

Calibration and Fusion:

- Scores quantify the similarity of a pair of samples while also taking their typicality into account, but they are simply scores, they are not interpretable as log likelihood ratios.
- Calibration converts a single score into a log likelihood ratio.
 - scores from multiple offender data points in GMM-UBM
- Fusion converts a parallel set of scores into a log likelihood ratio.
 - scores from multiple systems run on the same voice samples
 - scores from multiple phonetic units in the same voice recordings
- Logistic regression
 - with equal priors (prior odds = 1) optimisation objective is C_{llr}

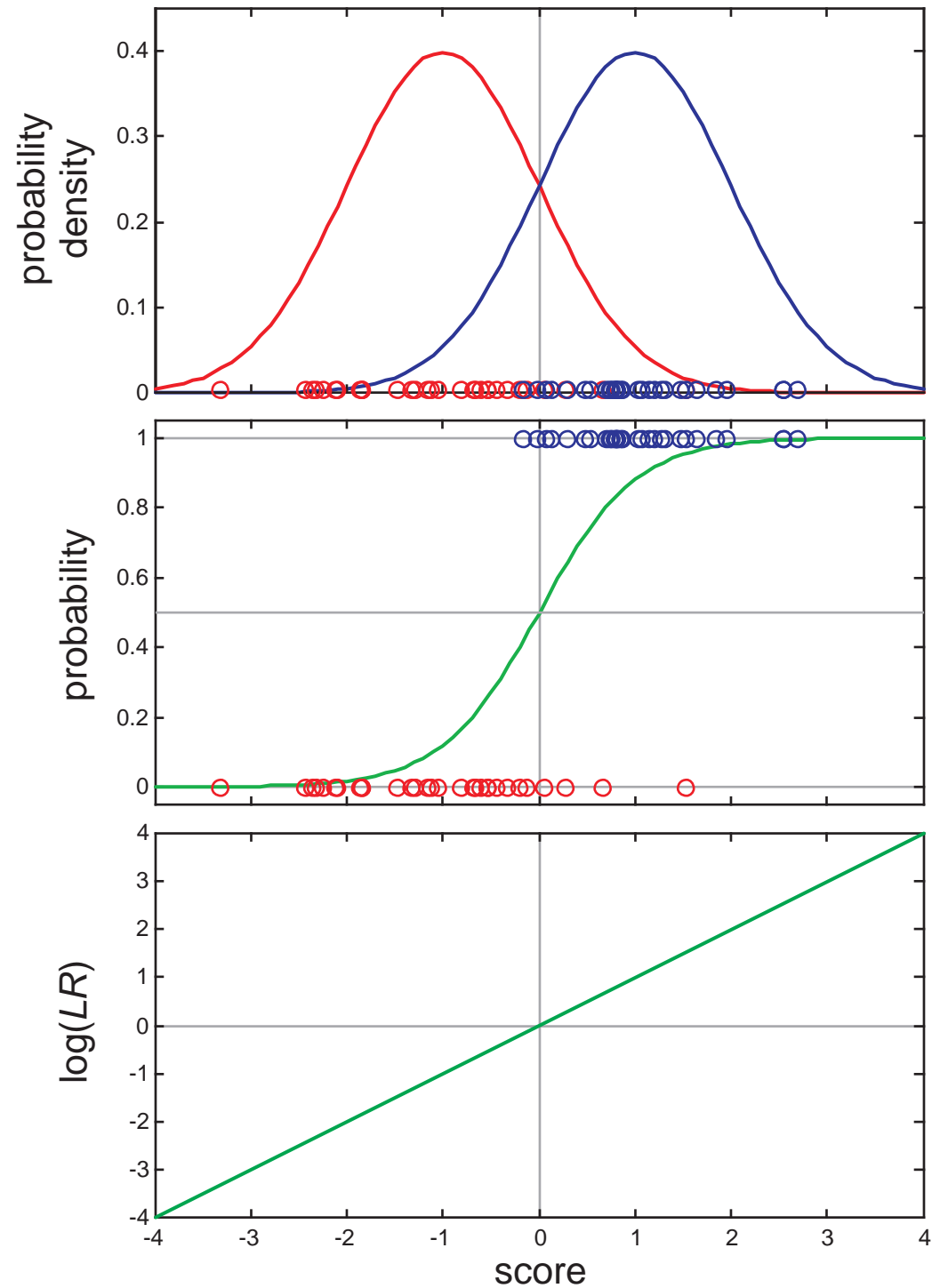
Calibration

Univariate-Logistic-Regression Calibration:

