

Internet Performance, Measuring and Modelling

Iain W. Phillips and José A. Hernández

Computer Science

Loughborough University

`I.W.Phillips@lboro.ac.uk, J.A.Hernandez@lboro.ac.uk`

September 2004

1 Introduction

Double act:

- History of work at Loughborough
- Other Measurement projects
- Visualisation of Measurements
- *Mathematical Modelling*
- *Applications*

2 Ancient History

- In 1994 JANET → SuperJanet, contract won by BT
- Built over SMDS—Switched Multi-megabit Data Service, and ATM networks
- University Research Initiative—Managing Multiservice Networks

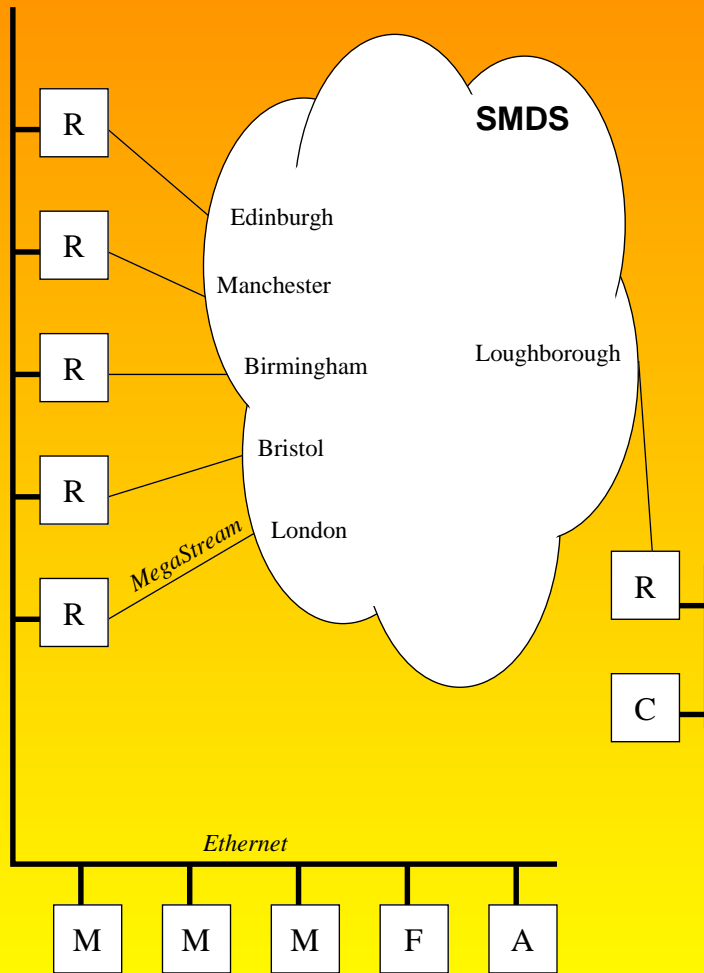
2.1 URI—MMN

- 6 Universities:
 - Loughborough EE
 - University College London CS
 - Imperial London CS
 - Oxford Brookes University CMS
 - Cambridge University CL
 - Lancaster University CS
- 5 Years, £1.5 million
- 1994–1999

2.2 MMN—Loughborough

- Performance Monitoring and Measurement
- Researched and built a delay measurement tool
- Active Sender
- Used GPS for synchronisation
- Accurate to about $10\mu\text{s}$

3 Performance Monitoring



3.1 What is performance?

- Loss
- Delay
- Throughput

3.2 What causes performance problems?

- Routing misconfiguration
- Link or Node failure
- Aggressive Applications
 - Peer-to-peer, video streaming, online gaming etc
- Denial of Service attacks

