# Dynamics of Resource Sharing in Networks

Frank Kelly
**www.statslab.cam.ac.uk/~frank**

MITACS International Focus Period
Network and Internet Economics
Workshop, Vancouver, 1 June 2011

# Outline

- Fairness in networks
- Rate control in communication networks (relatively well understood)
- Ramp metering (early models)
- Energy networks (preliminary remarks)

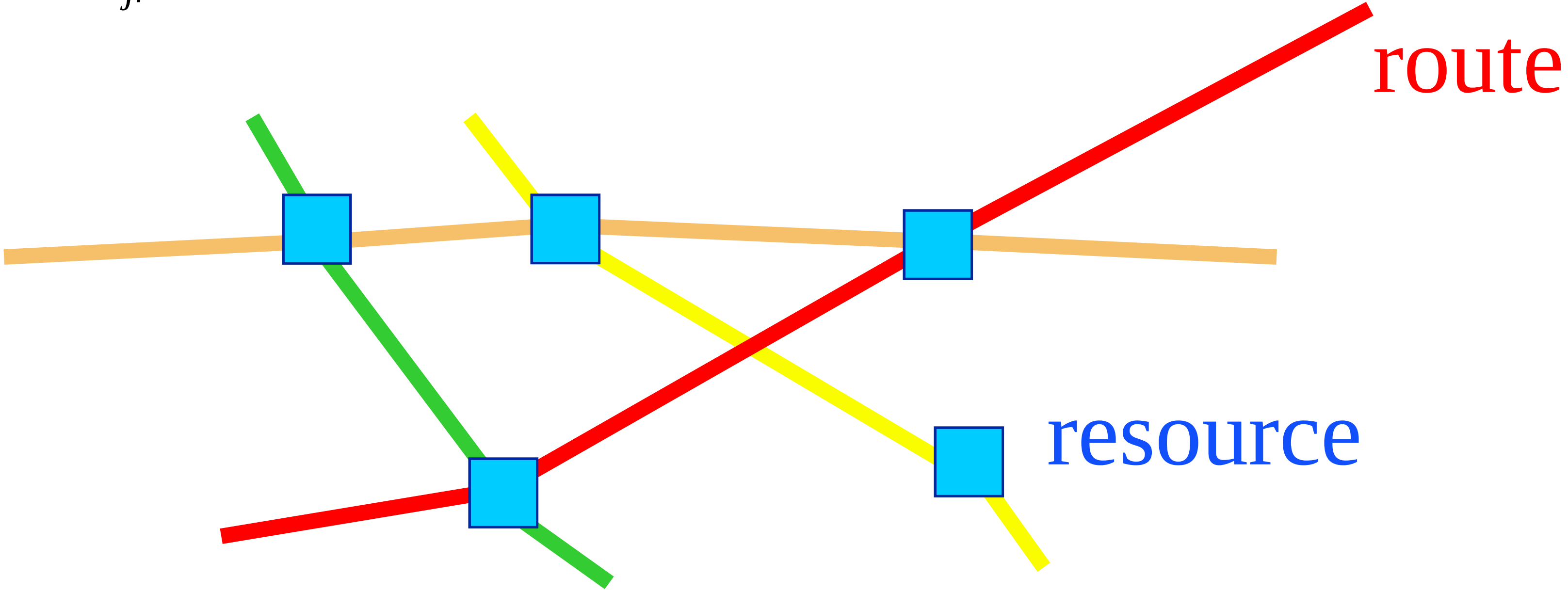
# Network structure

$J$ - set of resources

$R$ - set of routes

$A_{jr} = 1$ - if resource $j$ is on route $r$

$A_{jr} = 0$ - otherwise



# Notation

$J$ - set of resources

$R$ - set of users, or routes

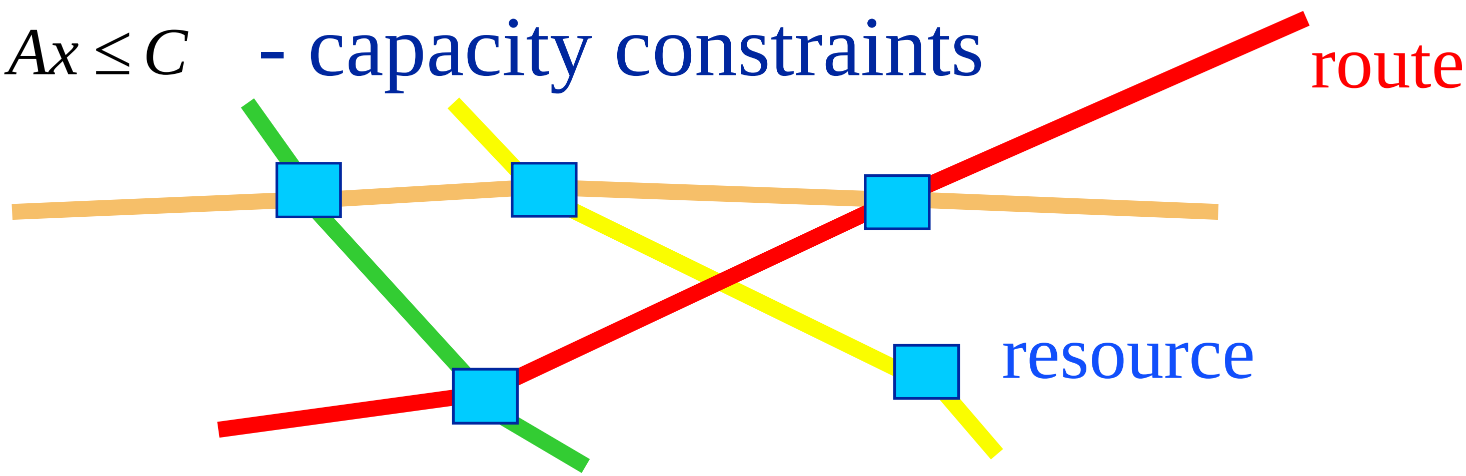
$j \in r$ - resource $j$ is on route $r$

$x_r$ - flow rate on route $r$

$U_r(x_r)$ - utility to user $r$

$C_j$ - capacity of resource $j$

$Ax \le C$ - capacity constraints



# The system problem

**SYSTEM(*U*,*A*,*C*):**

$$\begin{aligned} &\textit{Maximize} && \sum_{r \in R} U_r(x_r) \\ &\textit{subject to} && Ax \le C \\ &\textit{over} && x \ge 0 \end{aligned}$$

Maximize aggregate utility,
subject to capacity constraints

# The user problem

**USER$_r$($U_r$;$\lambda_r$):** $$\textit{Maximize } U_r\left(\frac{w_r}{\lambda_r}\right) - w_r$$

$$\textit{over} \qquad w_r \geq 0$$

User $r$ chooses
an amount to pay per unit time, $w_r$,
and receives in return a flow $x_r = w_r / \lambda_r$

# The network problem

**NETWORK(*A*,*C*;*w*):**

$$\begin{aligned} &\text{Maximize} && \sum_{r \in R} w_r \log x_r \\ &\text{subject to} && Ax \le C \\ &\text{over} && x \ge 0 \end{aligned}$$

*As if* the network maximizes a logarithmic utility function, but with constants $\{w_r\}$ chosen by the users

# Problem decomposition

Theorem: the system problem
may be solved
by solving simultaneously
the network problem and
the user problems