Brümmer & du Preez (2006), van Leeuwen & Brümmer (2007)

- Linear shifting and scaling of scores
- Heavily penalises scores which provide strong support for contrary-to-fact hypotheses
 - gradient
- Log likelihood ratios as output

Calibration



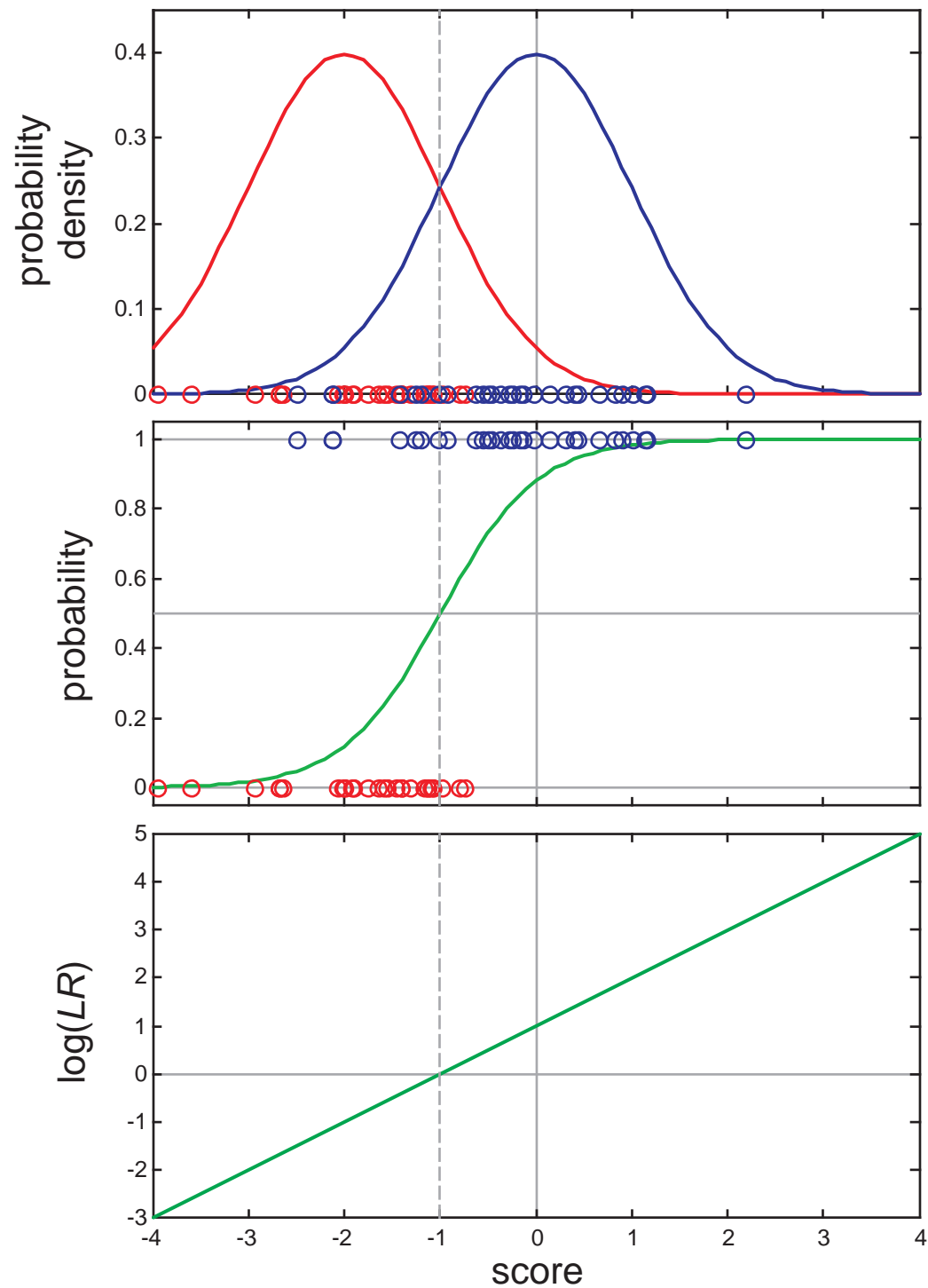
$$LR = \frac{p}{1-p}$$

$$y = a + bx$$

$$a = 0$$

$$b = 1$$

Calibration

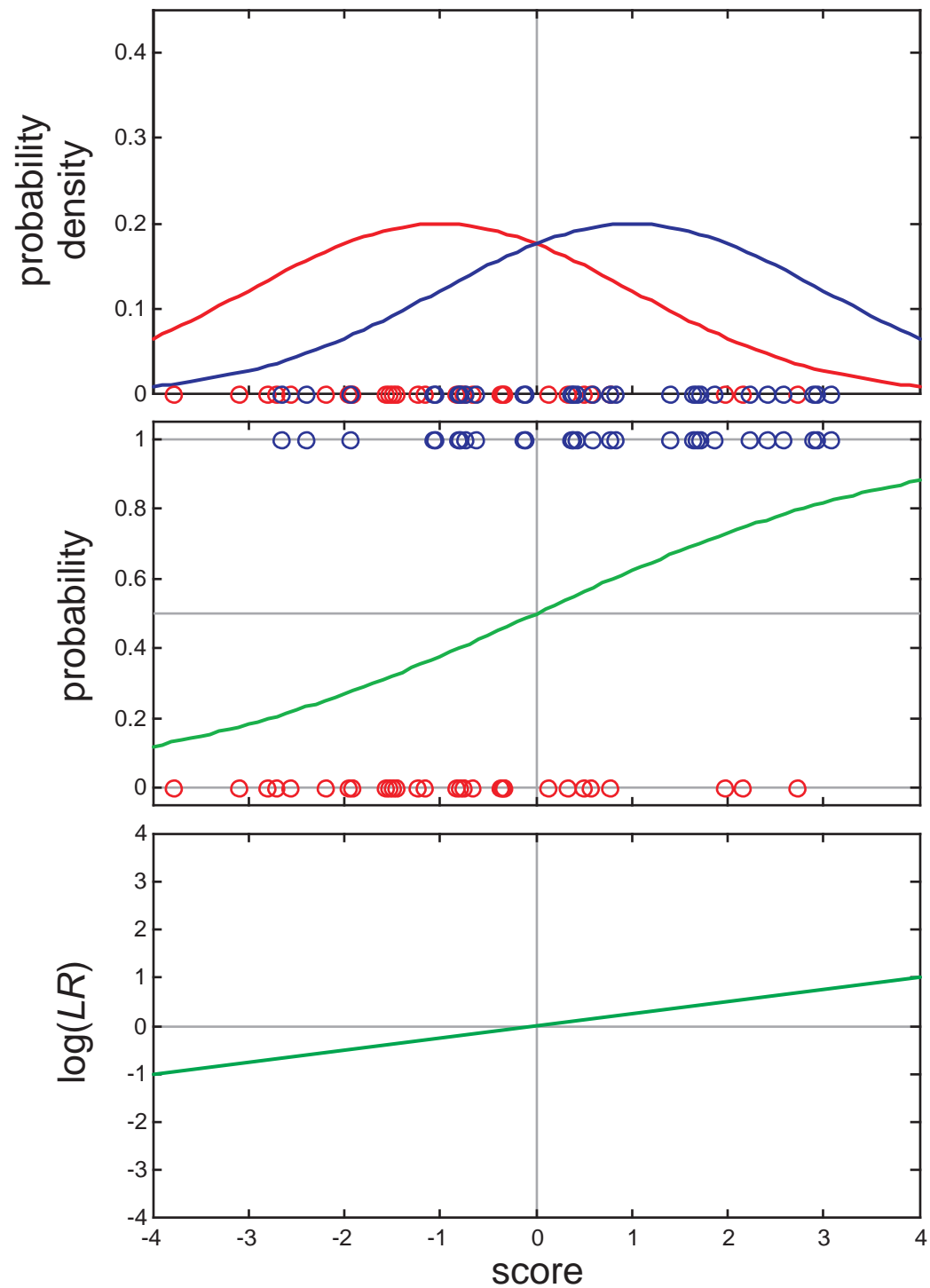


$$y = a + bx$$

$$a = 1$$

$$b = 1$$

Calibration



$$y = a + bx$$

$$a = 0$$

$$b = 1/4$$

Calibration Example Morrison, Zhang, & Rose (2010)

- 2 non-contemporaneous recordings of each of 64 male speakers of Standard Chinese
- Answering questions
- Telephone recordings
- Mean formant values (F1, F2, F3) of /a/ tokens
- ~10 tokens from each recording
- Multivariate kernel density (MVKD) procedure for calculation of likelihood ratios Aitken & Lucy (2004)
- Cross validated