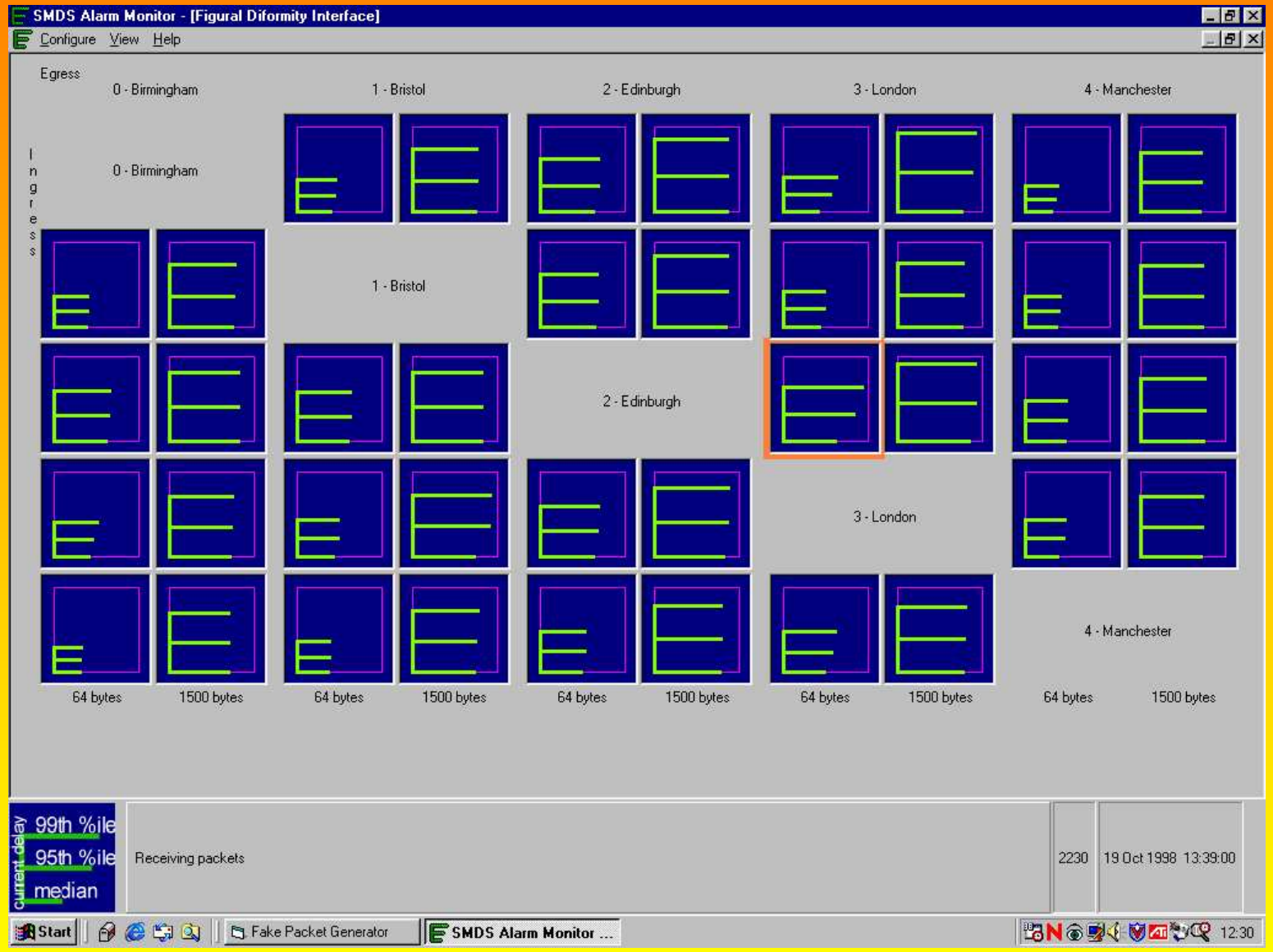
3.3 Monitoring Project

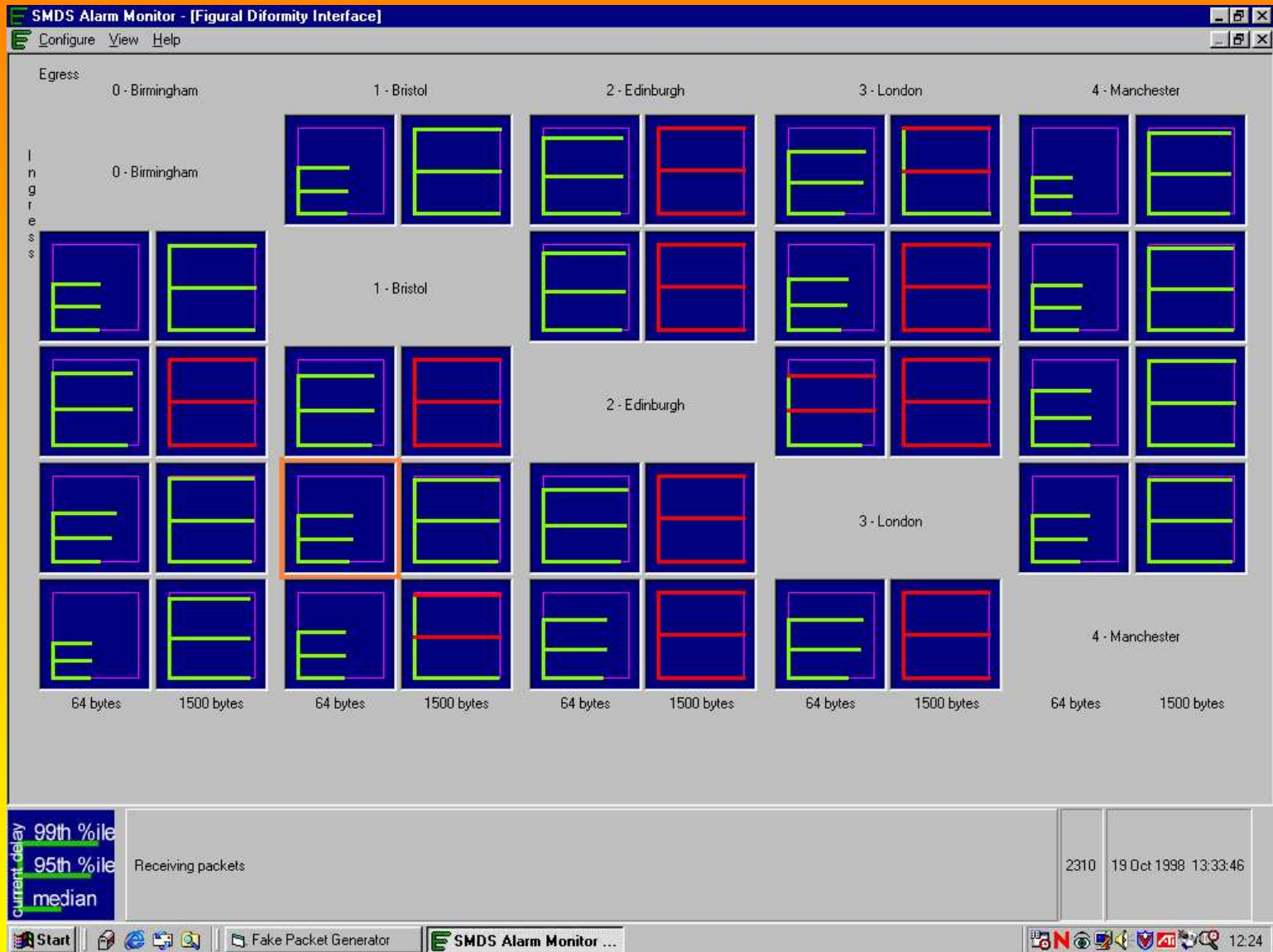
Separate funding:

- To produce weekly and monthly reports on performance
- From 1996 to the present (and just renewed)
- Achieved many further developments

4 Visualisation

- Tools to reduce working load of network operators
- FDV—Figurable Deformity Visualisation
- TMT—Trunk Monitoring Tool

