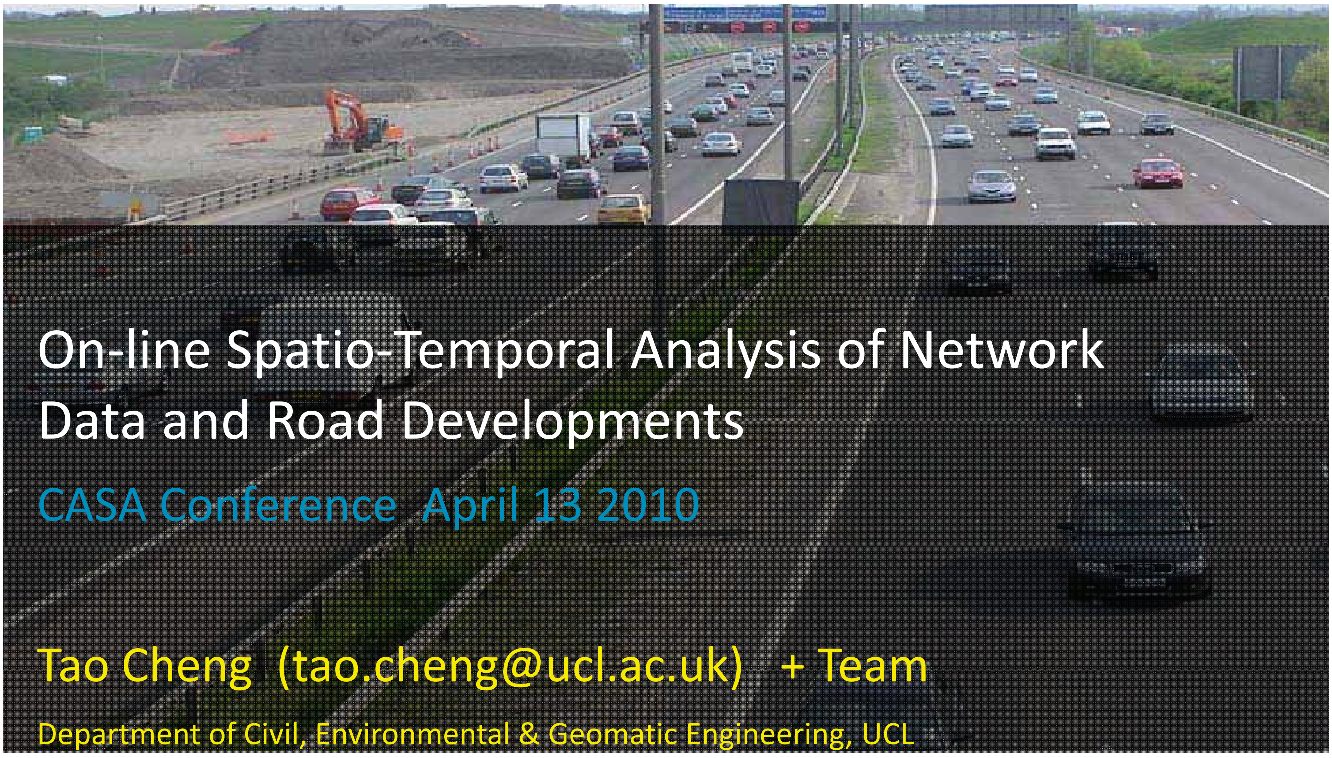


On-line Spatio-Temporal Analysis of Network Data and Road Developments

CASA Conference April 13 2010

Tao Cheng (tao.cheng@ucl.ac.uk) + Team

Department of Civil, Environmental & Geomatic Engineering, UCL

EPSRC

Transport for London

STANDARD
Spatio-Temporal Analysis of Network Data And Road Developments

Outline

- Introduction
  - Background and aim
- Methodology – integrated ST Data Mining
  - Statistical approach
  - Machining learning
  - Visualization
  - Simulation
- Programme and Progress

STANDARD
Spatio-Temporal Analysis of Network Data And Road Developments

Background

- Large cities are increasingly crowded - population & mobility
- Traffic congestion affects both the economy and daily life.
- It is difficult and expensive to increase the capacity of the road network.

City of London

- Traffic levels in the Congestion Charging Zone are falling but congestion levels are rising.
- cost of congestion
  - £3 billion per year
- Mayor’s traffic priorities
  - reduce congestion and smooth traffic flow
- Removal of western extension of CC (27/11/2008)
- Olympic Games 2012
  - travel time to London Olympic sites

london

STANDARD
Spatio-Temporal Analysis of Network Data And Road Developments

Why is this?

- Reduction in network capacity?
- Reallocations of capacity to other uses?
- Reduced resilience of the network?

Aim - To understand the traffic congestion in central London

- To quantitatively measure road network performance
- To understand causes of traffic congestion
  - association between traffic and interventions
    - traffic flow, speed/journey time
    - incidents, road works, signal changes and bus lane changes
- Case study – London

EPSRC

Transport for London

STANDARD
Spatio-Temporal Analysis of Network Data And Road Developments

Challenge (1) Network Complexity

1) Dynamics
2) Spatial dependence
3) Spatio-temporal interactions
4) Heterogeneity

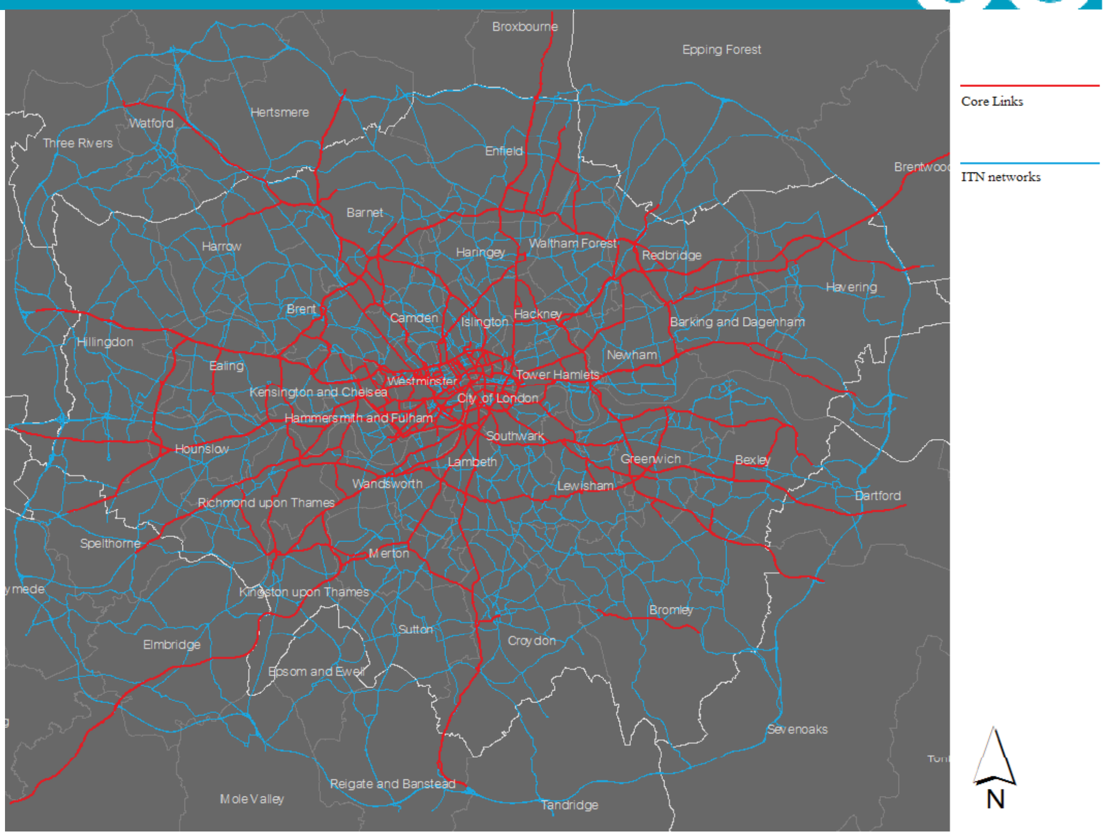
STANDARD
Spatio-Temporal Analysis of Network Data And Road Developments