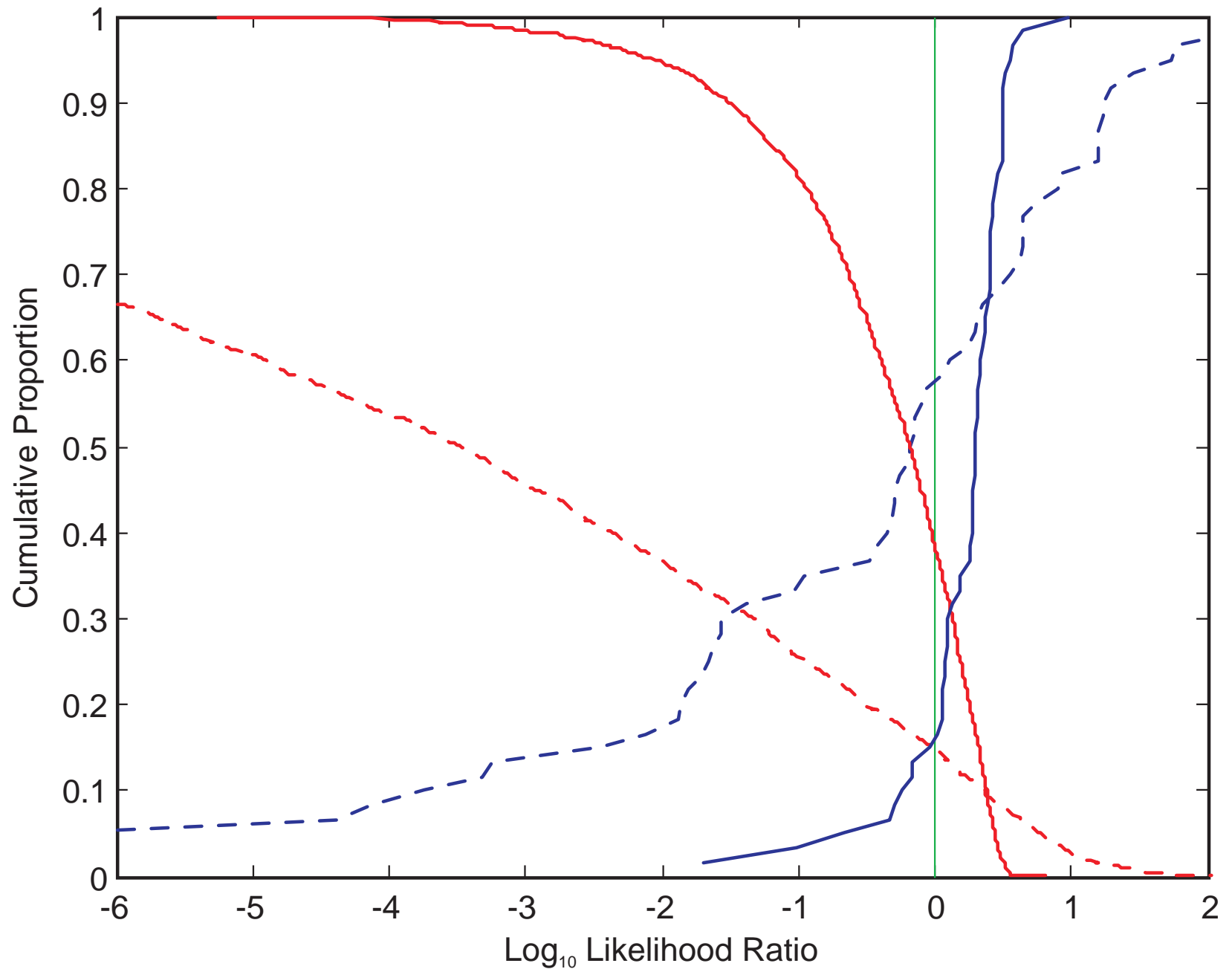
Calibration Example

Raw

$$C_{llr} = 1.802$$

Calibrated

$$C_{llr} = 0.750$$



Fusion

Multivariate-Logistic-Regression Fusion:

Pigeon, Druyts, & Verlinde (2000), Brümmer et al. (2007)