K 1997,
Johari, Tsitsiklis 2005,
Yang, Hajek 2006

# Max-min fairness

Rates $\{x_r\}$ are *max-min fair* if they are feasible:

$$x \geq 0, \qquad Ax \leq C$$

and if, for any other feasible rates $\{y_r\}$,

$$\exists r : y_r > x_r \;\Rightarrow\; \exists s : y_s < x_s < x_r$$

Rawls 1971,
Bertsekas, Gallager 1987

# Proportional fairness

Rates $\{x_r\}$ are *proportionally fair* if they are feasible:

$$x \geq 0,\ Ax \leq C$$

and if, for any other feasible rates $\{y_r\}$, the aggregate of proportional changes is negative:

$$\sum_{r \in R} \frac{y_r - x_r}{x_r} \leq 0$$

# Weighted proportional fairness

A feasible set of rates $\{x_r\}$ are such that are *weighted proportionally fair* if, for any other feasible rates $\{y_r\}$,

$$\sum_{r \in R} w_r \frac{y_r - x_r}{x_r} \leq 0$$

# Fairness and the network problem

Theorem: a set of rates $\{x_r\}$
solves the network problem,
**NETWORK(*A*,*C*;*w*),**
if and only if the rates are
weighted proportionally fair

# Bargaining problem (Nash, 1950)

Solution to **NETWORK(*A*,*C*;*w*)** with ***w* = *1*** is unique point satisfying

- Pareto efficiency
- Symmetry
- Independence of Irrelevant Alternatives

(General ***w*** corresponds to a model with unequal bargaining power)

# Market clearing equilibrium (Gale, 1960)

Find prices ***p*** and an allocation ***x*** such that

$$p \geq 0, \quad Ax \leq C$$ (feasibility)

$$p^T(C - Ax) = 0$$ (complementary slackness)

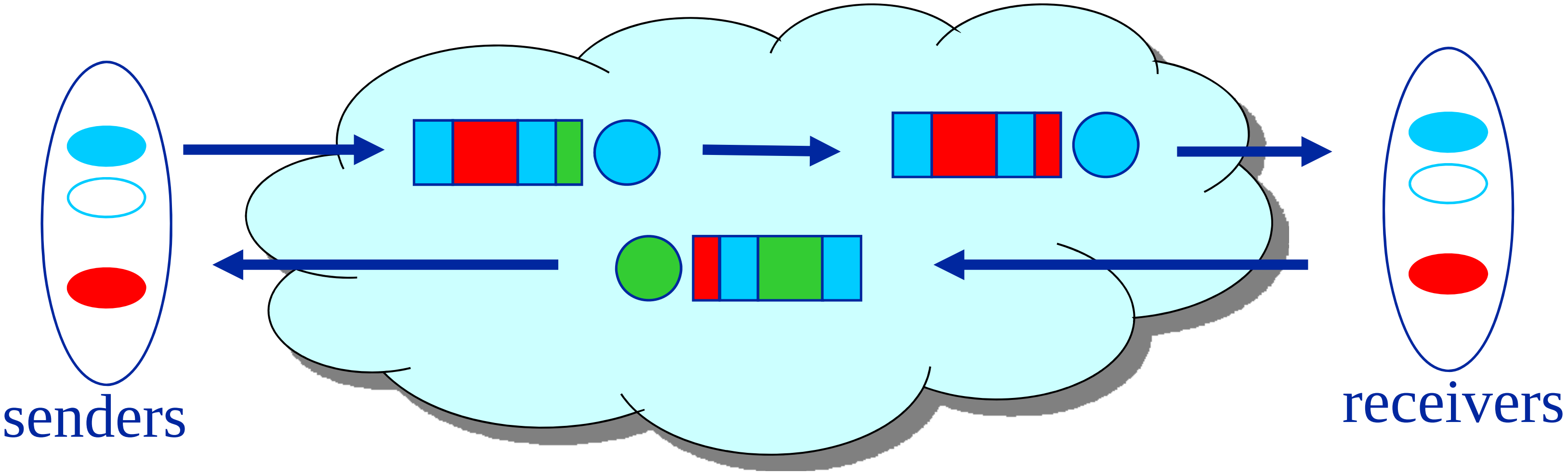
$$w_r = x_r \sum_{j \in r} p_j, \quad r \in R$$ (endowments spent)

Solution solves **NETWORK(*A*,*C*;*w*)**

# Outline

- Fairness in networks
- Rate control in communication networks (relatively well understood)
- Ramp metering (early models)
- Energy networks (preliminary remarks)

# End-to-end congestion control



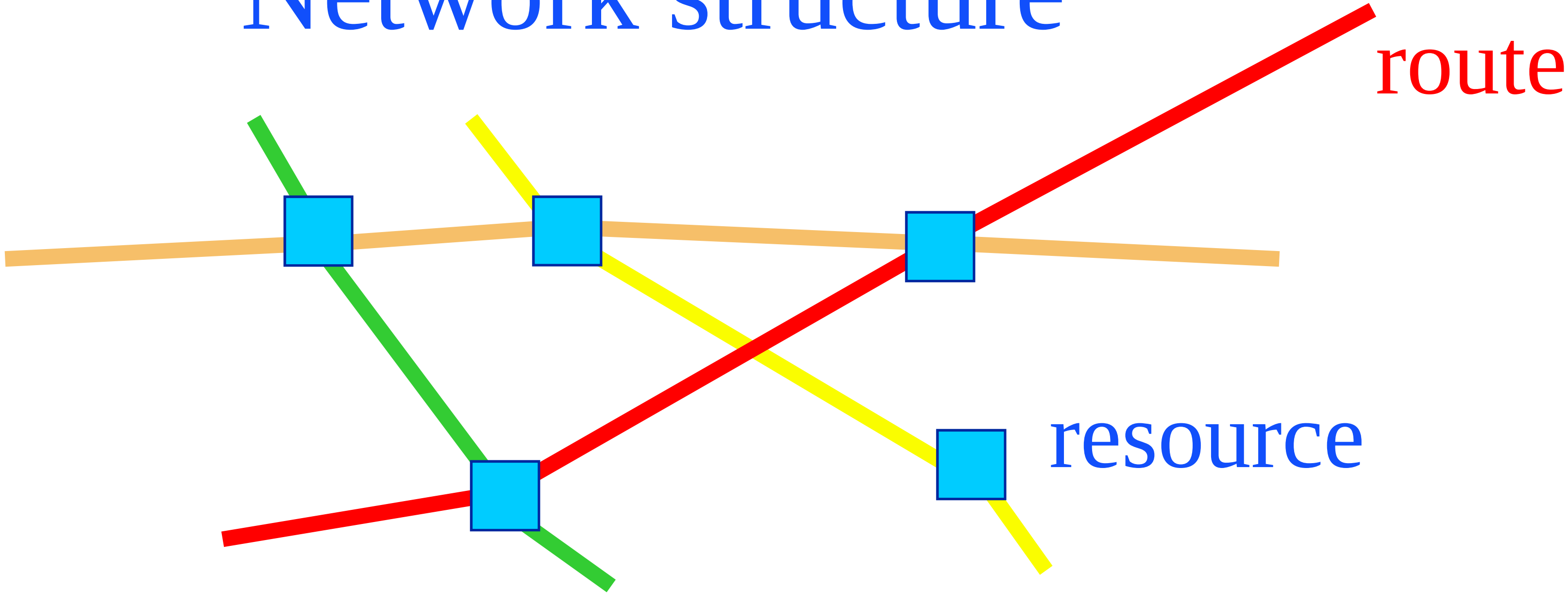

Senders learn (through feedback from receivers) of congestion at queue, and slow down or speed up accordingly. With current TCP, throughput of a flow is proportional to

$$\boxed{1/(T\sqrt{p})}$$

$T$ = round-trip time, $p$ = packet drop probability.