Challenge (2) - Data issues

- massive – 20GB monthly
- multi-sourced related to 5 different networks
- different scales (density & frequency)
- variable data quality
- contain conflicts, errors, mistakes and gaps

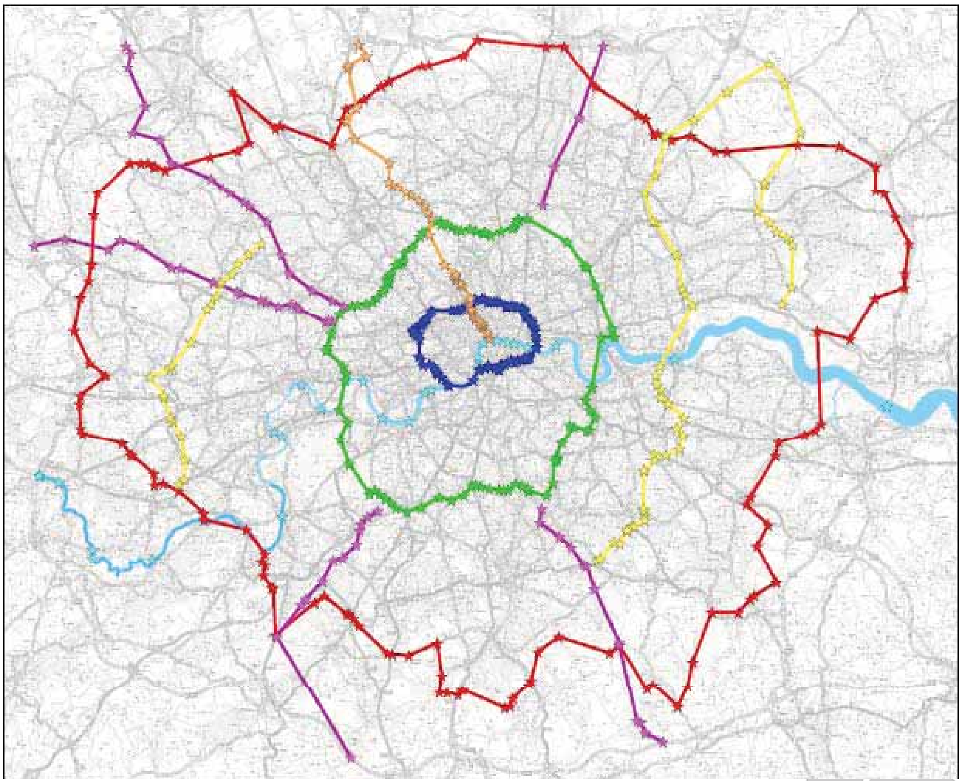
STANDARD
Spatio-Temporal Analysis of Network Data And Road Developments

DATA COVERAGE



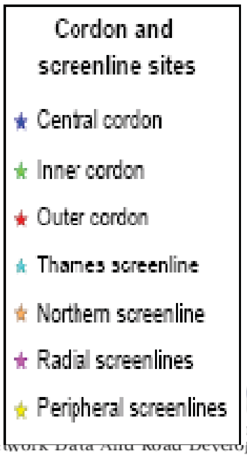
UCL

London Road Networks



Cordons
Central, Inner, Outer

Screenlines
Thames, Northern, five radials four peripherals



UCL

Traffic Flow Surveys

- NMC (National manual count annual data)
- ATC (Automatic Count) - 20 MB
- different time periods, intervals and accuracy

Traffic speed (and hence journey time) data

- **MCOS** (Moving Car Observer Surveys)
  - Centre, Inner, and Outer
  - least accurate of the datasets
- **ITIS** (GPS vehicle tracking system) - 2GB
  - major A roads and bus routes in town with 2000 probes
  - medium accuracy
- **ANPR** (Automatic Number Plate Reading) – 6 GB
  - main roads in the central and west extensions of CCZs
  - 5-minute intervals, 5 vehicle groups,
  - high accuracy
  - available since March 2008
- **At least 5 networks**
  - boundaries **do not fully align**

STANDARD
Spatio-Temporal Analysis of Network Data And Road Developments

LTIS incident and event data - 20MB

- **works, hazards, accidents, signal faults, special events, breakdowns, security, and other causes**
- **DfT have all these data as map or as text files**

- Minimal, Moderate, Serious or Severe

subjective ?

unrecorded?

not geocoded?

not broadcast on the traffic Link website, creating problems in analysis and reporting.

There are uncertainties and gaps in the intervention data

STANDARD
Spatio-Temporal Analysis of Network Data And Road Developments

Methodology - ISTDM

What's new - (1) data-driven, top-down

Transition in data availability

- Data scarcity:
  - High cost
  - Low volume
  - Intensive validation
- Data abundance:
  - High volume
  - Multiple kinds and sources
  - Extensive application

Transition in modelling approach

- Bottom-up:
  - Mechanistic
  - Explicit representation of behaviour (origin, dest^n, model, time …)
  - System properties by aggregation
- Top-down:
  - Phenomenological
  - Describe system gross of all behavioural responses
  - Direct to objectives

STANDARD
Spatio-Temporal Analysis of Network Data And Road Developments

UCL

What's new: (2) integrated space and time

Existing ST analysis methods

- time series analysis + spatial correlation
- spatial statistics + the time dimension
- time series analysis + artificial neural networks

ST dependence ≠ space + time

Integrated modelling of ST is needed -

- seamless & simultaneous
- ST-association/autocorrelation

STANDARD
Spatio-Temporal Analysis of Network Data And Road Developments

UCL

What's new: (3) hybrid/quantitative approach

- combine regression analysis with machine learning
  - improve the sensitivity and explanatory power
- study the heterogeneity and scale of road performance
  - optimal scale for monitoring
- Quantitative assessment of network performance
  - Sensible decision making & policy evaluation

STANDARD
Spatio-Temporal Analysis of Network Data And Road Developments

Principle of ST Modelling

Space-time data = global (deterministic) space-time trends + local (stochastic) space-time variations

$$Z_i(t) = \mu_i(t) + e_i(t)$$

Z(t)=u(t)+e(t)

