Dynamic chain graphs for high-dimensional time series: an application to real-time traffic flow forecasting

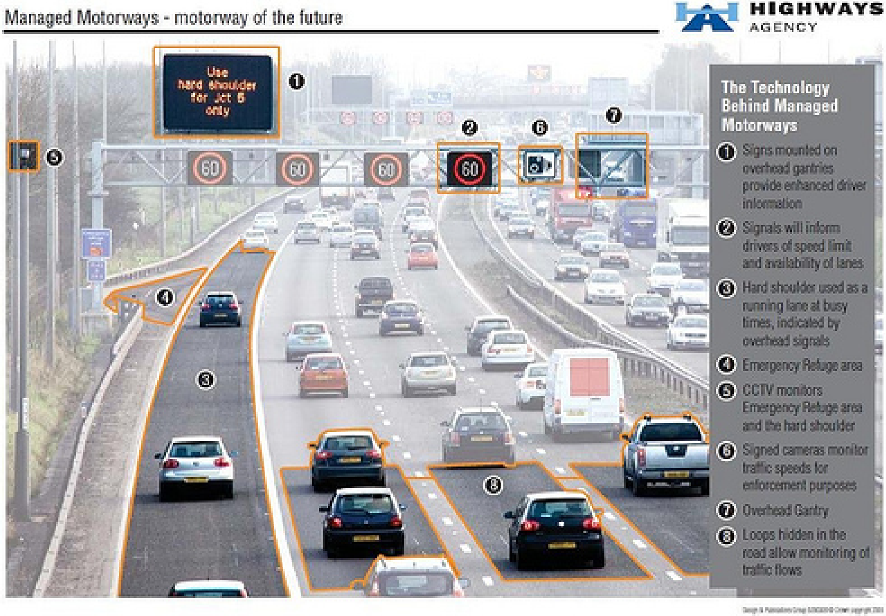
Catriona Queen
Joint work with Osvaldo Anacleto [1]

Department of Mathematics and Statistics
The Open University

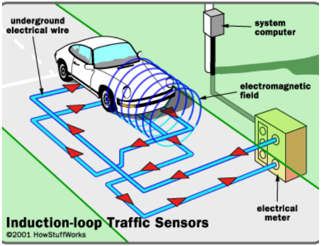
University of Bristol Seminar
February 2014

[1]Now at Roslin Institute, Edinburgh University

The Managed Motorways project



Traffic modelling



- Minute-by-minute data at different locations in a network.
- Online, real-time environment.
- Road managers need to take decisions given traffic conditions.

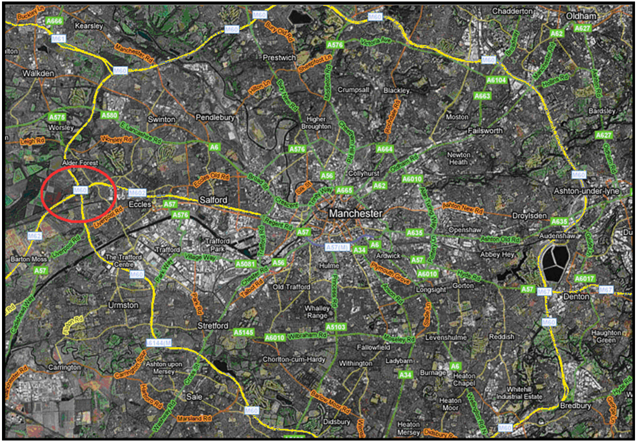
Real-time flow forecasts are crucial for road management

Traffic modelling background

- Several approaches applied to traffic modelling: mathematical models, neural networks, ARIMA, state space models...
- Usually difficult to provide real-time forecasts.
- Most models are univariate.