(Jacobson 1988, Mathis, Semke, Mahdavi, Ott 1997, Padhye, Firoiu, Towsley, Kurose 1998, Floyd & Fall 1999)

# Network structure




| | |
|---|---|
| $J$ | - set of resources |
| $R$ | - set of routes |
| $j \in r$ | - resource $j$ is on route $r$ |
| $x_r(t)$ | - flow rate on route $r$ at time $t$ |
| $\mu_j(t)$ | - rate of congestion indication, at resource $j$ at time $t$ |

# A primal algorithm

$$\frac{d}{dt}x_r(t) = \kappa_r(x_r(t))\left(w_r - x_r(t)\sum_{j\in r}\mu_j(t)\right)$$

$$\mu_j(t) = p_j\left(\sum_{s:j\in s}x_s(t)\right)$$

$x_r(t)$ - rate changes by linear increase, multiplicative decrease

$p_j(.)$ - proportion of packets marked as a function of flow through resource

# Global stability

Theorem: the above dynamical system has a stable point to which all trajectories converge. The stable point is proportionally fair with respect to the weights $\{w_r\}$, and solves the network problem, when

$$p_j(x) = 0 \qquad x \le C_j$$
$$\qquad = \infty \qquad x > C_j$$

K, Maulloo, Tan 1998

# General TCP-like algorithm

Source maintains window of sent, but not yet acknowledged, packets - size *cwnd*

$$cwnd \approx xT$$

On route $r$,

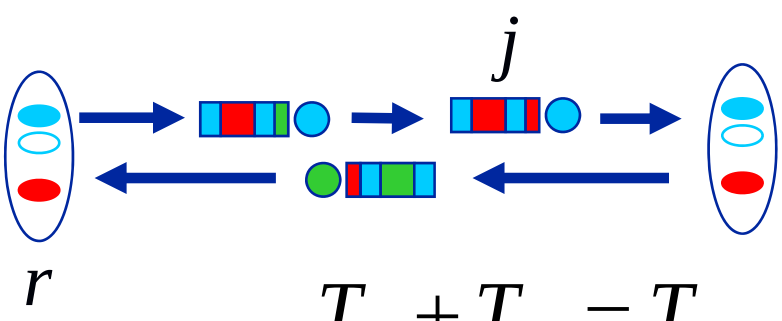
- $cwnd$ incremented by $a_r \, cwnd^{\,n}$ on positive acknowledgement
- $cwnd$ decremented by $b_r \, cwnd^{\,m}$ for each congestion indication $(m>n)$
- $a_r = 1$, $b_r = 1/2$, $m=1$, $n= -1$ corresponds to Jacobson's TCP

# Differential equations with delays

$$\frac{d}{dt} x_r(t) = \frac{x_r(t - T_r)}{T_r} \cdot \Big( a_r \, (x_r(t) T_r)^n \, (1 - \lambda_r(t)) - b_r \, (x_r(t) T_r)^m \, \lambda_r(t) \Big)$$

$$\lambda_r(t) = 1 - \prod_{j \in r} \big( 1 - \mu_j(t - T_{jr}) \big)$$

$$\mu_j(t) = p_j \left( \sum_{r: j \in r} x_r(t - T_{rj}) \right)$$




$$T_{rj} + T_{jr} = T_r$$

# Equilibrium point

$$x_r = \frac{1}{T_r}\left(\frac{a_r}{b_r}\frac{1-\lambda_r}{\lambda_r}\right)^{1/m-n} \qquad r \in R$$

- $a_r = 1$, $b_r = 1/2$, $m=1$, $n= -1$ corresponds to Jacobson's TCP, and recovers square root formula

- But what is the impact of delays on stability? Can we choose *m*, *n*,… arbitrarily?

# Delay stability

Johari, Tan 1999,
Massoulié 2000,
Vinnicombe 2000,
Paganini, Doyle, Low 2001

Equilibrium is locally stable if there exists a global constant $\beta$ such that

$$x p_j'(x) < \beta \; p_j(x), \qquad a_r (x_r T_r)^n < \frac{\pi}{2\beta}$$

condition on sensitivity for each resource $j$

condition on aggressiveness for each route $r$

# Outline

- Fairness in networks
- Rate control in communication networks (relatively well understood)
- Ramp metering (early models)
- Energy networks (preliminary remarks)

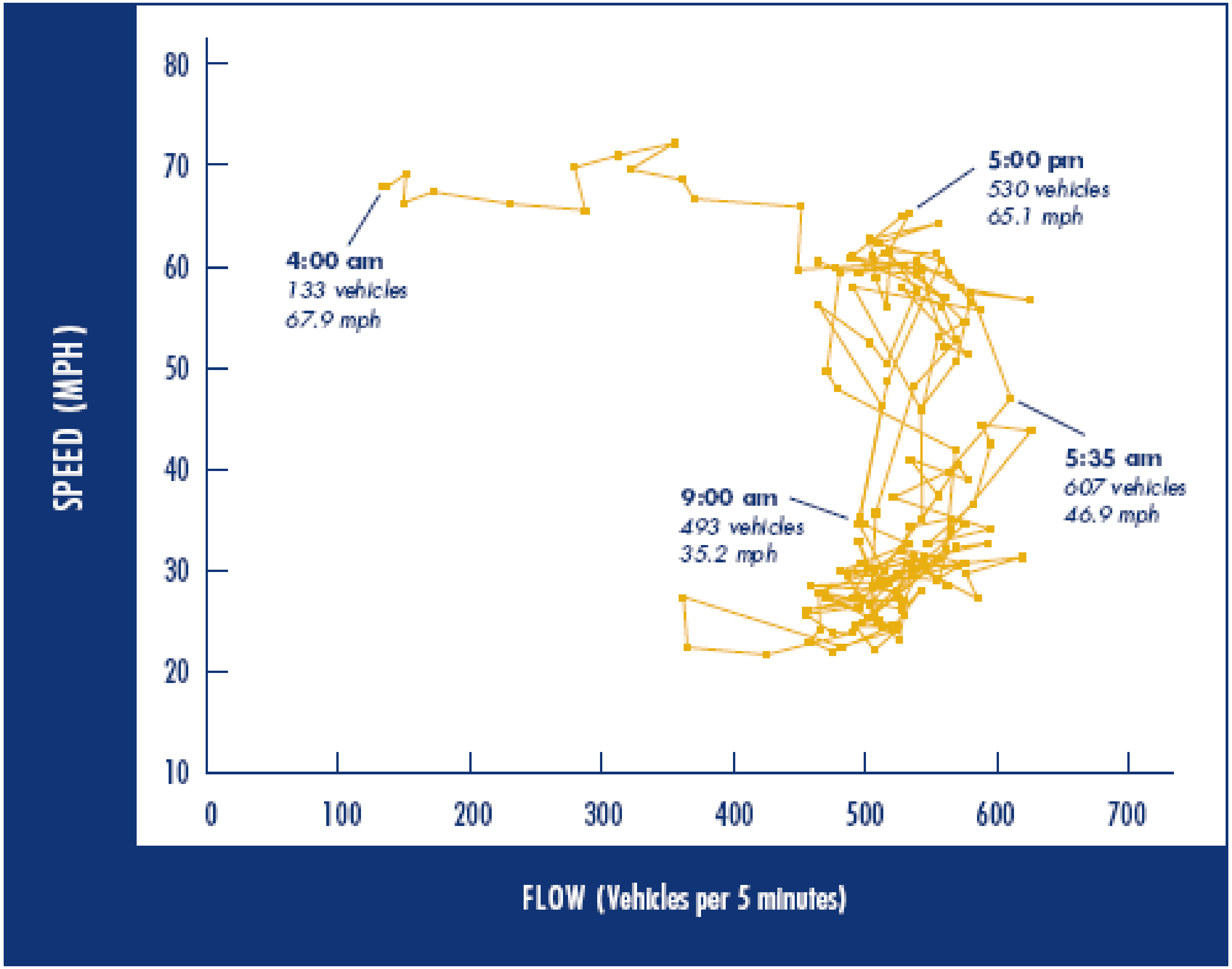



**FIGURE 1**

Speed vs. flow on I-10 westbound in 5 minute intervals from 4:00 am to 6:00 pm

[What we've learned about highway congestion]

*P. Varaiya*, Access 27, Fall 2005, 2-9.