$Z_i=u_i+e_i$

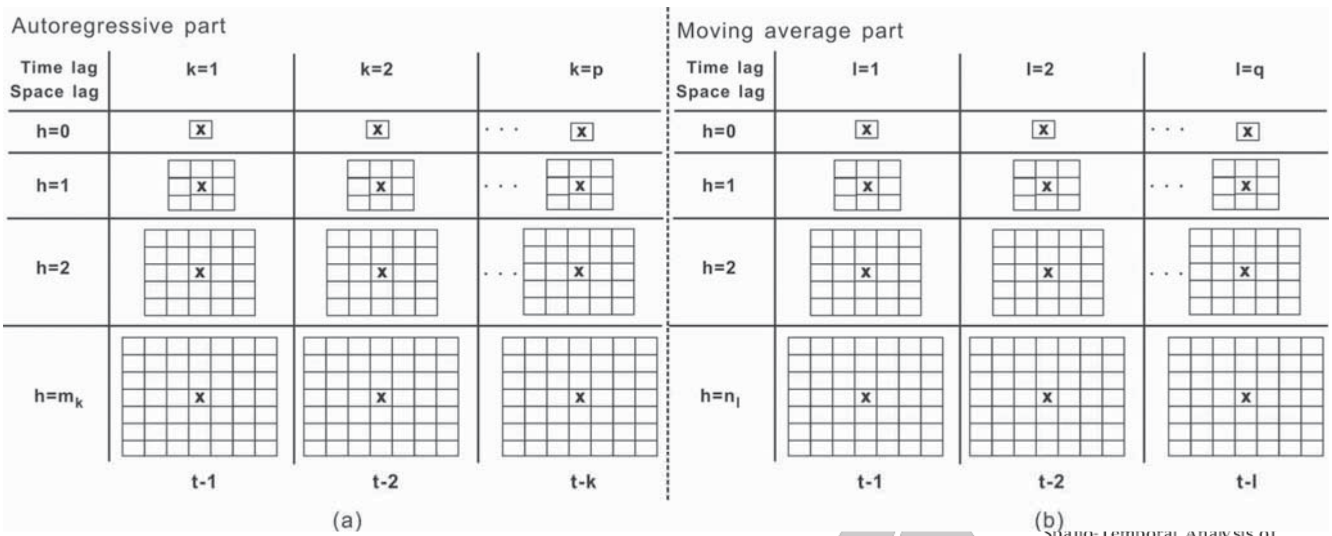
- $z_i(t)$ - the observation of the data series at spatial location *i* and at time *t*;
- $\mu_i(t)$ - space-time patterns that explain large-scale deterministic space-time trends and can be expressed as a nonlinear function in space and time.
- $e_i(t)$ - the residual term, a zero mean space-time correlated error that explains small-scale stochastic space-time variations.

Cheng, Wang, Li (forthcoming, Geographical Analysis)

STANDARD
Spatio-Temporal Analysis of Network Data And Road Developments

Model 1 - STARIMA - Spatio-Temporal Auto-Regressive Integrated Moving Average

$$z_i(t) = \sum_{k=1}^{p} \sum_{h=0}^{m_k} \phi_{kh} W^{(h)} z(t-k) - \sum_{l=1}^{q} \sum_{h=0}^{n_l} \theta_{lh} W^{(h)} \varepsilon(t-l) + \varepsilon(t)$$



(Pfeifer P E and Deutsch S J, 1980)

Our approach – Integrated modelling of ST

Model 1 – STARIMA

$$z_i(t) = \sum_{k=1}^{p}\sum_{h=0}^{m_k} \phi_{kh} W^{(h)} z(t-k) - \sum_{l=1}^{q}\sum_{h=0}^{n_l} \theta_{lh} W^{(h)} \varepsilon(t-l) + \varepsilon(t)$$

- define weights based upon spatial distance and spatial adjacency
- consider anisotropy
- able to model spatially continued phenomena

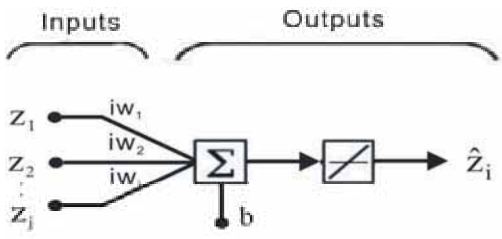
Tao Cheng, Jiaqiu Wang, Xia Li, 2010 A hybrid approach for space-time series of environmental data, *Geographical Analysis (forth coming)*

Model 2 - ANN - Artificial Neural Networks

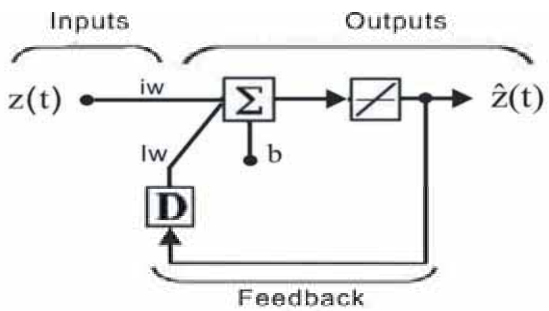
(Mandic D P and Chambers JA, 2001)

SFNN - spatial interpolation

DRNN - time series analysis



(a)



(b)

(a) static neuron

$$\hat{z}_i = \sum_{j=1}^{n} iw_{ij} \cdot z_j + b$$

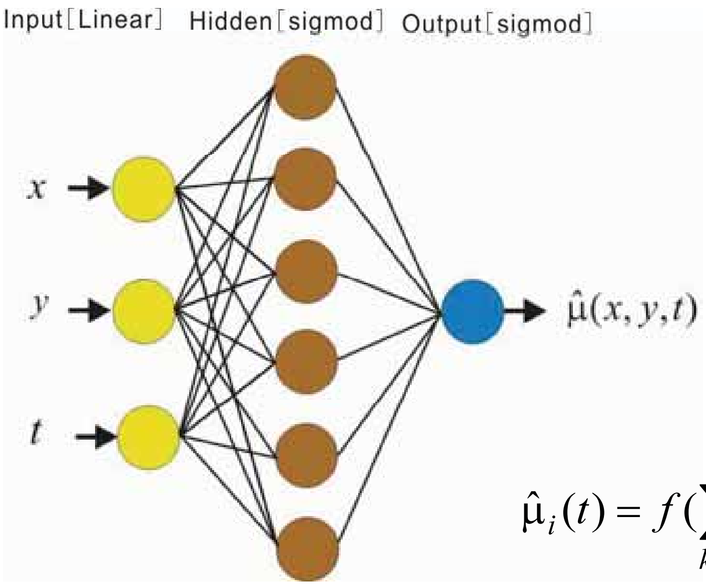
(b) dynamic neuron

$$\hat{z}(t) = iw \cdot z(t) + lw \cdot \hat{z}(t-1) + b$$

Cheng, Wang (2008, TGIS) Cheng, Wang (2009, CEUS)

STANDARD Spatio-Temporal Analysis of Network Data And Road Developments

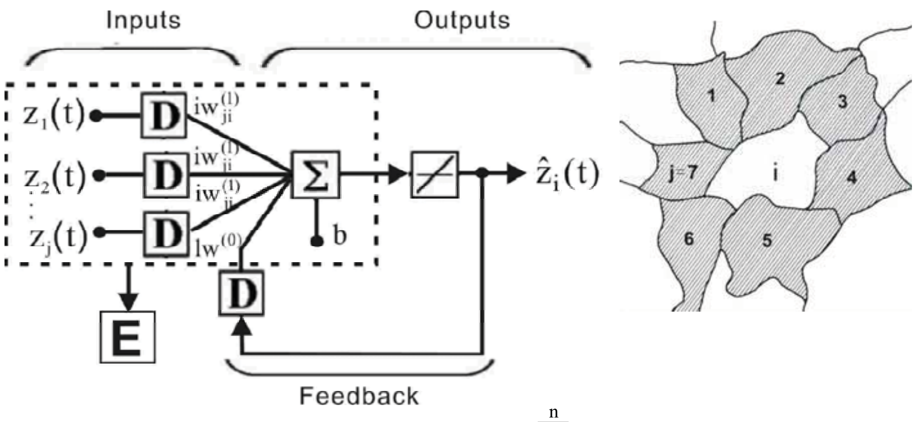
• ANN for space-time trend analysis



$$\hat{\mu}_i(t) = f(\sum_{k=1}^{n} \beta_k f(i,t) + \beta_0)$$

Tao Cheng, Jiaqiu Wang, 2009, Accommodating Spatial Associations in DRNN for Space-Time Analysis, *Computers, Environment and Urban System*, 33, 409-418.