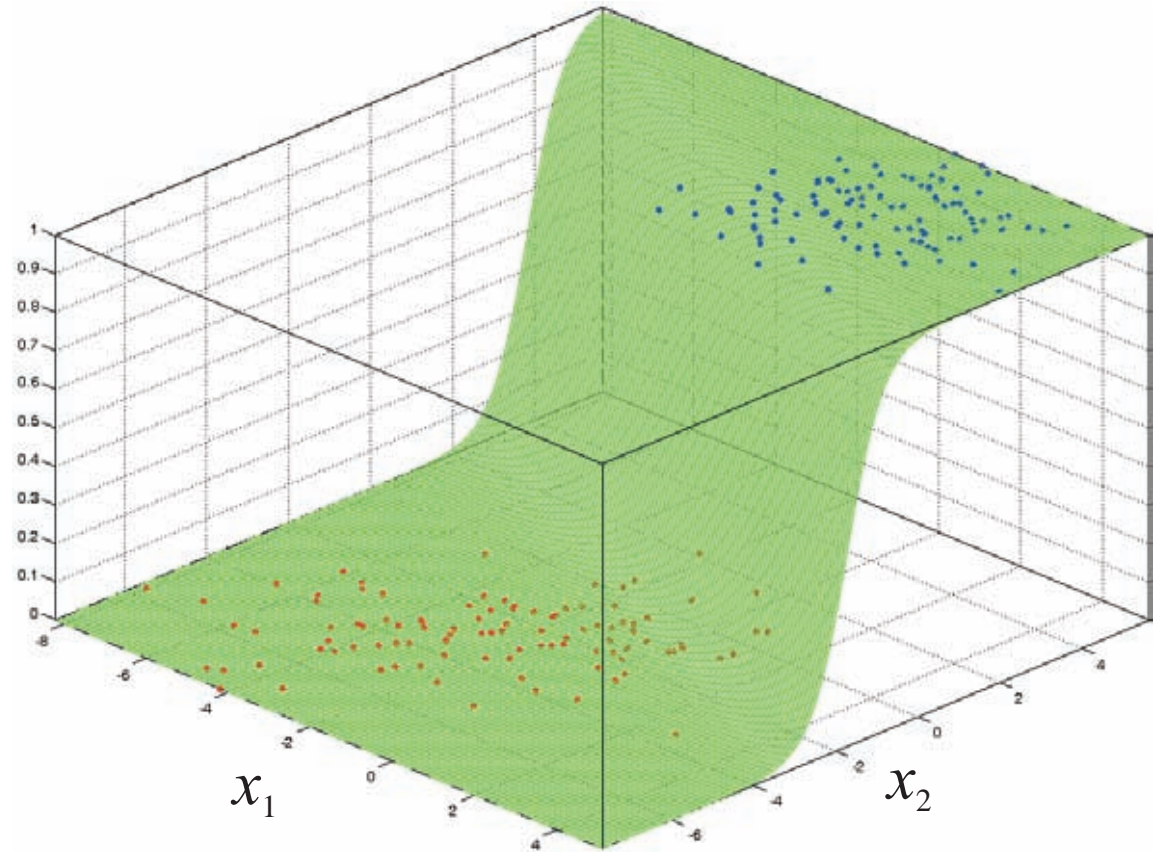
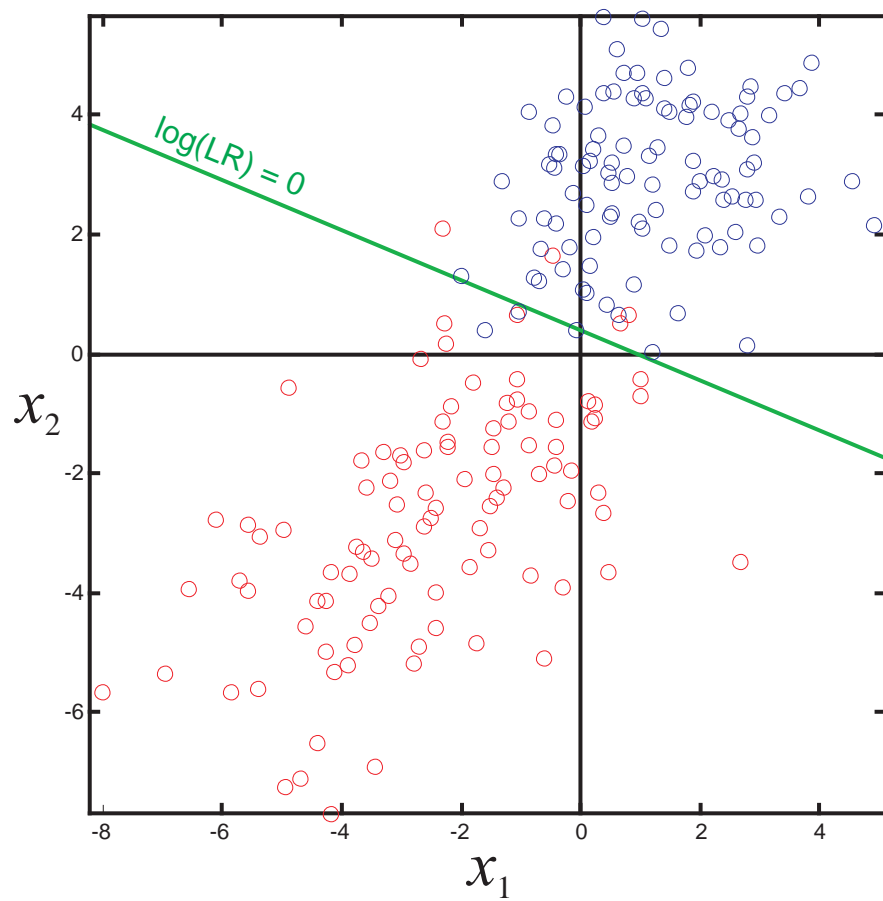
- Combines parallel sets of scores
 - scores from multiple systems run on the same voice samples
 - scores from multiple phonemes in the same voice recordings
- Takes account of correlation
 - weighted combination of scores
 - $y = a + bx_1 + cx_2 \dots$
- Calibrates
- Log likelihood ratios as output

Fusion

- Parallel sets of scores

	x_1	x_2
Spk01 v Spk01	+5	+1
Spk01 v Spk02	-3	-2
Spk01 v Spk03	-1	-1
⋮		
Spk02 v Spk02	+6	+3
Spk02 v Spk03	+1	-2
⋮		