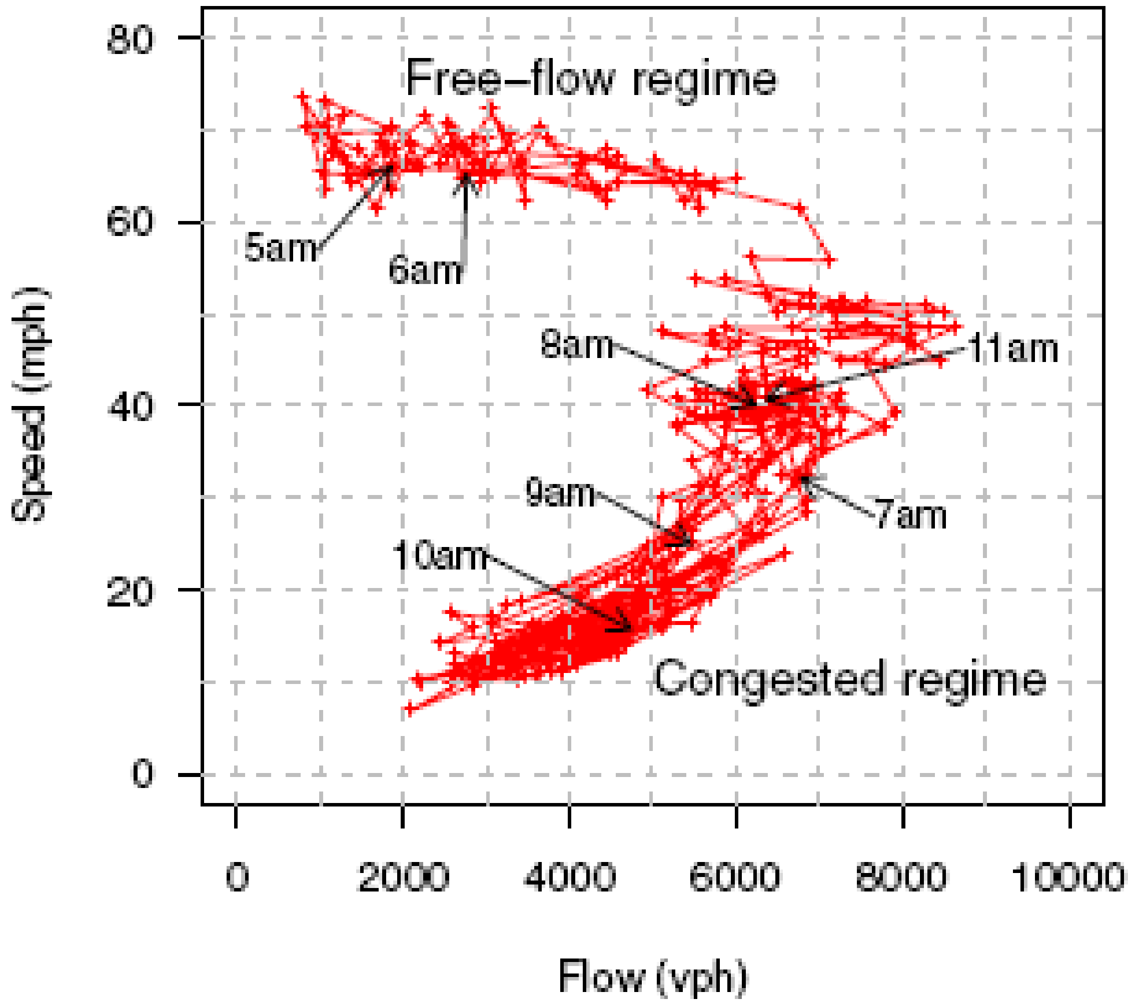
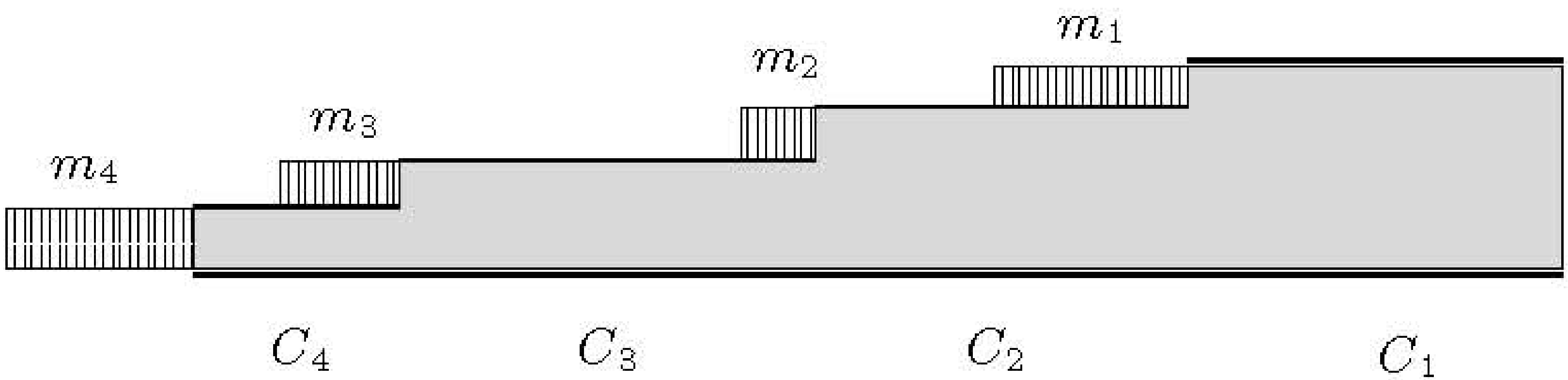



Figure 2. The relationship between the speed and flow of vehicles observed on the morning of Wednesday, 14 July 2004 using the M25 midway between junctions 11 and 12 in the clockwise direction. In the free-flow regime, flow rapidly increases with only a modest decline in speeds. Above a critical occupancy of vehicles there is a marked drop in speed with little, if any, improvement in flow which is then followed by a severe collapse into a congested regime where both flow and speed are highly variable and attain very low levels. Finally, the situation recovers with a return to higher flows and an improvement in speeds

Data, modelling and inference in road traffic networks
*R.J. Gibbens and Y. Saatci*
Phil. Trans. R. Soc. A366 (2008), 1907-1919.

# A linear network




$$m_i(t) = m_i(0) + e_i(t) - \int_0^t \Lambda_i(m(s))\mathrm{d}s, \quad t \geq 0$$

queue size

cumulative inflow

metering rate

# Metering policy

Suppose the metering rates can be chosen to be any vector $\Lambda = \Lambda(m)$ satisfying

$$\sum_i A_{ji}\,\Lambda_i \le C_j, \quad j \in J$$

$$\Lambda_i \ge 0, \quad i \in I$$

$$\Lambda_i = 0, \quad m_i = 0$$

and such that

$$m_i(t) = m_i(0) + e_i(t) - \int_0^t \Lambda_i(m(s))\mathrm{d}s \ge 0, \quad t \ge 0$$

# Optimal policy?

For each of $i = I, I\text{-}1, \ldots\ldots 1$ in turn choose

$$\int_0^t \Lambda_i(m(s))\mathrm{d}s \geq 0$$

to be maximal, subject to the constraints.