Model 2 - STANN



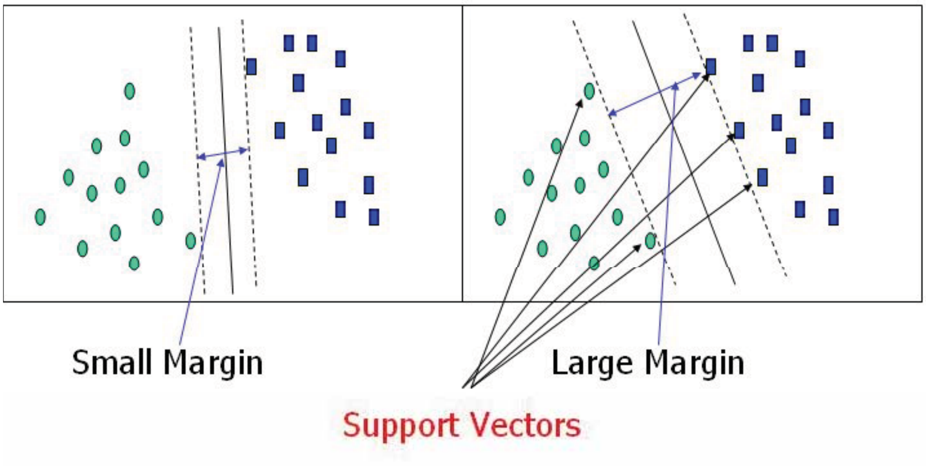
Space-Time Neuron $\hat{z}_i(t) = \sum_{j=1}^{n} iw_{ji}^{(1)} \cdot z_j(t-1) + lw^{(0)} \cdot \hat{z}_i(t-1) + b$

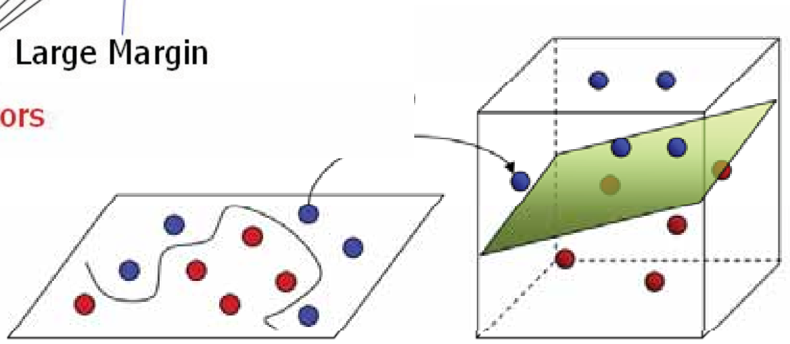
- One step implementation of ANN+ STARIMA
- Accommodate ST associations in ANN
- Deal with nonlinearity & heterogeneity in BP learning

Jiaqiu Wang, Tao Cheng, STANN – Modeling Space-Time Series by Artificial Neural Networks, *International Journal of Geographical Information Science*, under review

STANDARD
Spatio-Temporal Analysis of Network Data And Road Developments

Model 3 – SVM - Support Vector Machines





SVC & SVR (Vapnik et al, 1996) Input Space Feature Space

Model 3 - STSVR

- Nonlinear Spatio-Temporal Regression by SVM

$$\phi(f_r, f_s) = \sum_{i=1}^{N} (\alpha_i^* - \alpha_i) K((f_r, f_s)', (f_r, f_s))$$

- Develop ST kernel function
- Overcome over-fitting in STANN
- Deal with errors
- Model nonlinearity & heterogeneity

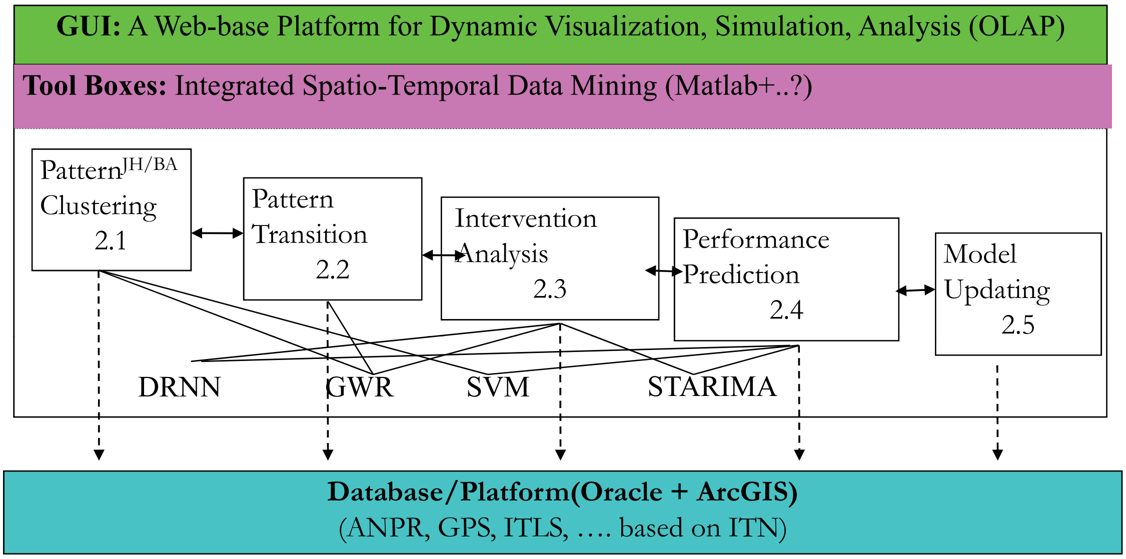
Jiaqiu Wang, Tao Cheng, James Haworth, Space-Time Kernels, submitted to Spatial Data Handling (SDH) 2010, Hong Kong, May 26-28.

Other methods

- Geographically Weighted Regression (GWR)
  - -> STWR?
- Permutation Scan Statistics (PSS)
  - -> STPSS? (or STC)
- Exploratory Visualization (DM) + ST+OLAP
  - -> STOEV?
- Simulation (Multi-scales)
  - -> STMSS?