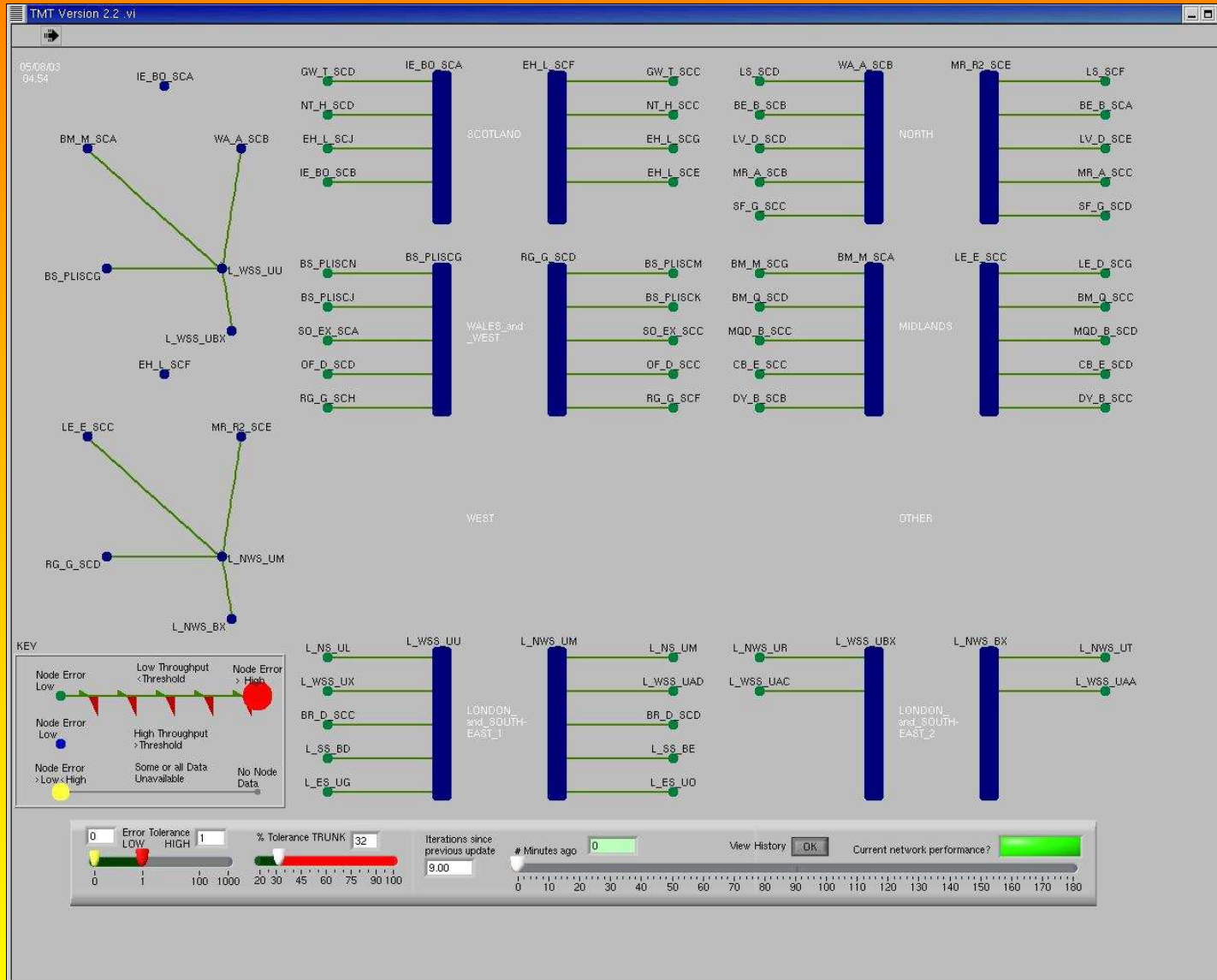


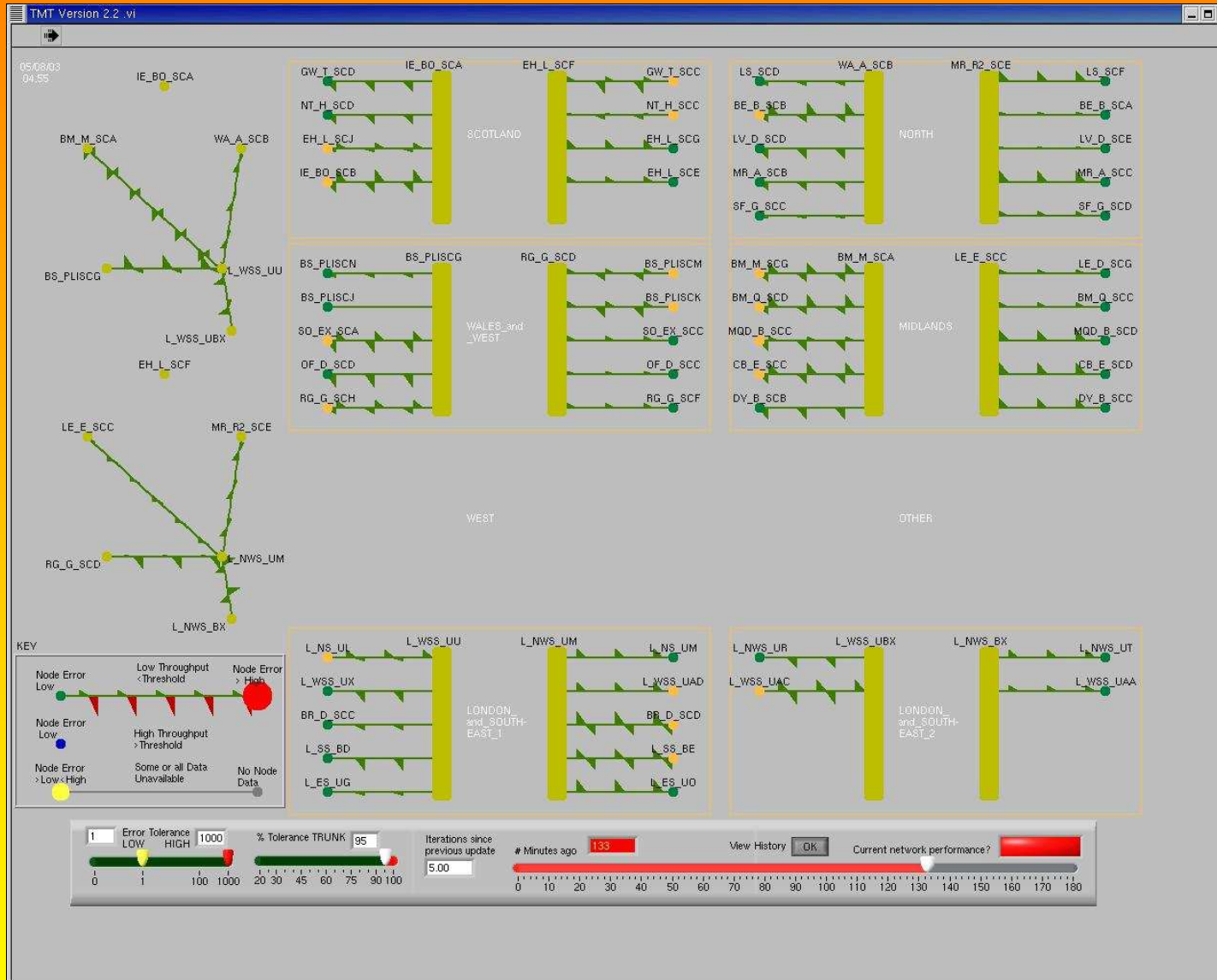


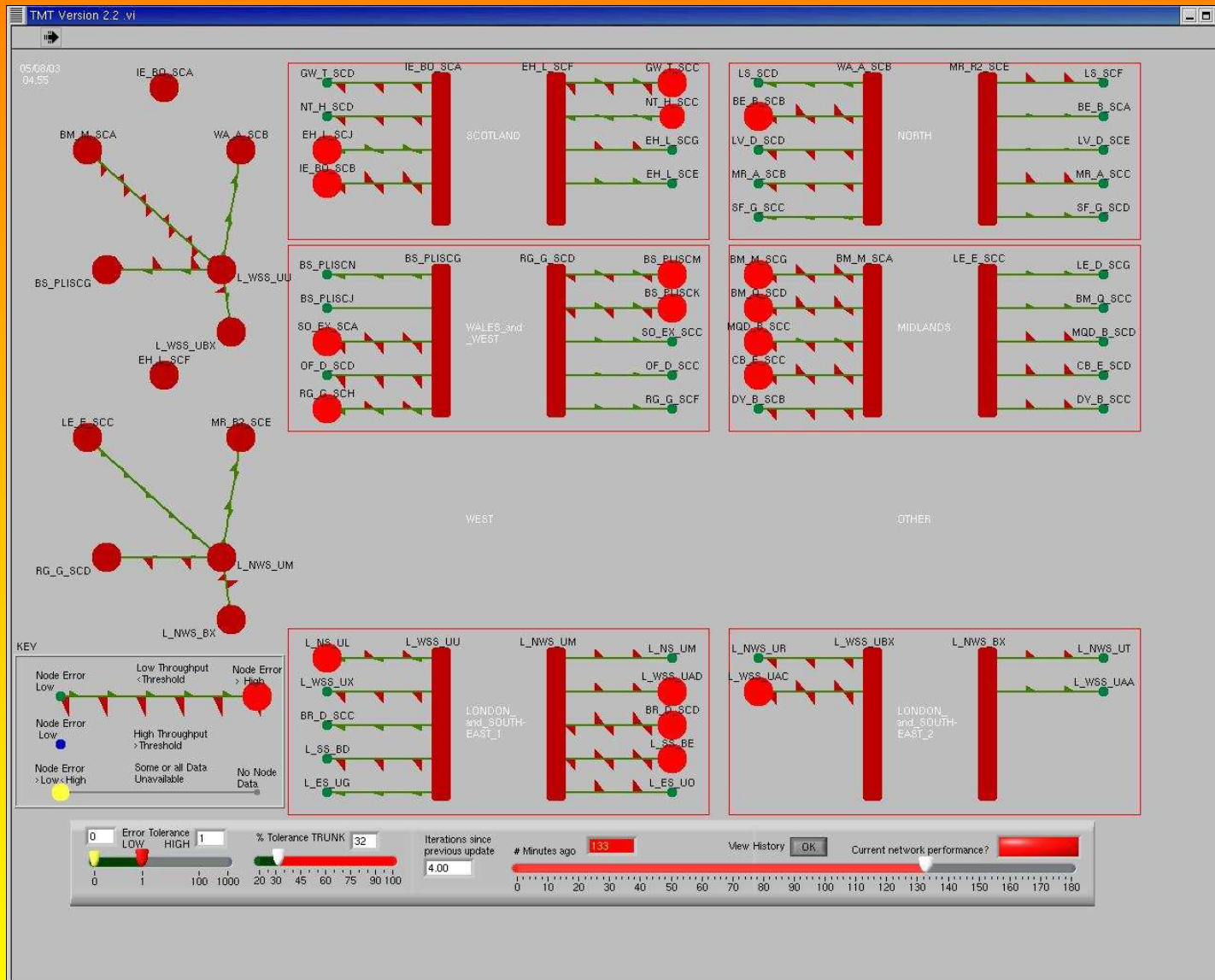
4.2 TMT

Trunk Monitoring Tool

- Uses SNMP to query trunk information from SMDS switches
- Presents this in a “single-look” view to operators.
- Deployed April 200







4.3 Exceptions

Interesting Network Events

Detected by:

- Manual
- Rule-based
- Neural networks

All based on simple statistics, max in day, min in day, mean, max - min, variance etc

5 Other Monitoring Projects

- RIPE-NCC—Monitoring (mostly) European Delays
- SPRINT (US)—Monitoring for Traffic Engineering
- NLANR—Traceroute/ping delays
- Waikato (NZ) DAG hardware traffic capture
- Cambridge/Loughborough (EE) passive monitoring
- *new* UKLIGHTmas

6 What to do next ...

- Can statistics/mathematics improve such displays?
- Can we predict Internet performance like the weather?
- How do we model?

7 Contents

- Motivation
- Traffic modelling review
- Mixing Weibull distributions
- Expectation Maximisation algorithm
- Experiments and results
- Applications and discussion

8 Motivation:

The need to model network performance

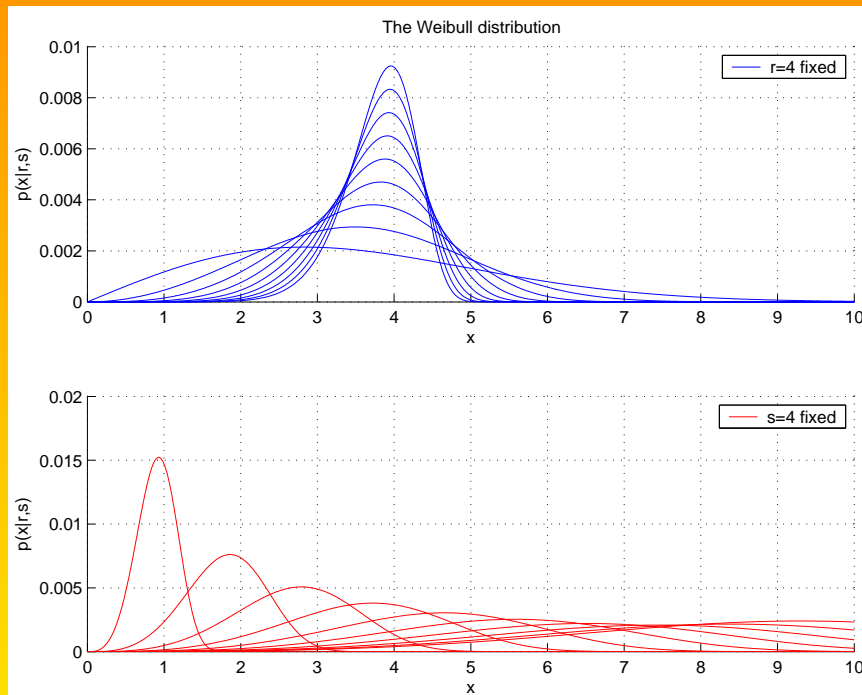
- Metrics to define network performance
- Low-level quantities: delay and loss
- End-to-end network performance status
- Packet probes such as ping or one-way delay UDP packets

9 Previous work:

Traffic modelling and delay distributions

- Network traffic shows self-similarity and long-range dependency.
- Current traffic strategies search for models compliant to these empirical properties: fBm, fARIMA, FSD, etc.
- When inputting such traffics into routers, the queue distribution exhibit heavy-tail distributions. Such distribution can be approximated to Weibull for the particular case of fBm.
- Such result has been previously validated in a single hop scenario.
- Our aim is to model multiple-hop (or end-to-end) delays with a combination of several Weibull distributions.

10 Mixing Weibull distributions: The Weibull distribution



- $p(x|r, s) = \frac{sx^{s-1}}{r^s} \exp\left(-\left(\frac{x}{r}\right)^s\right)$
- r is concerned with the mode location.
- s is related to tail behaviour.