This policy minimizes

$$\sum_i m_i(t)$$

for all times $t$

# Proportionally fair metering

Suppose $\Lambda(m) = (\Lambda_i(m), i \in I)$ is chosen to

maximize $$\sum_i m_i \log \Lambda_i$$

subject to $$\sum_i A_{ji} \Lambda_i \leq C_j, \quad j \in J$$

$$\Lambda_i \geq 0, \quad i \in I$$

$$\Lambda_i = 0, \quad m_i = 0$$

# Proportionally fair metering

$$\Lambda_i(m) = \frac{m_i}{\sum_j p_j A_{ji}}, \quad i \in I$$

where

$$\boxed{\begin{aligned}
&\Lambda_i \ge 0, \quad i \in I \\
&\sum_i A_{ji}\,\Lambda_i \le C_j, \quad j \in J \\
&p_j \ge 0, \quad j \in J \\
&p_j\left(C_j - \sum_i A_{ji}\,\Lambda_i\right) \ge 0, \quad j \in J
\end{aligned}}$$

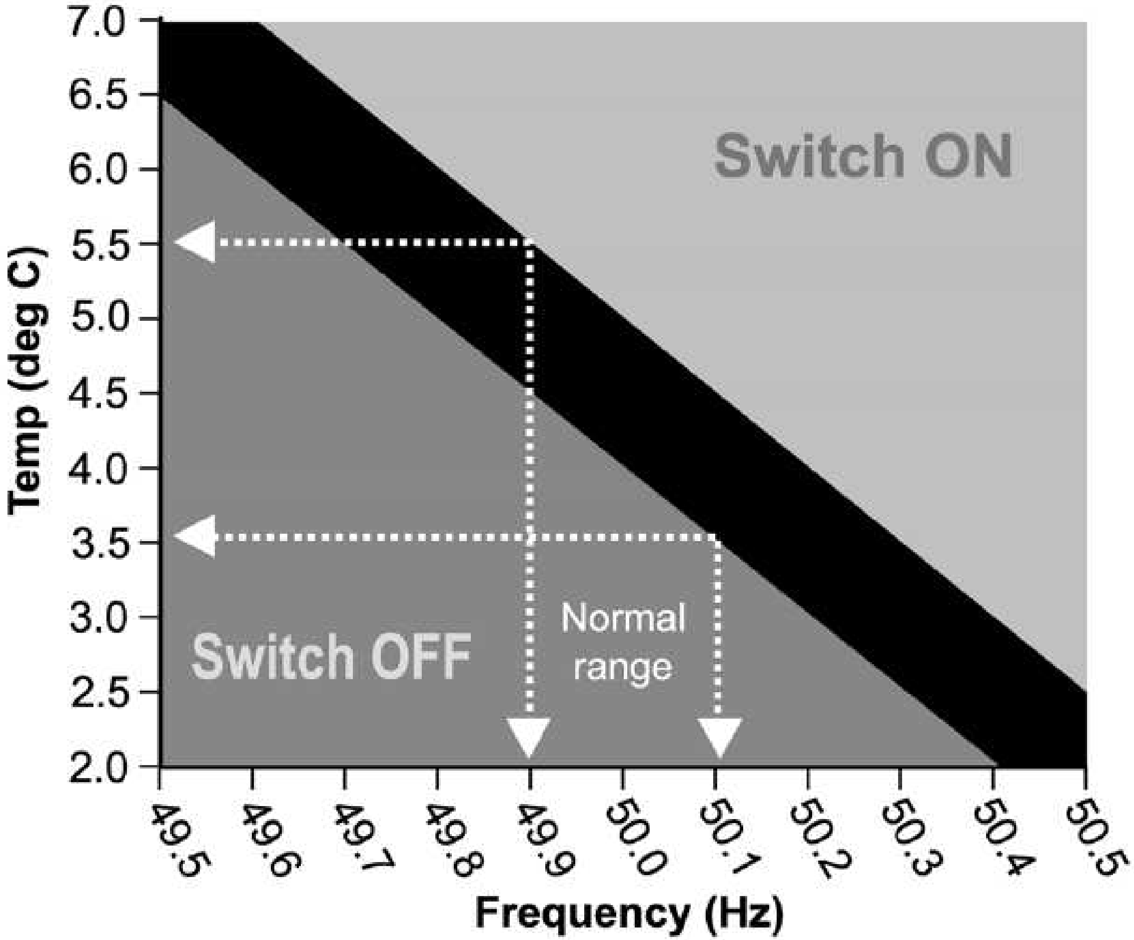
KKT conditions

$p_j$ - *shadow price* (Lagrange multiplier) for the resource $j$ capacity constraint

# Outline

- Fairness in networks
- Rate control in communication networks (relatively well understood)
- Ramp metering (early models)
- Energy networks (preliminary remarks)

# Dynamic demand




Use system frequency as a signal to control domestic loads, particularly refrigerators and freezers, to provide operating reserve