STANDARD
Spatio-Temporal Analysis of Network Data And Road Developments

Progress

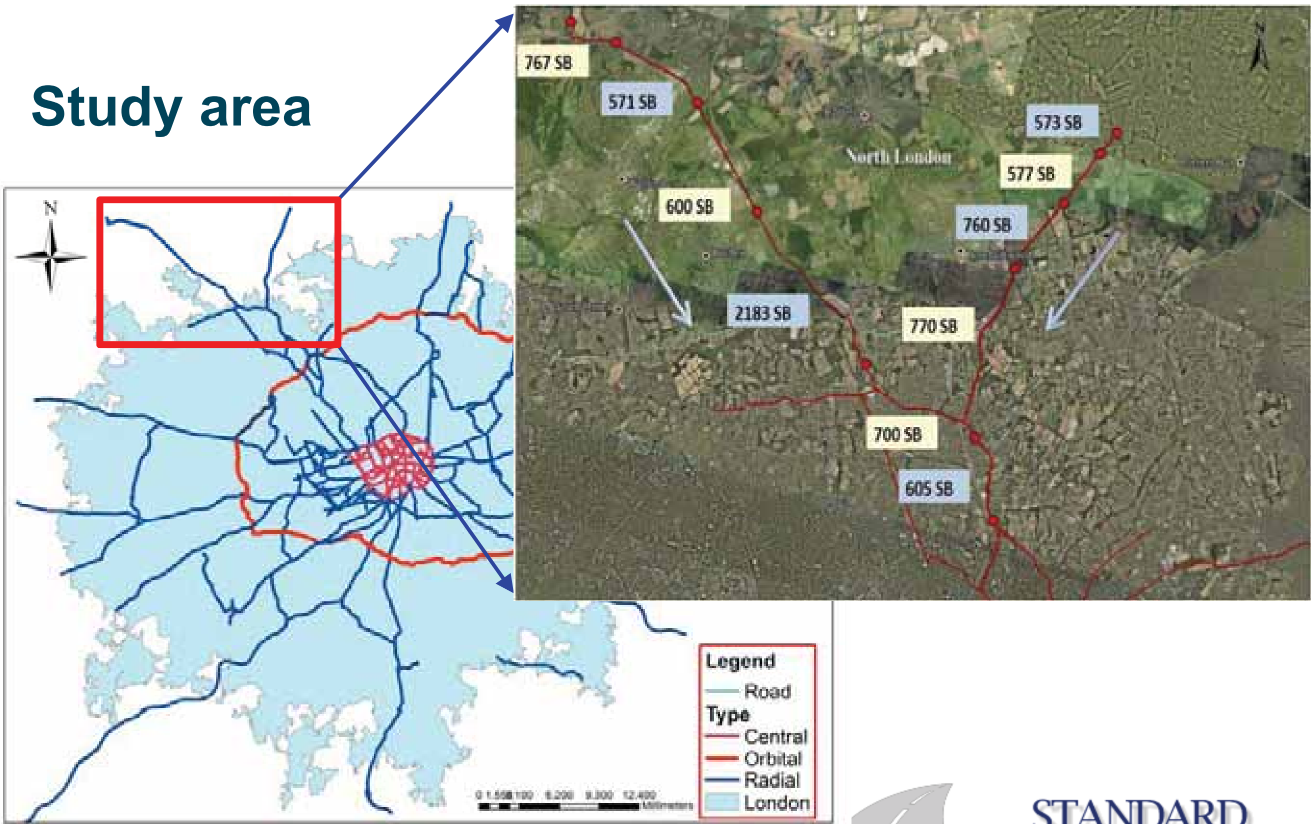


STANDARD Platform Structure

STANDARD
Spatio-Temporal Analysis of Network Data And Road Developments

STARIMA for Journey Time Prediction in London

Study area



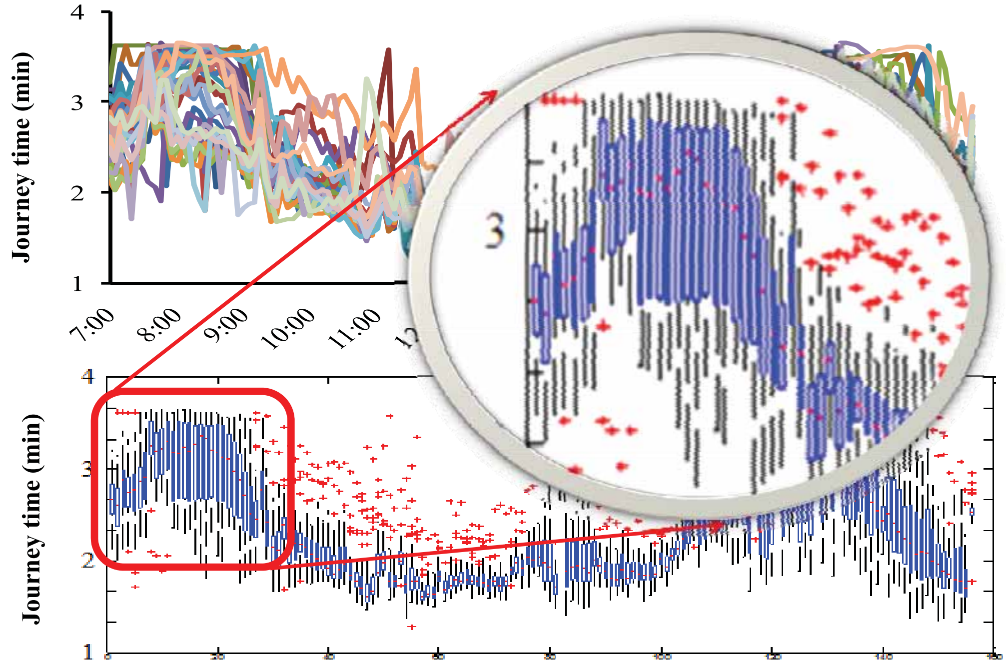
London Arterial Road Map

STANDARD
Spatio-Temporal Analysis of Network Data And Road Developments

Pattern analysis of journey time

(a)

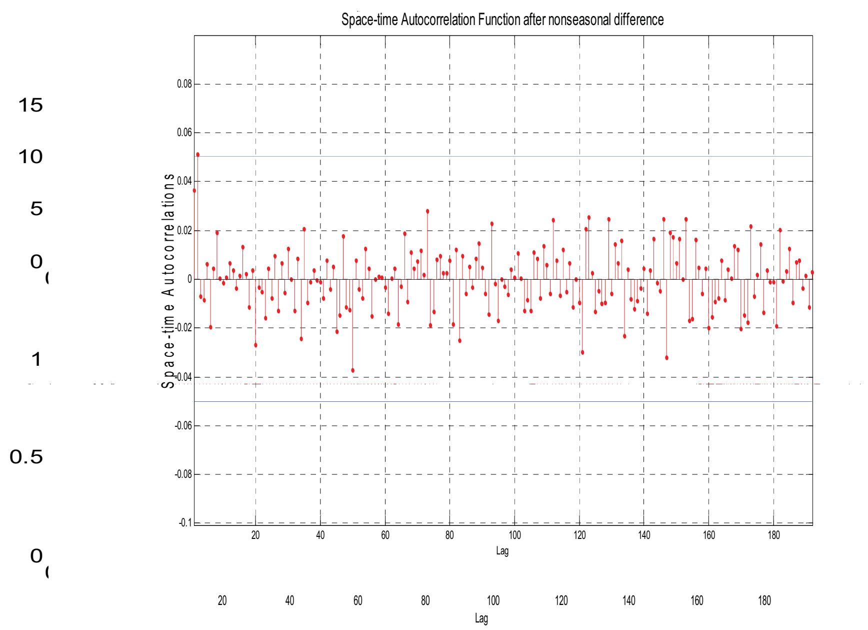
(b)



The distribution plot of 33 Mondays journey times link 605 during 07:00 to 19:00 (2009 Jan. – Aug.)