11 Mixing Weibull distributions:

Problem statement

- Let us assume we are given a sample of N delay measurements $\mathbf{x} = [x_1, \dots, x_N]$, which are supposed to be drawn from M Weibull distributions: $[p(x|\theta_1), \dots, p(x|\theta_M)]$
- The result is: $p(x|model) = \sum_{j=1}^M \alpha_j p(x|\theta_j)$
- $\alpha_j =$ weight of the j -th component of the mixture. Obviously, $\sum_j \alpha_j = 1$
- $\theta_j = [r_j, s_j]$ shape and scale parameters of the j -th Weibull distribution
- Finding α and θ appropriate to best fit delay histograms represented by the measurements sample \mathbf{x}

12 Mixing Weibull distributions: Expectation Maximisation

- To proceed, second random variable y , referred to as labels, is necessary to complete the problem formulation.
- $p(y_i = j|x_i, \Theta)$ = the probability of data x_i being drawn from the j -th component of the mixture. Obviously,
 - $p(x_i|y_i = j, \Theta) = p(x_i|\theta_j)$, and
 - $p(y_i = j|\Theta) = \alpha_j$
- With this formulation EM defines an iterative procedure to obtain the maximum likelihood estimates, based on two steps:
 - E-step: $Q(\Theta, \Theta^{(t)}) = E[\log L(\Theta|x, y)|\mathbf{x}, \Theta^{(t)}]$
 - M-step: $\Theta^{(t+1)} = \arg \max_{\Theta} Q(\Theta, \Theta^{(t)})$

13 Mixing Weibull distributions: Computing EM

- Expanding E-step:

$$Q(\Theta, \Theta^{(t)}) = \sum_{j=1}^M \sum_{i=1}^N (\log p(x_i|\theta_j)) p(y_i = j|x_i, \Theta^{(t)}) + \\ + \sum_{j=1}^M \sum_{i=1}^N (\log \alpha_j) p(y_i = j|x_i, \Theta^{(t)})$$

- Maximising:

$$\frac{\partial Q(\Theta, \Theta^{(t)})}{\partial \alpha_j} = 0$$

$$\frac{\partial Q(\Theta, \Theta^{(t)})}{\partial \theta_j} = 0$$

14 Mixing Weibull distributions:

EM applied to mixtures of Weibull distributions

1. Computing parameters:

$$\alpha_j = \frac{1}{N} \sum_{i=1}^N p(y_i = j | x_i, \Theta)$$

$$r_j = \left(\frac{\sum_{i=1}^N x_i^{s_j} p(y_i = j | x_i, \Theta)}{\sum_{i=1}^N p(y_i = j | x_i, \Theta)} \right)^{1/s_j}$$

$$s_j = \frac{\sum_{i=1}^N p(y_i = j | x_i, \Theta)}{\sum_{i=1}^N \left(\frac{x_i^{s_j}}{r_j^{s_j}} - 1 \right) \log \left(\frac{x_i}{r_j} \right) p(y_i = j | x_i, \Theta)}$$

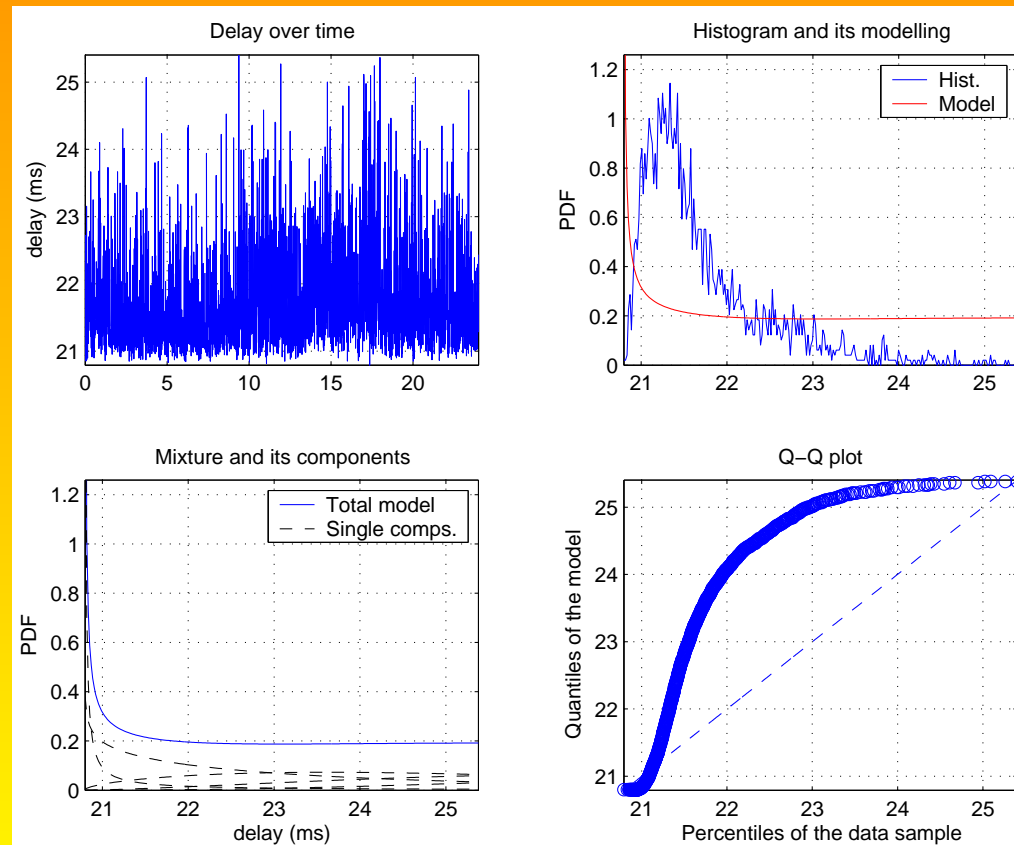
2. Updating hidden probs

$$p(y_i = j | x_i, \Theta) = \frac{\alpha_j p(x_i | \theta_j)}{\sum_{k=1}^M \alpha_k p(x_i | \theta_k)}$$

Table 1: EM procedure for mixtures of Weibull distributions.