From: Frequency responsive loads, Jeremy Colandairaj, NIE

# Distribution of frequency

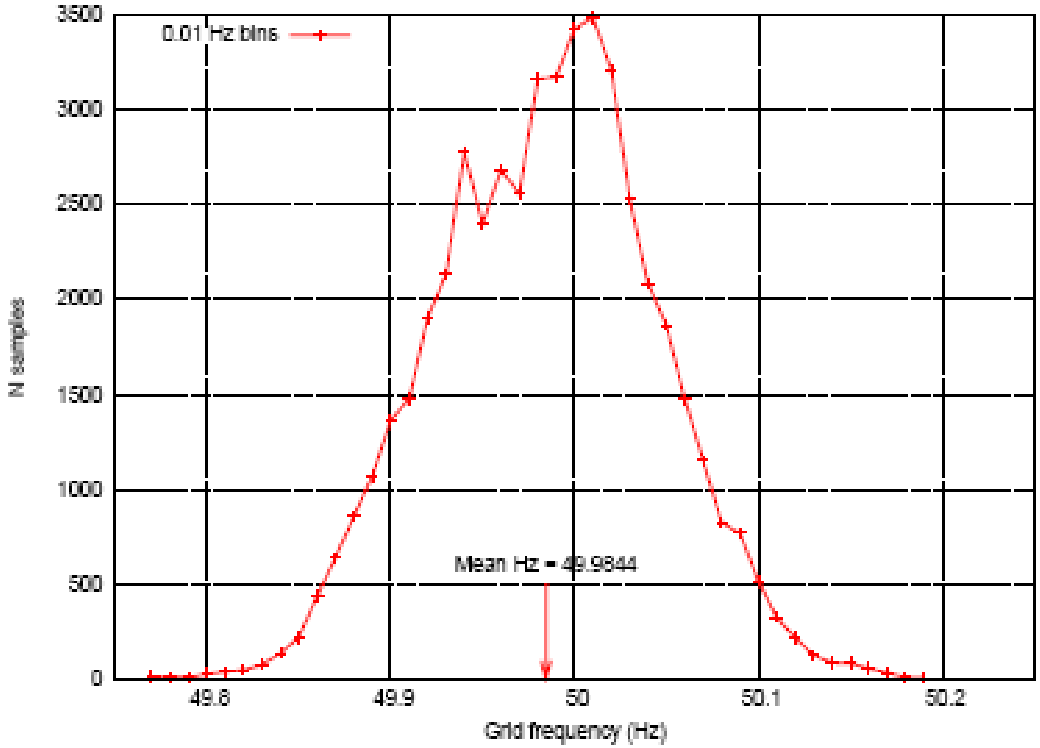
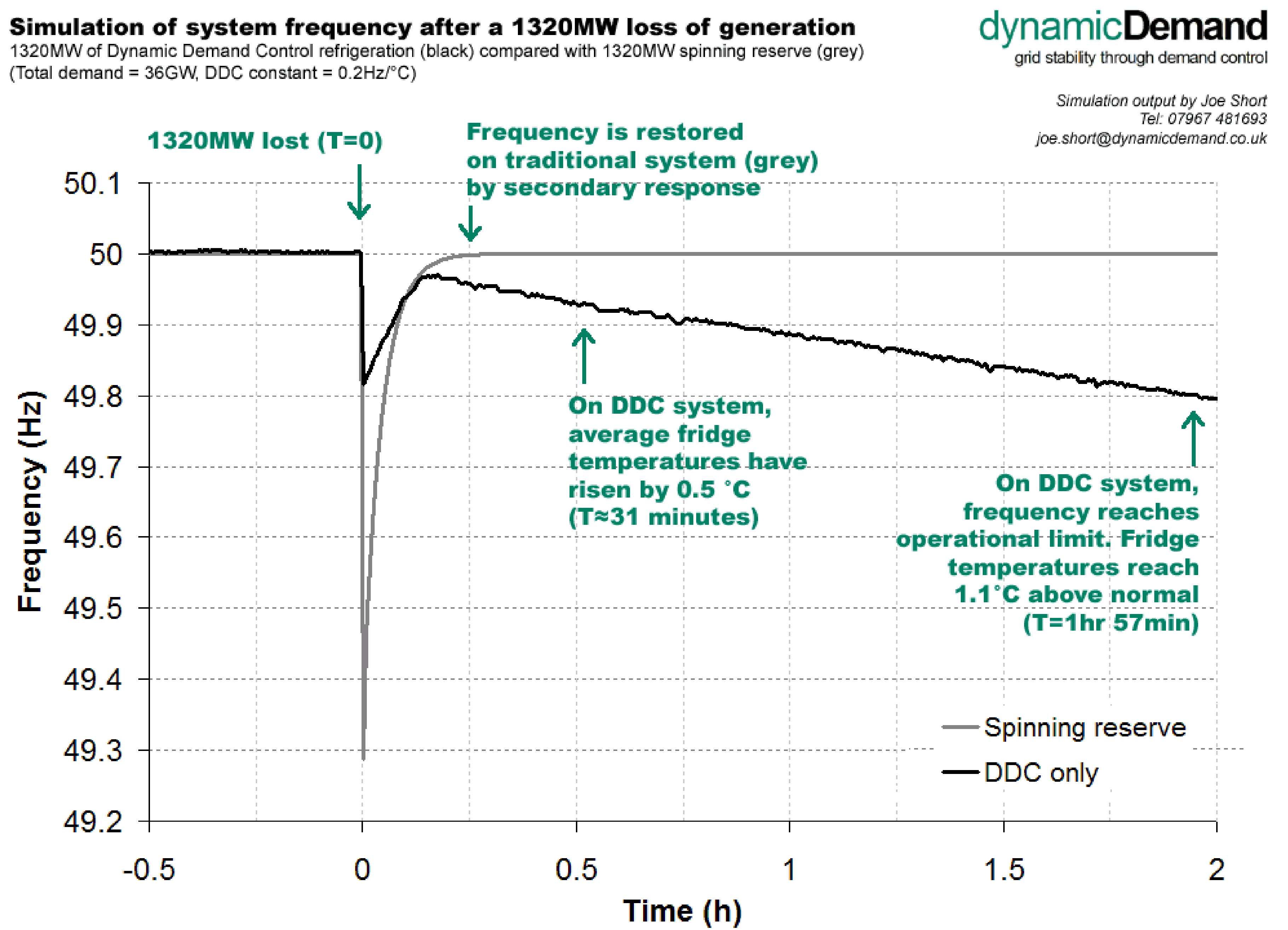


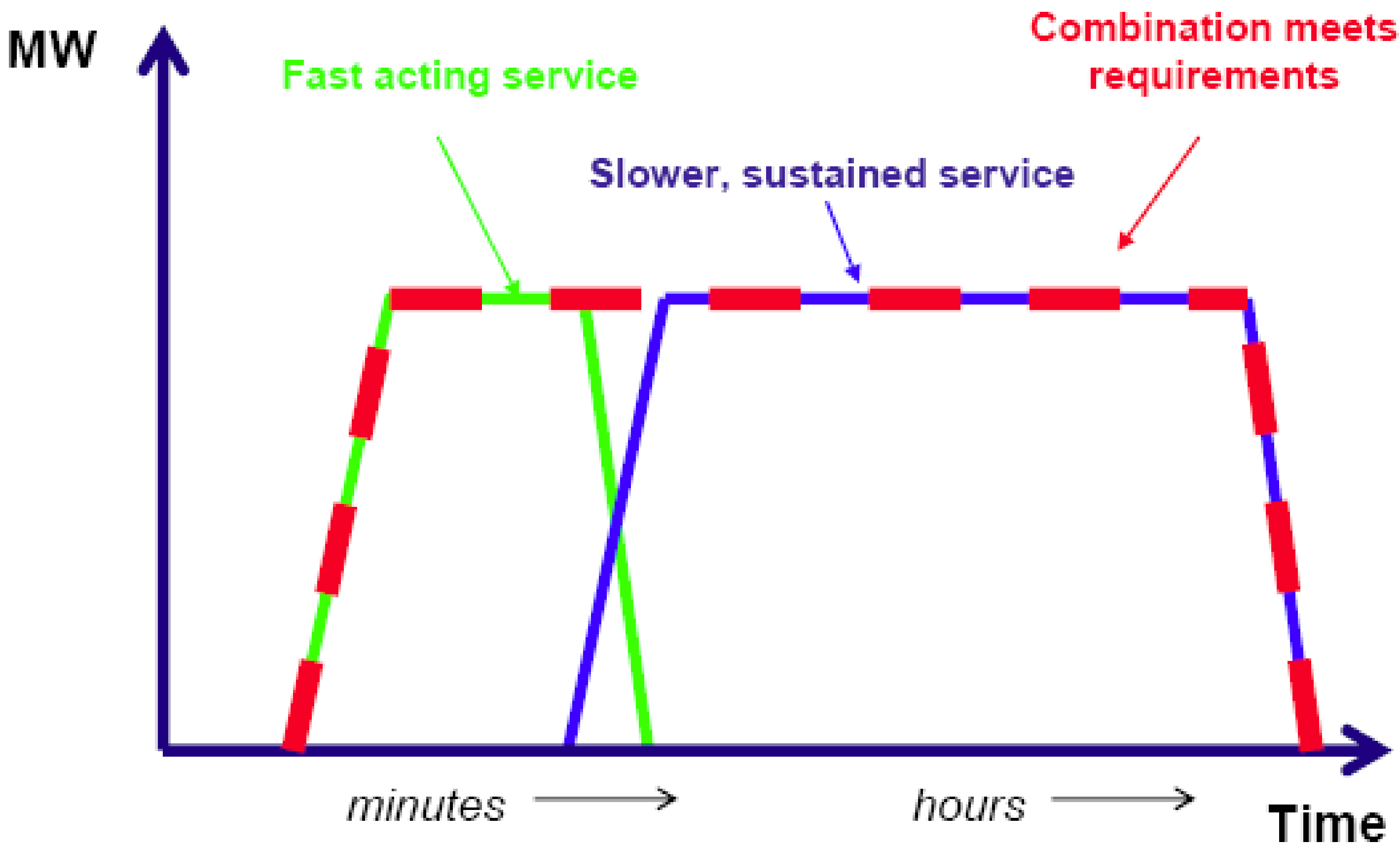

**Figure 5:** Distribution of grid frequency for a 30 hour period starting 31/08/2005 18:54:00. The mean frequency was slightly lower than nominal.