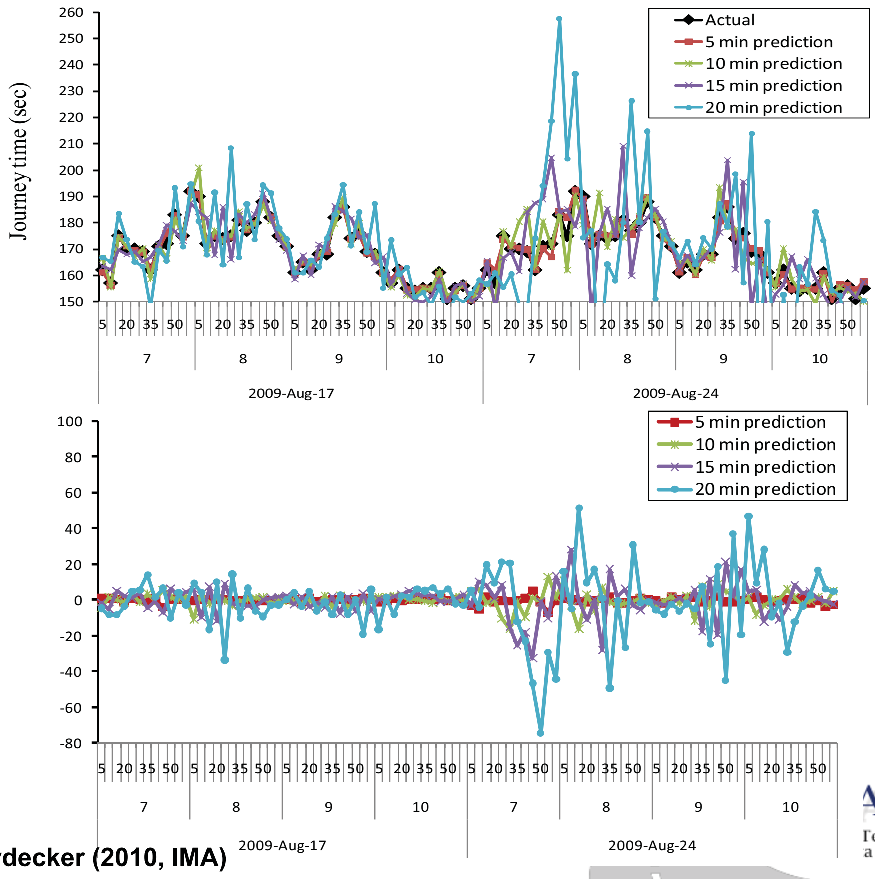
STANDARD
Spatio-Temporal Analysis of Network Data And Road Developments

Space-time analysis



STANDARD
Spatio-Temporal Analysis of Network Data And Road Developments

Results



Wang, Cheng, Heydecker (2010, IMA)

Accuracy

Prediction Accuracy at different prediction intervals

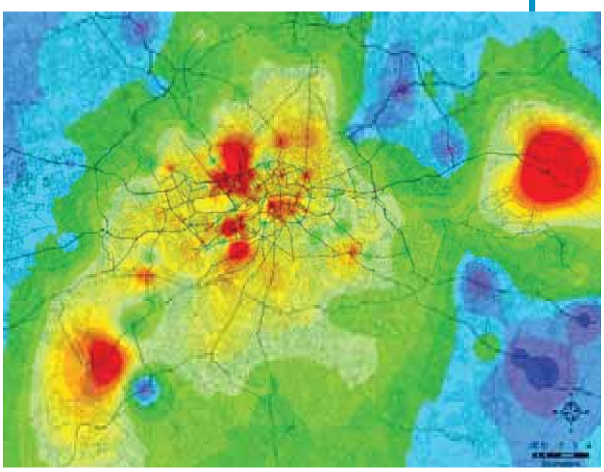
	Forecasting Horizon			
	5 min	10 min	15 min	20 min
Number of validate prediction	96	93	86	70
Mean relative error	0.07%	0.25%	0.44%	0.81%
Standard deviation of relative error	0.16%	0.38%	0.77%	1.27%

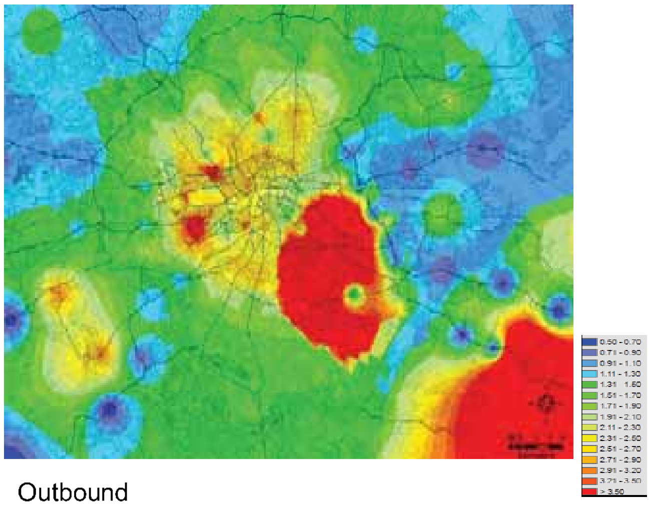
Comparison of results from extended STARIMA model and standard STARIMA model (Kamarianakis and Prastacos, 2005) at 5 min interval

	Number of validate prediction	Mean relative error	Standard deviation of relative error
Extended STARIMA	96	0.07%	0.16%
Standard STARIMA	95	0.11%	0.41%

STANDARD
Spatio-Temporal Analysis of Network Data And Road Developments

Visualization of traffic congestion in space-time



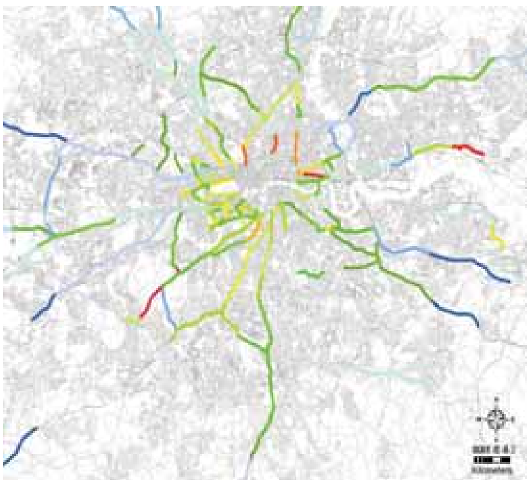


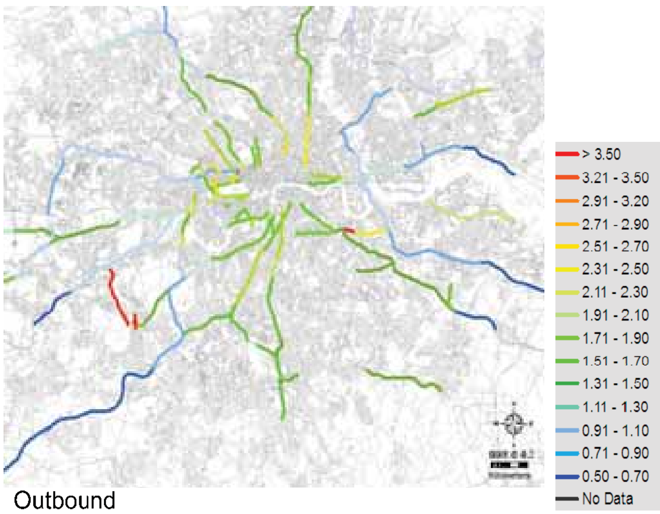
Inbound Outbound

Figure 1. Delay at 9 am on 12th April 2009

Cheng, Emmonds, Tanaksaranond, Sonoiki (2010, GISRUK)