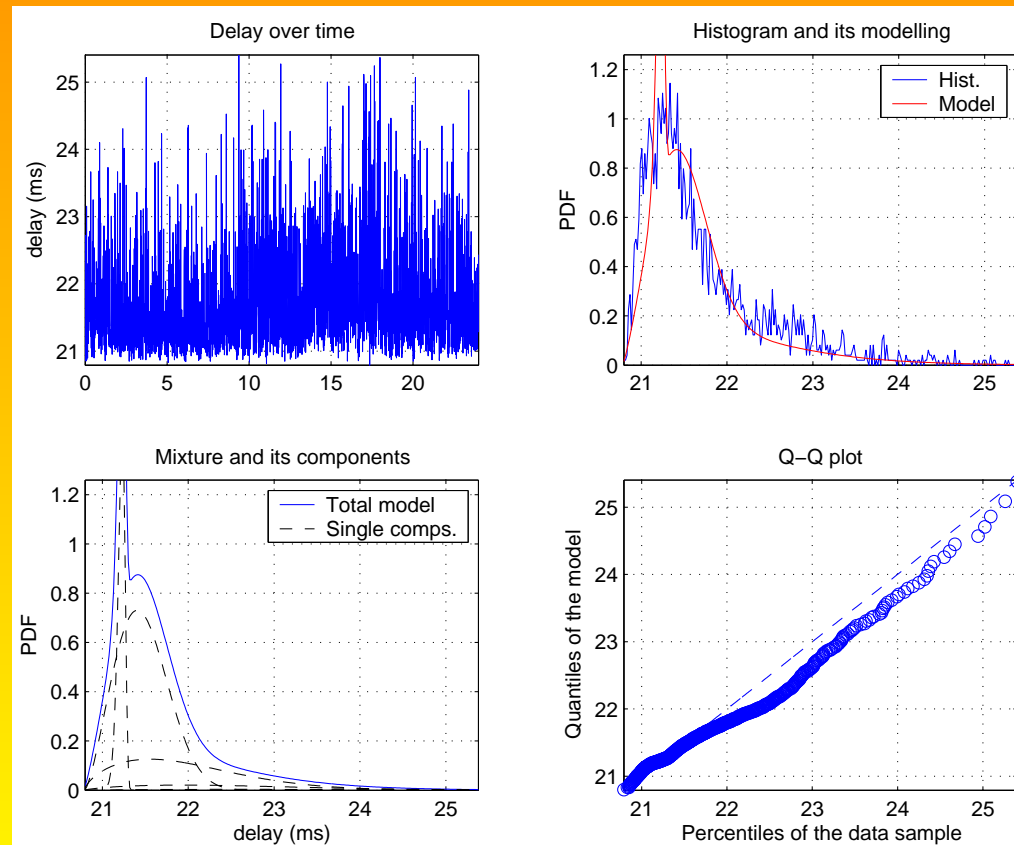
15 Mixing Weibull distributions

Convergence speed - Initialisation



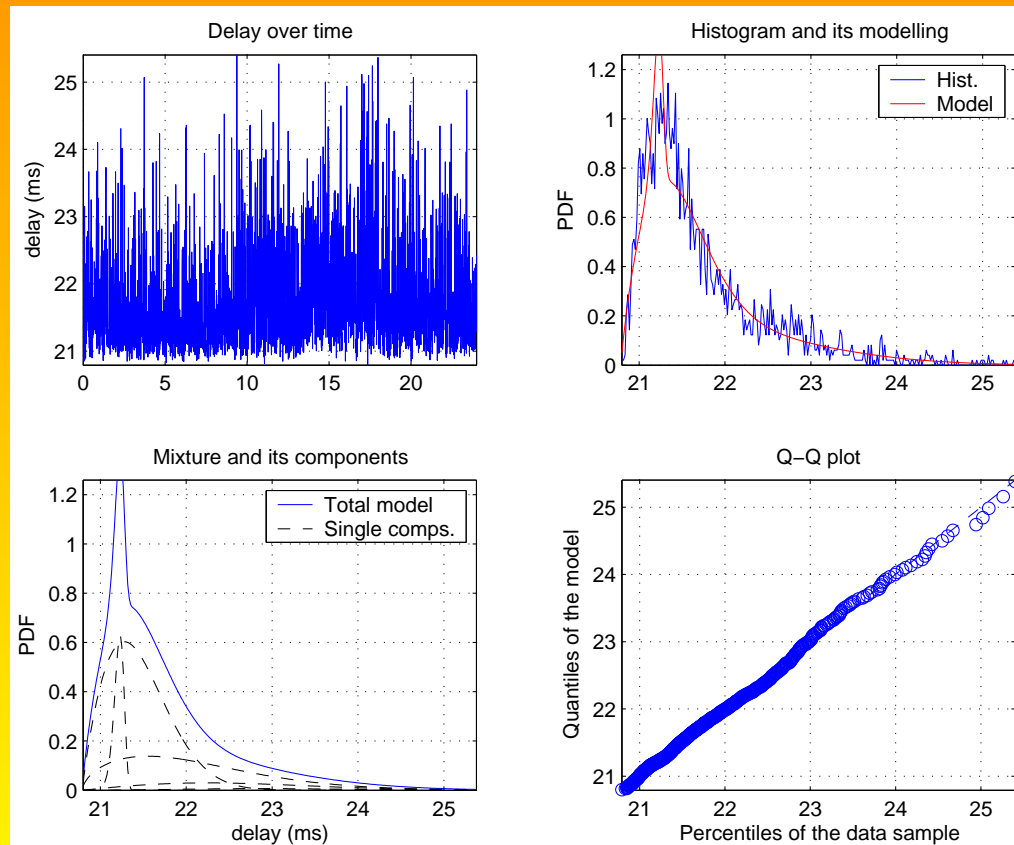
16 Mixing Weibull distributions

Convergence speed - After 1 iteration



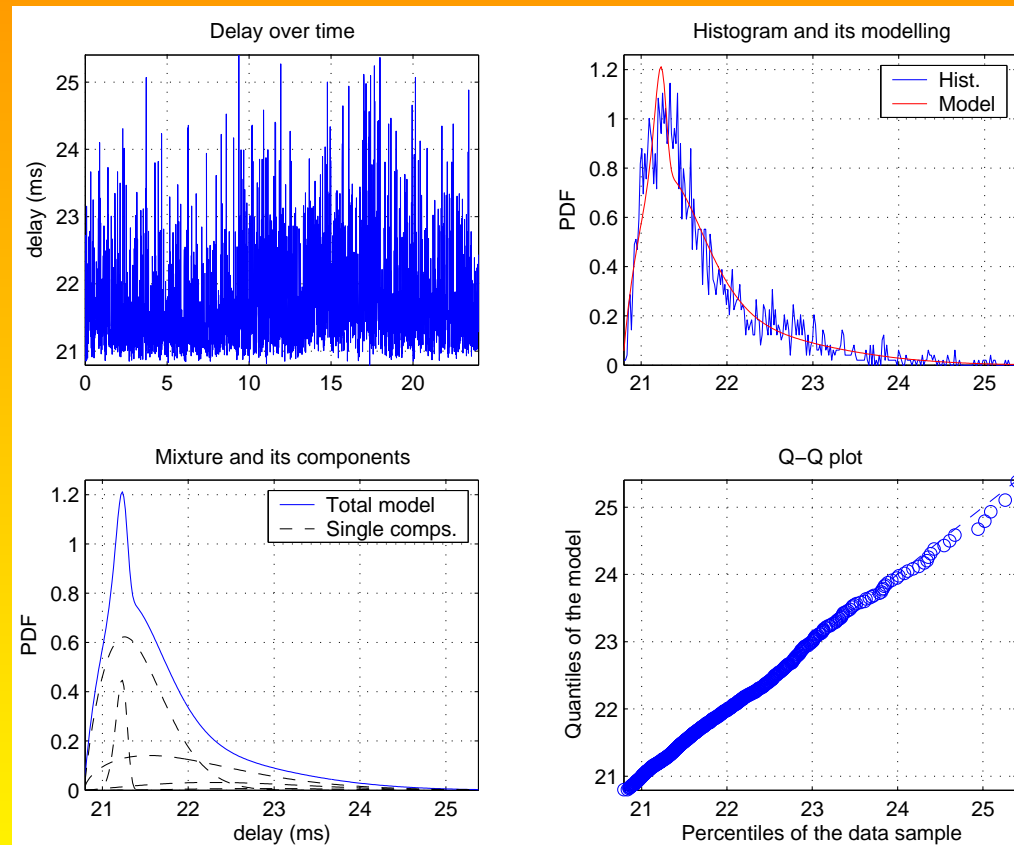
17 Mixing Weibull distributions

Convergence speed - After 2 iterations



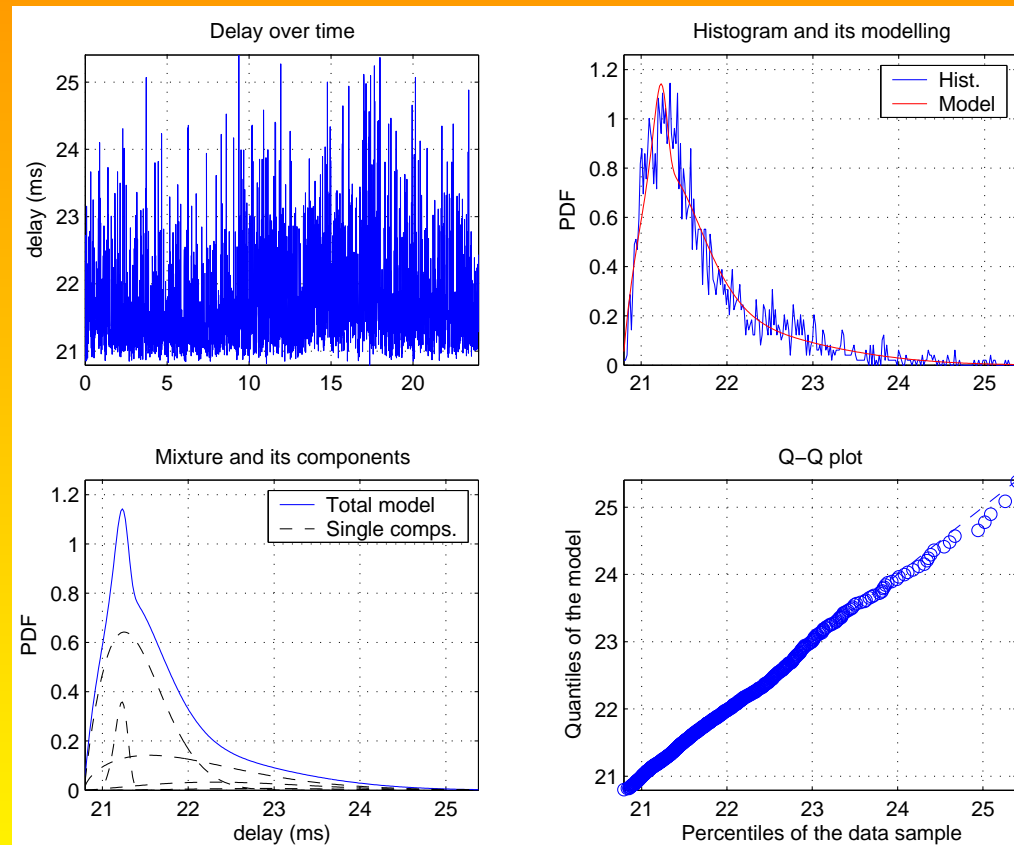
18 Mixing Weibull distributions

Convergence speed - After 3 iterations