From: www.dynamicDemand.co.uk (Dynamic Demand is a not-for-profit organisation set up by a grant from the Esmée Fairbairn Foundation)



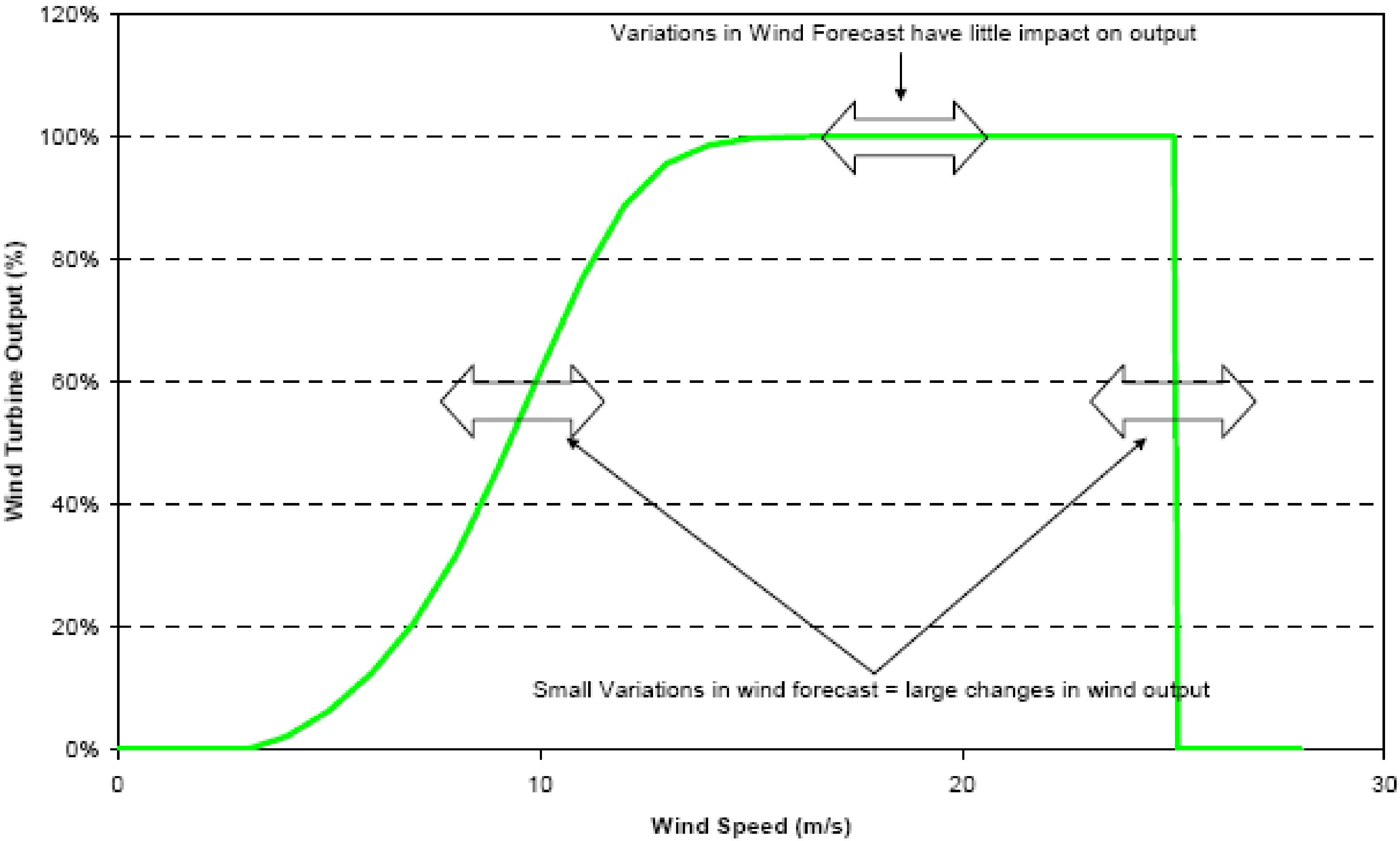
Simulation of system frequency after a 1320MW loss of generation
1320MW of Dynamic Demand Control refrigeration (black) compared with 1320MW spinning reserve (grey)
(Total demand = 36GW, DDC constant = 0.2Hz/°C)
dynamicDemand
grid stability through demand control
Simulation output by Joe Short
Tel: 07967 481693
joe.short@dynamicdemand.co.uk
1320MW lost (T=0)
Frequency is restored on traditional system (grey) by secondary response
On DDC system, average fridge temperatures have risen by 0.5 °C (T≈31 minutes)
On DDC system, frequency reaches operational limit. Fridge temperatures reach 1.1°C above normal (T=1hr 57min)
Frequency (Hz)
50.1
50
49.9
49.8
49.7
49.6
49.5
49.4
49.3
49.2
-0.5
0
0.5
1
1.5
2
Time (h)
Spinning reserve
DDC only

# Hybrid reserve service



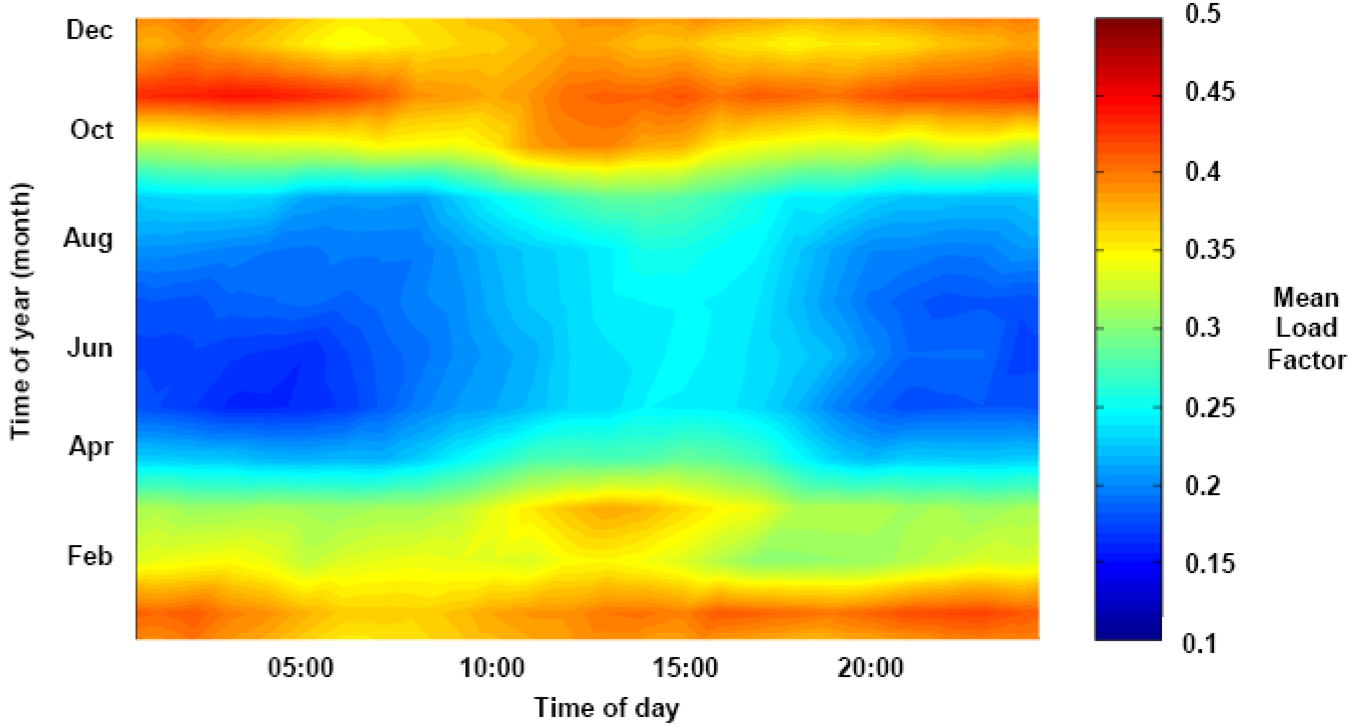

Operating the Electricity Transmission Networks in 2020, Follow Up Report, National Grid, February 2010