STANDARD
Spatio-Temporal Analysis of Network Data And Road Developments

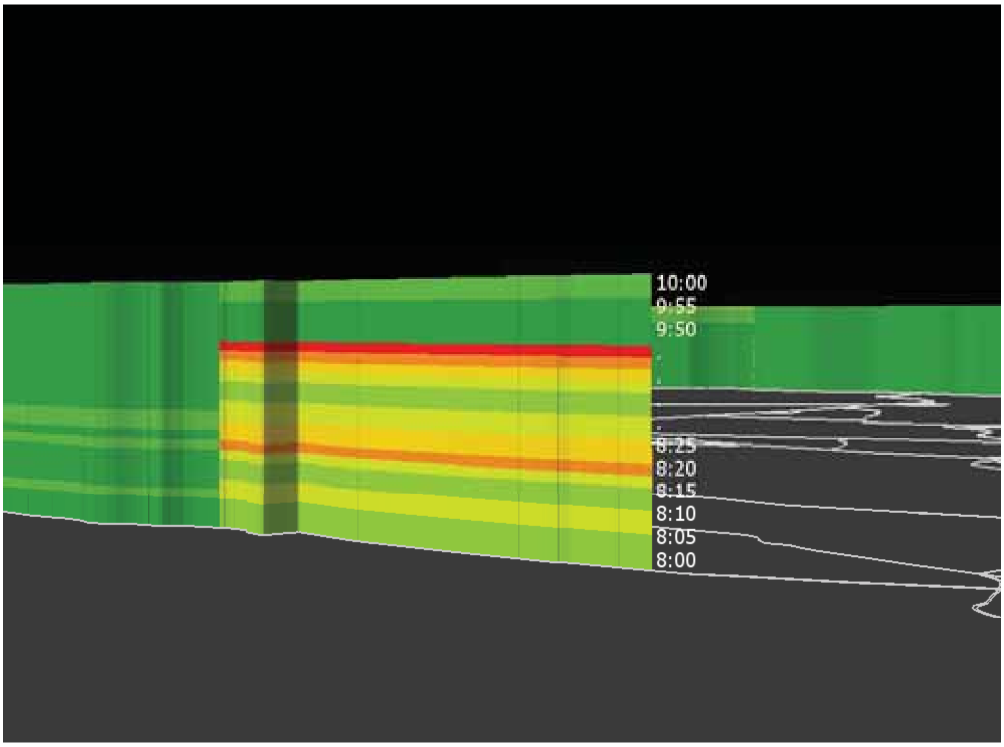




Inbound Outbound

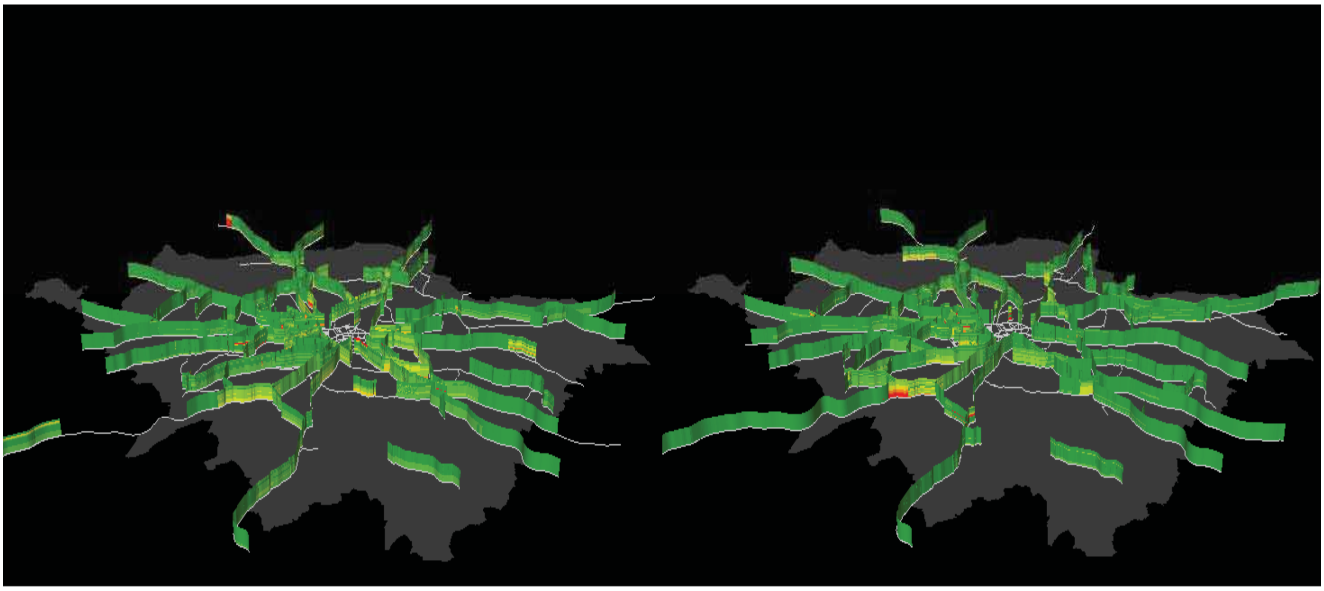
Figure 2. Delay at 9:15am on 12th April 2009