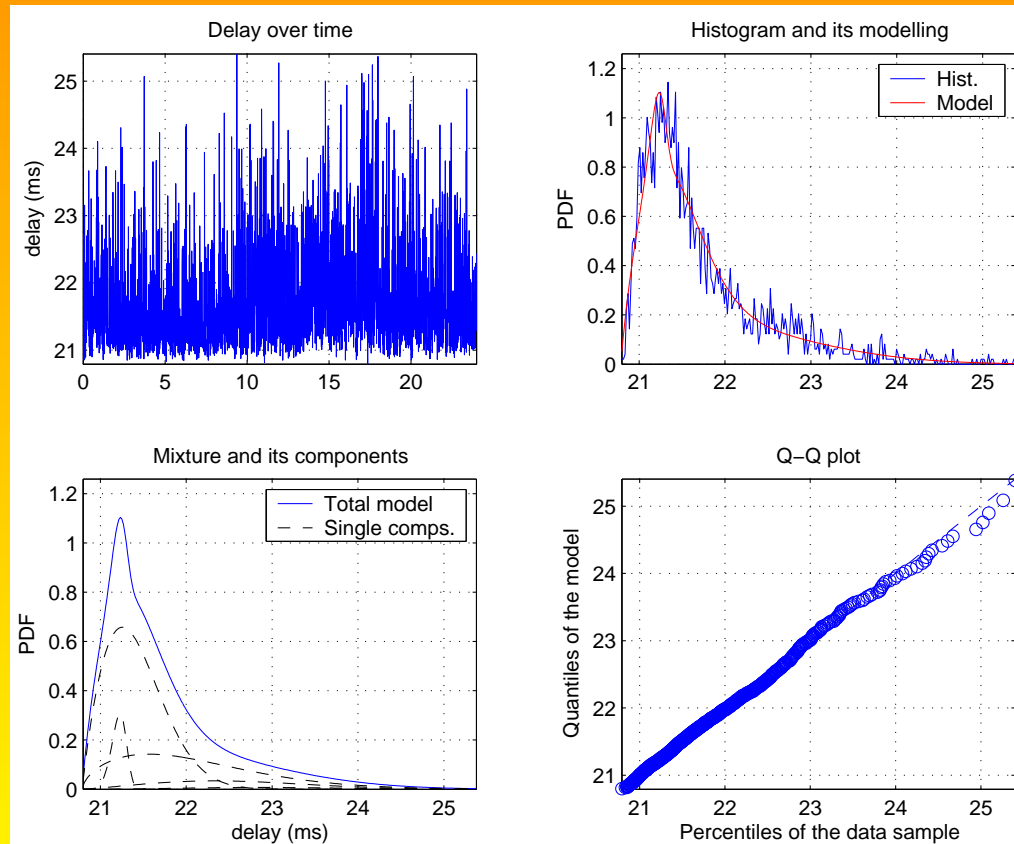
19 Mixing Weibull distributions

Convergence speed - After 4 iterations



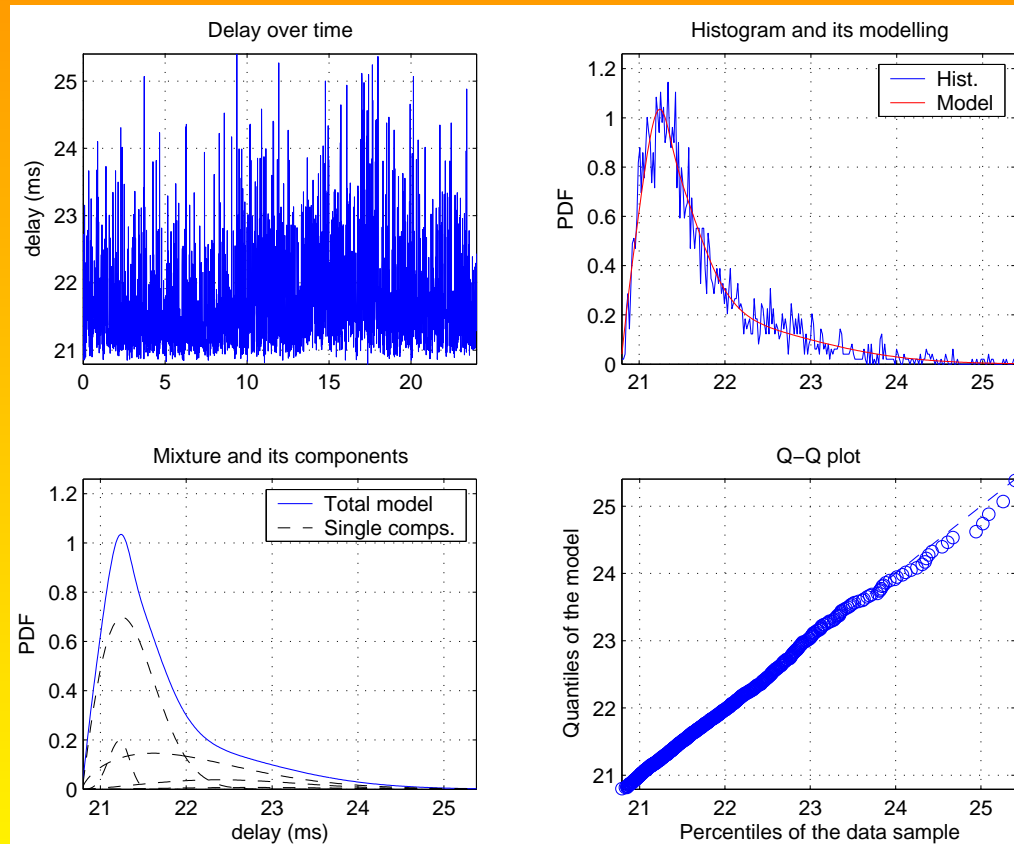
20 Mixing Weibull distributions

Convergence speed - After 5 iterations



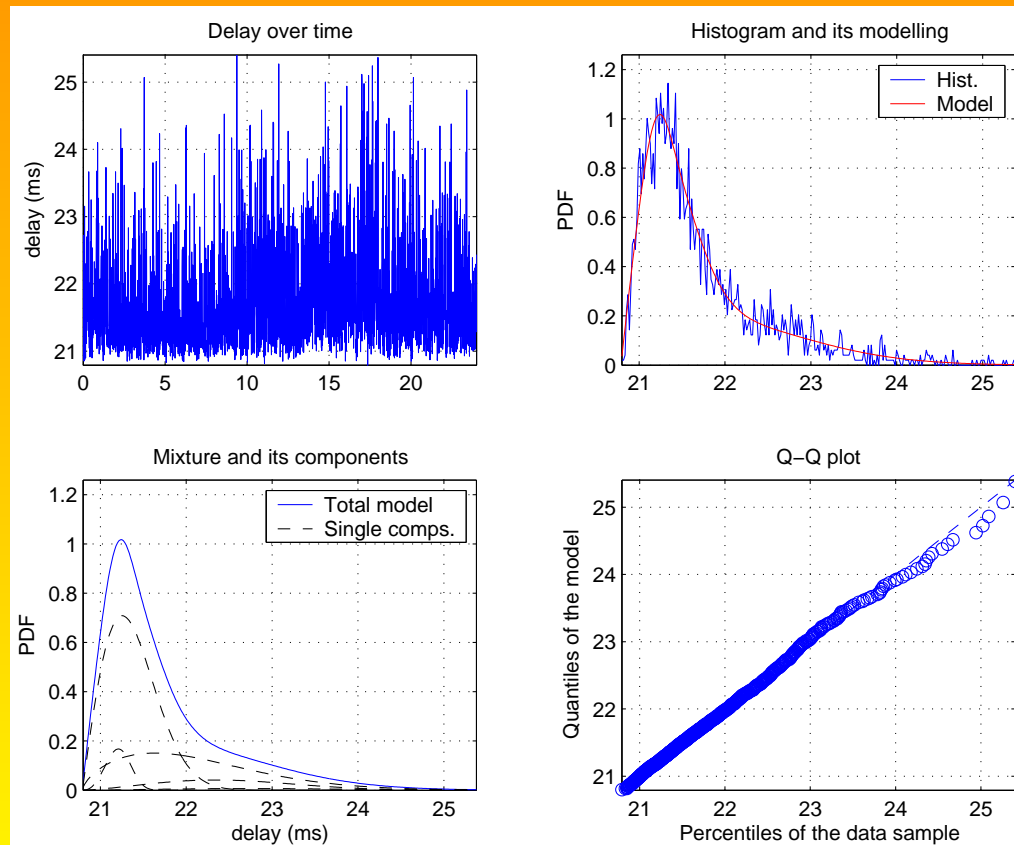
21 Mixing Weibull distributions

Convergence speed - After 10 iterations



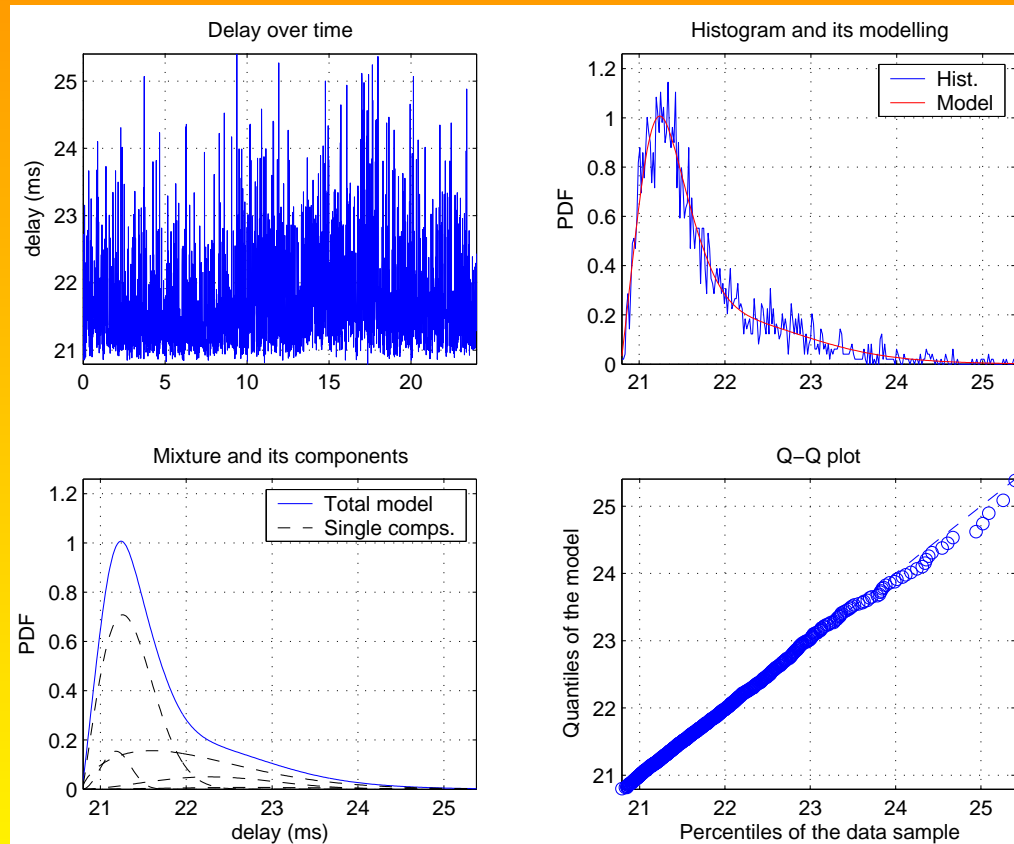
22 Mixing Weibull distributions

Convergence speed - After 15 iterations