Fusion



$$y = a + bx_1 + cx_2$$

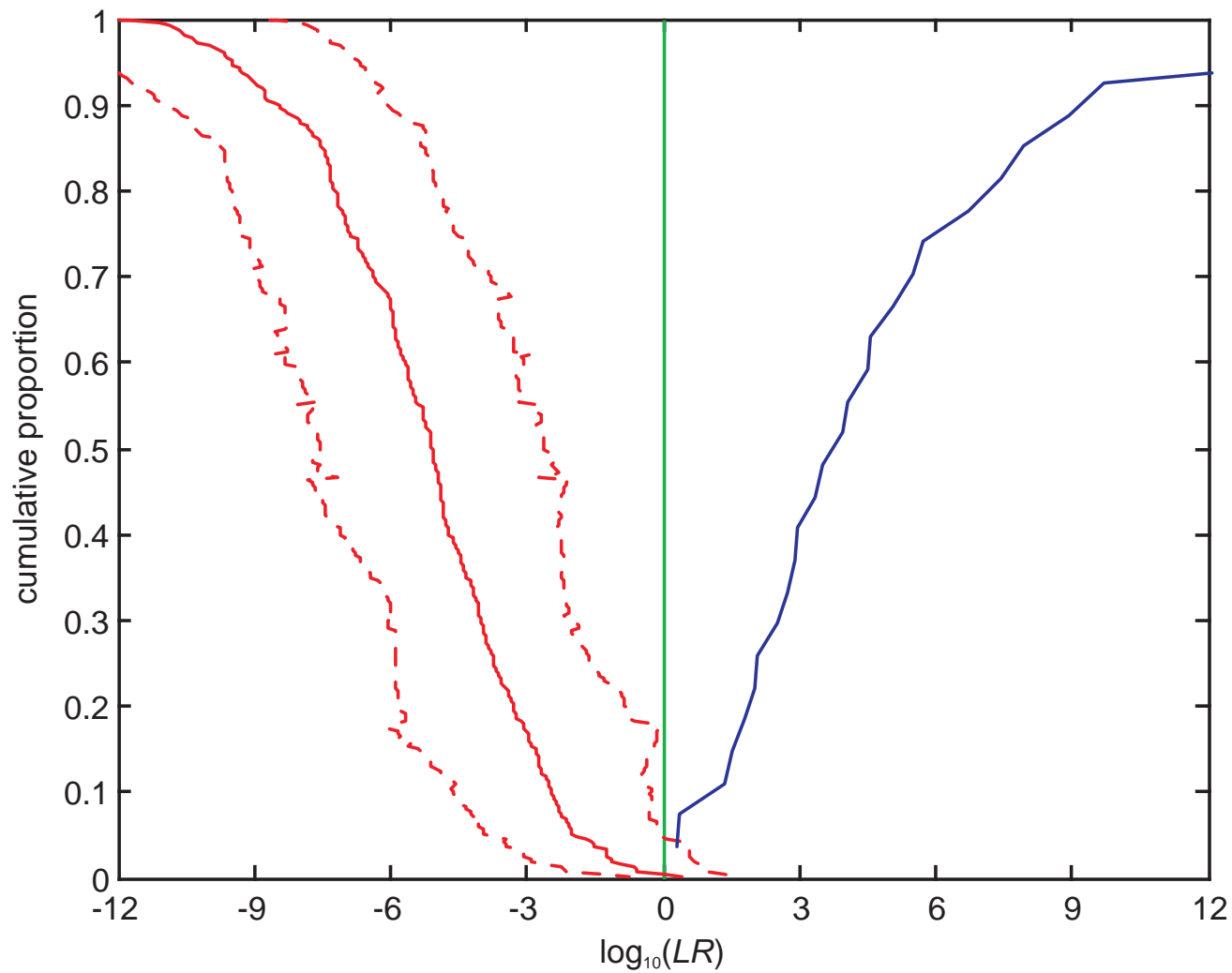
$$a = -0.93 \quad b = +0.97 \quad c = +2.32$$

Fusion Example Morrison (2010)

- 2 non-contemporaneous recordings of each of 27 male speakers of Australian English
- Read speech
- Studio recordings
- 3rd-order DCTs fitted to F2 trajectories of tokens of /aɪ/, /eɪ/, /oʊ/, /aʊ/, and /ɔɪ/
- 16–20 tokens of each phoneme from each recording
- GMM-UBM used to calculate likelihood ratios Reynolds, Quatieri, & Dunn (2000)
- Cross validated

Fusion Example

system	C_{llr}
/aɪ/	0.375
/eɪ/	0.367
/oʊ/	0.326
/aʊ/	0.455
/ɔɪ/	0.311
fused	0.035



Thank You

<http://geoff-morrison.net>

<http://forensic-voice-comparison.net>

<http://forensic.unsw.edu.au>