STANDARD
Spatio-Temporal Analysis of Network Data And Road Developments



STANDARD
Spatio-Temporal Analysis of Network Data And Road Developments

LCAP 15 January 2010 8:00-10:00 am

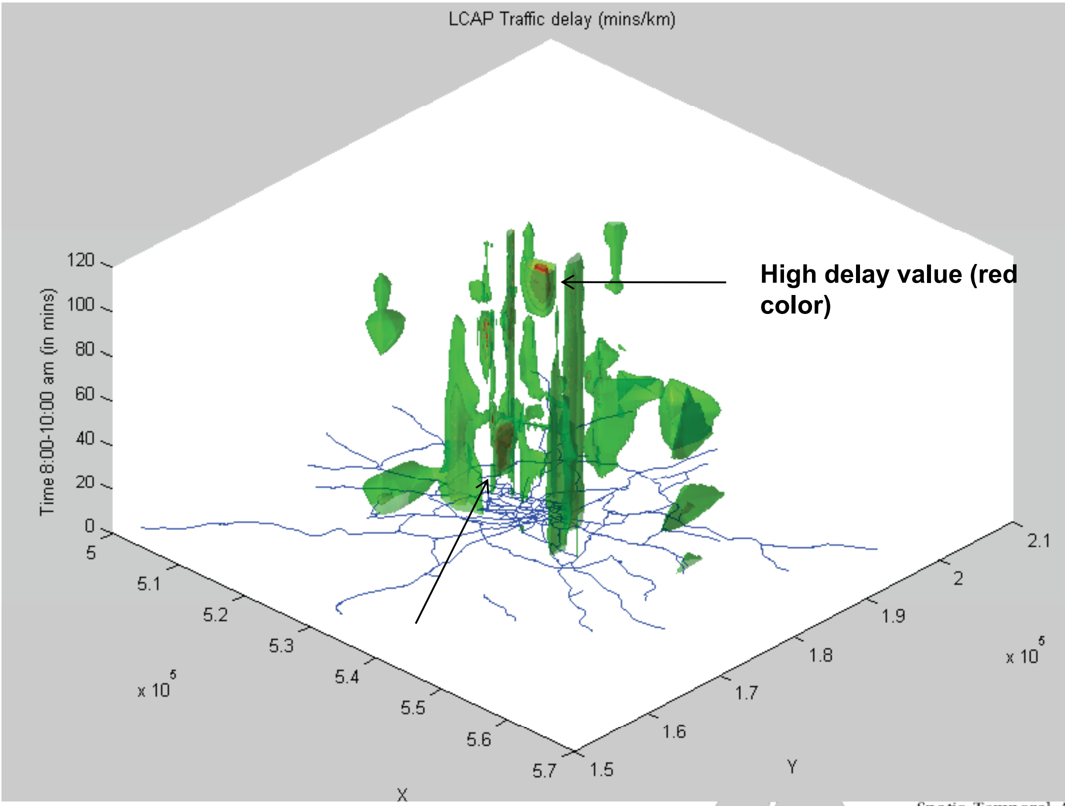


Inbound

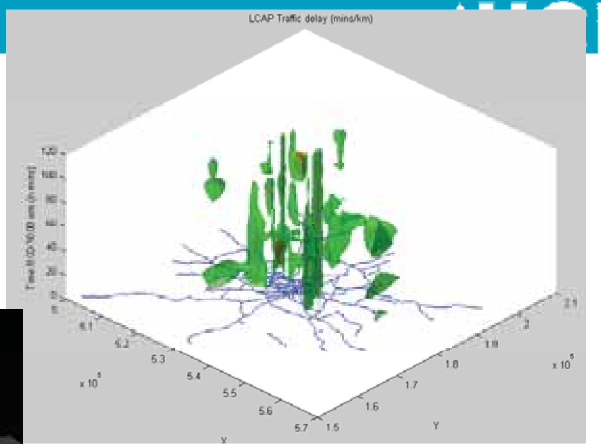
Outbound

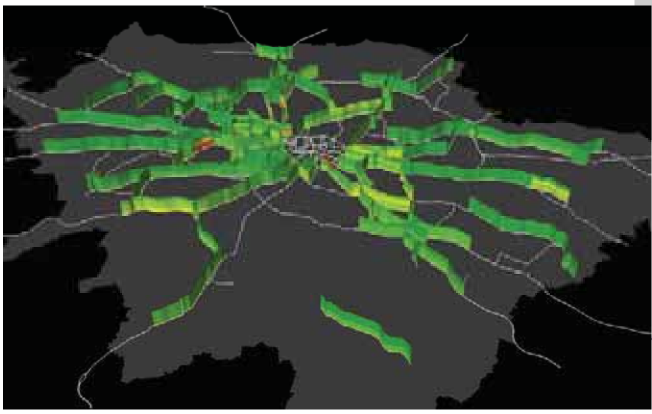
STANDARD
Spatio-Temporal Analysis of Network Data And Road Developments

Isosurface



Sideview





Topview

Spatio-Temporal Analysis of
Network Data And Road Developments

UCL

Links	Basic Temporal Intervals
[578,720]	[152,167,204,237,262]
[579,580]	[153,155,198,224,226,253,257,283]
[580,579]	[153,155,198,224,226,253,257,283]
[599,719]	[113,117,126,136,137,139,144,206,220,221]
[599,719,666]	139
[665,666]	[38,128,141,194,203,229]
[666,665]	[38,128,141,194,203,229]
[666,719]	[79,94,123,139,148,150,166,200,264]
[666,719,599]	139
[719,599]	[113,117,126,136,137,139,144,206,220,221]
[719,666]	[79,94,123,139,148,150,166,200,264]
[720,578]	[152,167,204,237,262]

Cheng, Anbaroglu (2010, SDH)

STANDARD
Spatio-Temporal Analysis of Network Data And Road Developments

UCL

James Haworth

Department of Civil, Environmental & Geomatic Engineering, UCL

Multi-scale analysis of road network performance

- Using spatio-temporal data mining techniques to look for patterns in congestion at varying spatial and temporal scales
- What patterns can be observed in inbound and outbound congestion...
  - Daily? Weekly? Seasonally?...
- Identification of recurrent and non-recurrent congestion in London

EPSRC

Transport for London

EP/G023212/1

STANDARD
Spatio-Temporal Analysis of Network Data And Road Developments

UCL

Understanding Road Congestion as an Emergent Property of Traffic Networks

Macroscopic

Flow and economic models based on 'known' road capacity

Microscopic

Individual behaviours simplistic and route choice macroscopic-driven

STANDARD
Spatio-Temporal Analysis of Network Data And Road Developments

Picture credits: DOT California, PTV, Paramics

UCL

Formal model of Emergence

Link level

What causes congestion to emerge at link level?

What is the effect of road layout?

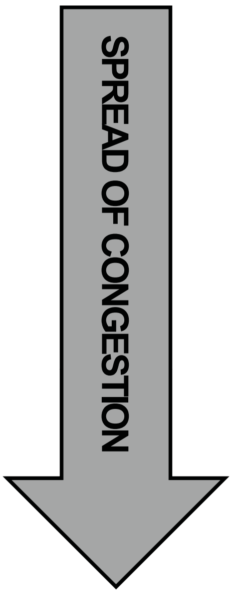
Junction level

Are junctions the key source of congestion?

What choices are available to drivers?

Network level

How does congestion spread to the whole network?

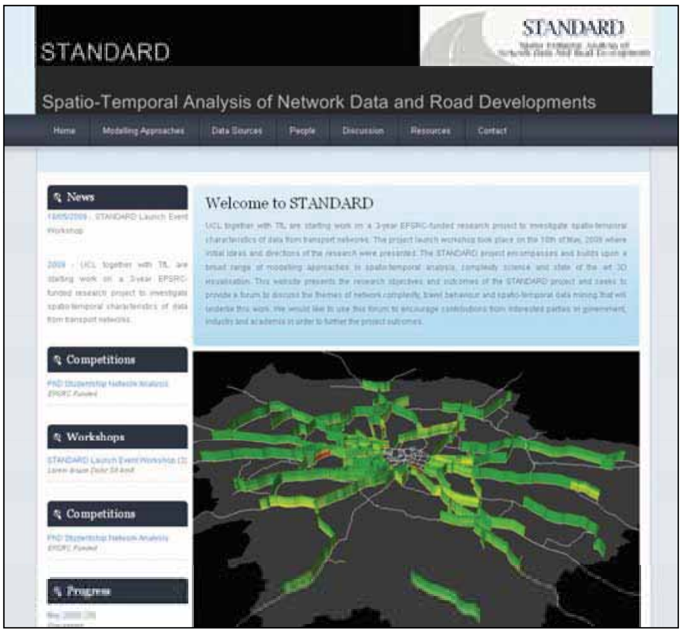


Manley, Cheng (2010, IMCIC)

STANDARD
Spatio-Temporal Analysis of Network Data And Road Developments

UCL

STANDARD Website – http://standard.cege.ucl.ac.uk



STANDARD
Spatio-Temporal Analysis of Network Data And Road Developments

UCL

Conclusion - Network Complexity

Can we predict/migrate emergence (congestion) of Road Network ?

Understand → predict/migrate

Detect → emergence

Model → (congestion)

Simulation → Road Network

STANDARD
Spatio-Temporal Analysis of Network Data And Road Developments

Acknowledgements

UCL

Transport for London

EPSRC
Pioneering research and skills

ARUP

ublox

Location and Timing

973
National Basic Research Program of China

中华人民共和国科学技术部
Science and Technology of the People's Rep

National High-tech R&D Program
(863 Program)

CASM Chinese Academy of Surveying and Mapping

NSF China

测绘遥感信息工程国家重点实验室
LIESMARS State Key Laboratory of Information Engineering in Surveying, Mapping and Remote Sensing