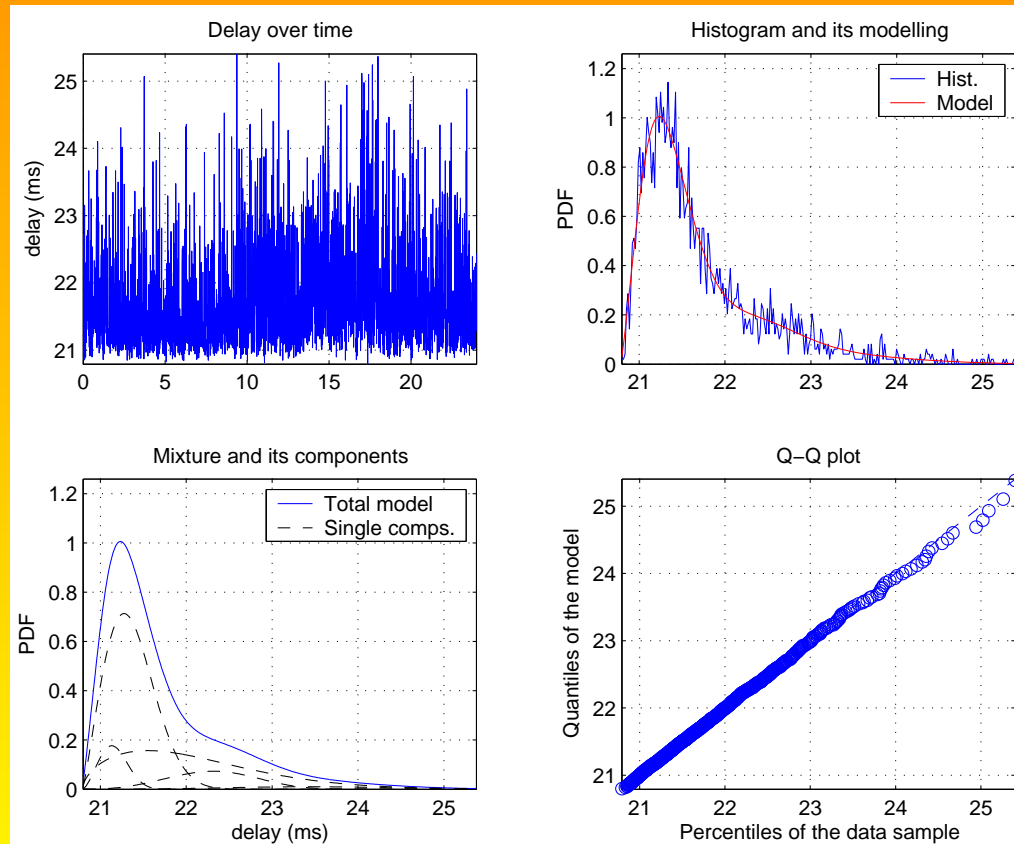
23 Mixing Weibull distributions

Convergence speed - After 25 iterations



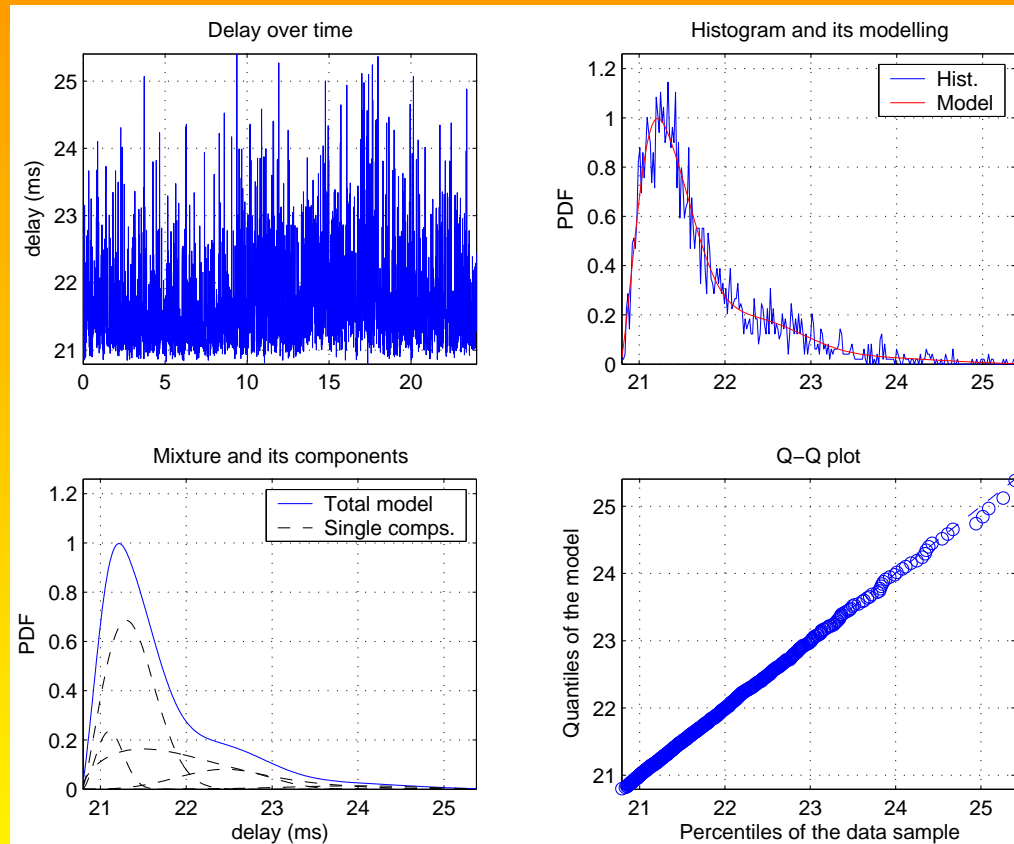
24 Mixing Weibull distributions

Convergence speed - After 50 iterations



25 Mixing Weibull distributions

Convergence speed - After 100 iterations



26 Experiments and results

Measurement testbed

- The following delay measurements, provided by RIPE NCC^a, have been utilised for this experiments.