# Typical wind turbine power curve



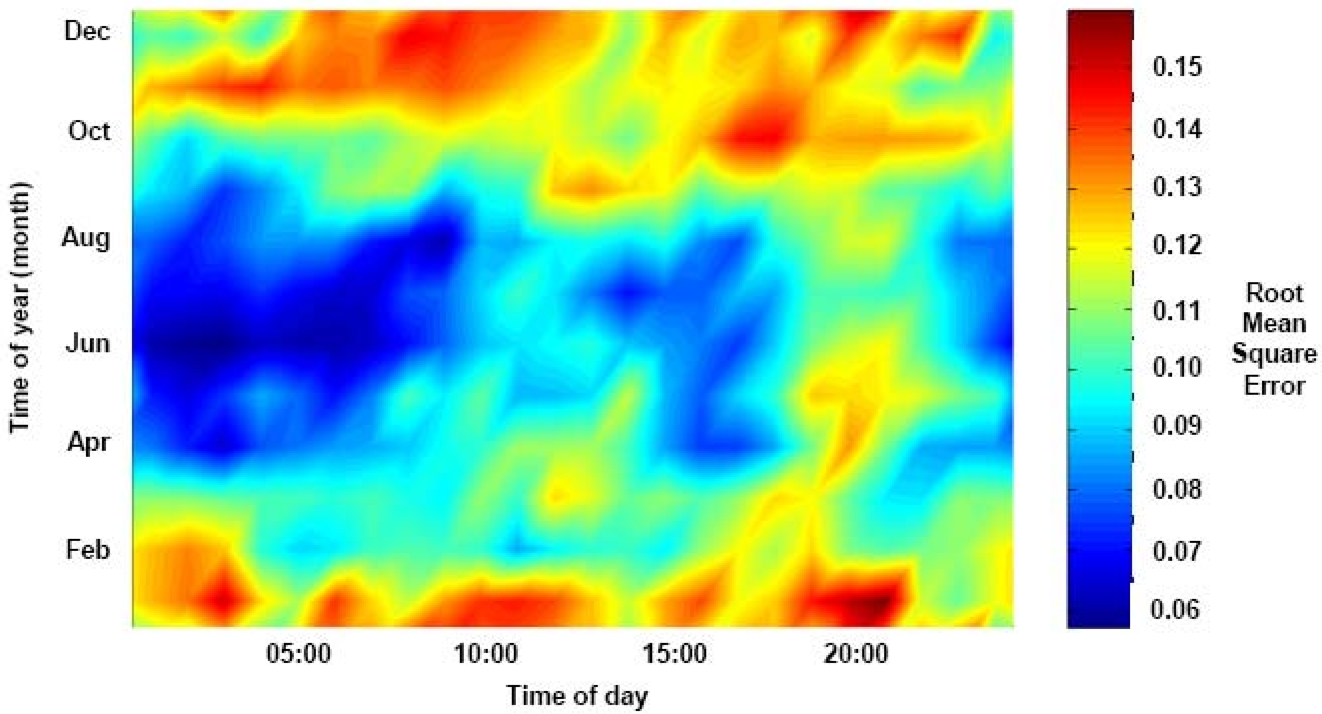

Operating the Electricity Transmission Networks in 2020
Initial Consultation, National Grid 2009

# Recorded wind load factors 2008



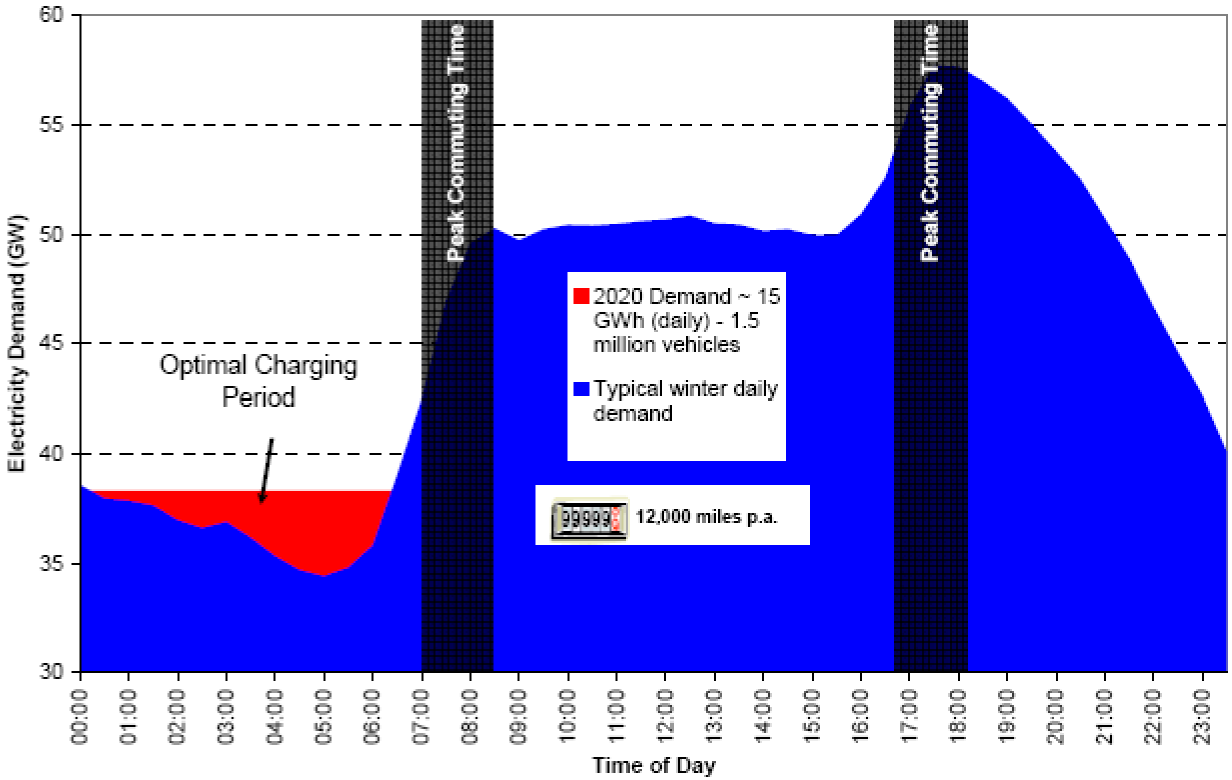

Operating the Electricity Transmission Networks in 2020
Initial Consultation, National Grid 2009

# Persistence errors in forecasting wind




Operating the Electricity Transmission Networks in 2020
Initial Consultation, National Grid 2009

# Matching vehicle charging to the current electricity demand profile




Operating the Electricity Transmission Networks in 2020
Initial Consultation, National Grid 2009

# British Electricity Transmission System

The  Transmission System broadly comprises all circuits operating at 400kV and 275kV. In Scotland transmission also includes 132kV networks.

The Transmission System is connected via interconnectors to transmission systems in France and Northern Ireland.


nationalgrid