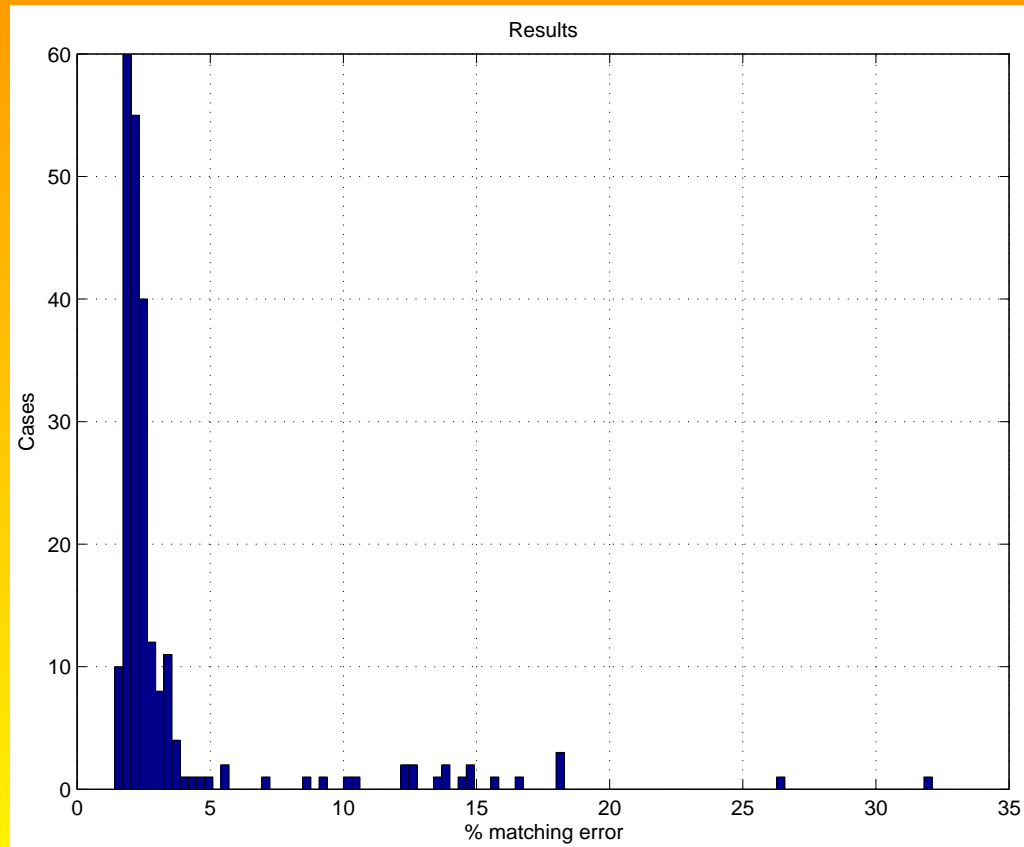
245 24-hour exps. } Total: $\approx 700,000$ measurements
 ≈ 3000 meas. per exp.

- GPS accuracy \approx few hundred of nanoseconds error.
- Matching error = $\frac{\sqrt{\sum(\text{hist-model})^2}}{\sum \text{hist}} \times 100\%$

^a<http://www.ripe.net>

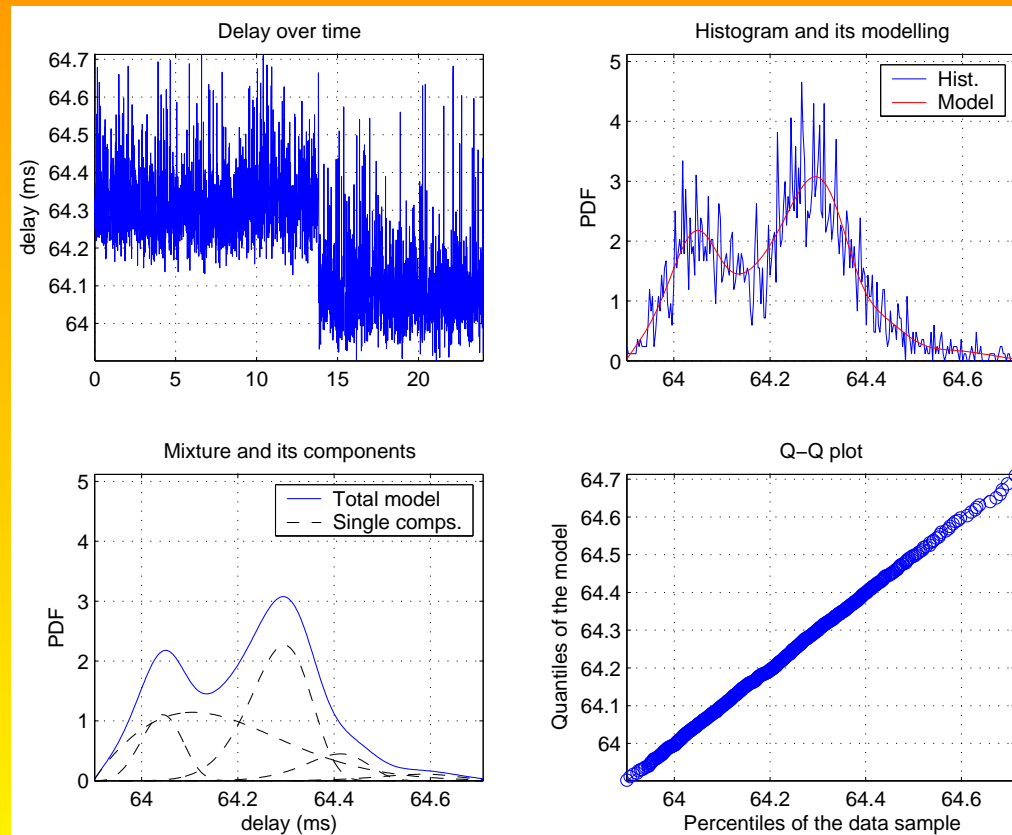
27 Experiments and results

Full experiments model validation



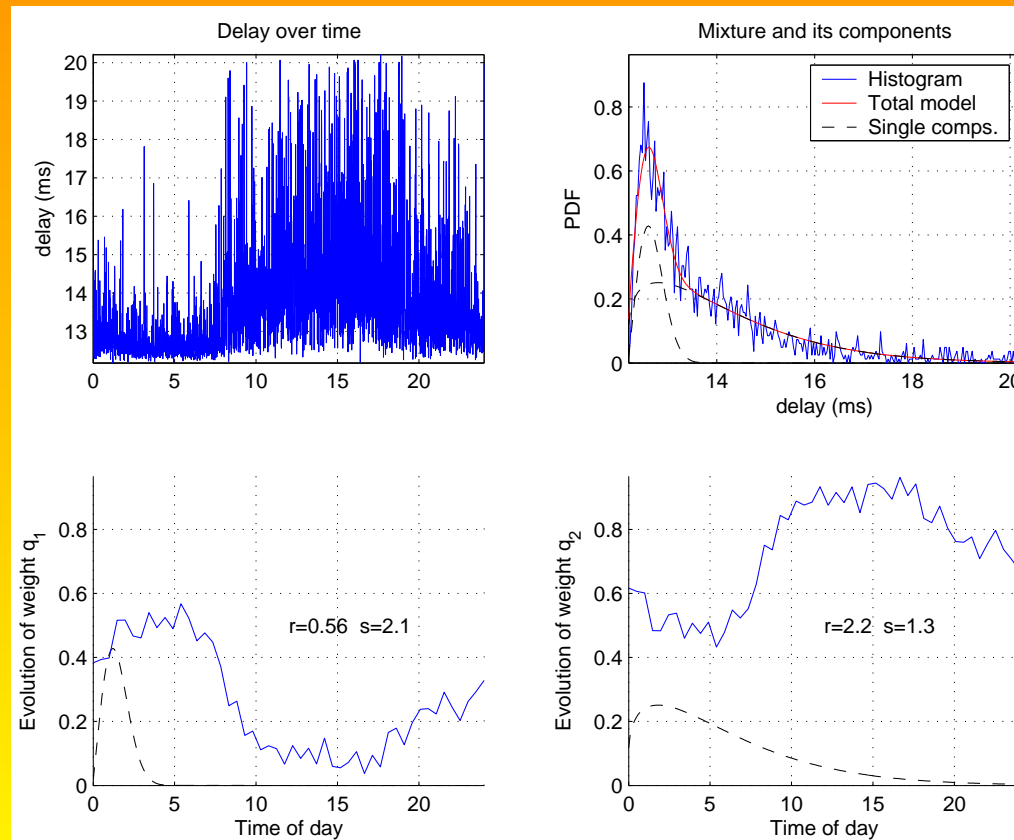
28 Experiments and results

Example of a five Weibull matching result



29 Experiments and results

Example of Parameter Evolution



30 Discussion

Two main conclusions arise from this work:

- A combination of Weibull distributions look very suitable to match end-to-end delay histograms.
- The Weibull parameters impact the appearance of the Weibull distribution.
 - r is related to the location of the mode/maximum/peak for that particular Weibull component.
 - s concerns tail behaviour: the smaller the slower the tail decays.
- Expectation Maximisation is a suitable algorithm to find the parameters defining such model, both easily and optimally.

31 Applications

Performance related applications:

- Traffic engineering.
- Fault tolerance and troubleshooting.
- Provisioning.
- Admission control.
- ...

32 Thanks

Other researchers at Loughborough especially: David Parish, Omar Bashir, Mark Sandford.

Jose's PhD is a Loughborough CS Scholarship.

RIPE for 1 year's (75GB) measurements.

Other URI partners.

Any questions ... ?