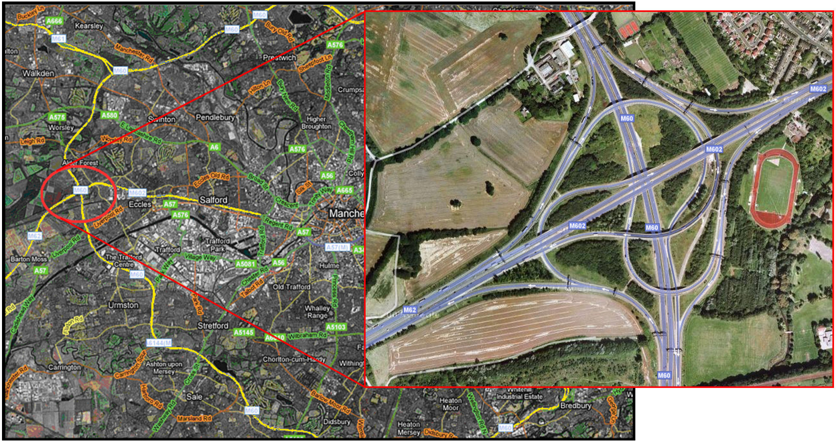
Challenge: develop multivariate flow forecasting models for an on-line environment

Focus: M60/M62/M602 intersection



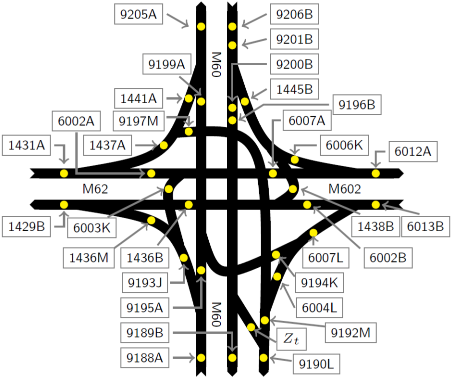
Pictures taken from Google Maps

Focus: M60/M62/M602 intersection



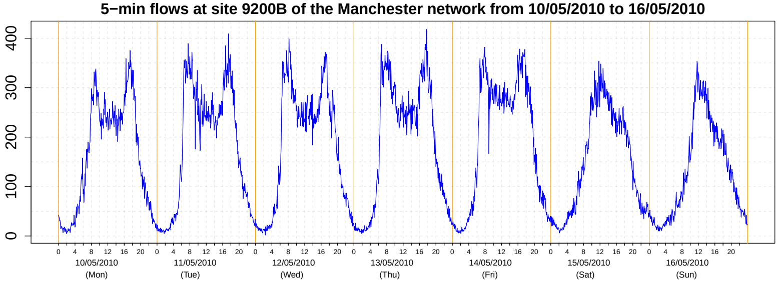
Pictures taken from Google Maps

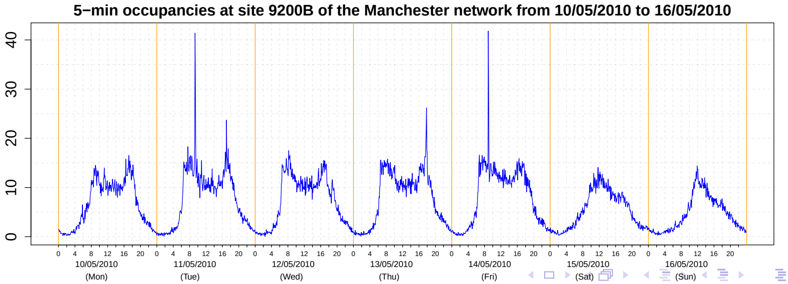
Manchester network



- Data available for 31 sites.
- Minute-by-minute flow data.
- Aggregate data to 5 min intervals.
- It takes less than five minutes to traverse the network.
- Flow modelled using occupancy, speed and headway as predictors.

Flow and occupancy weekly series: example





The statistical problem

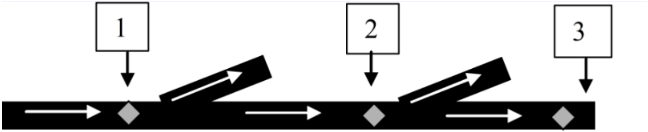
Multivariate time series: observations are taken simultaneously over time.

Goal: develop a multivariate model which accommodates the interrelationships among the series

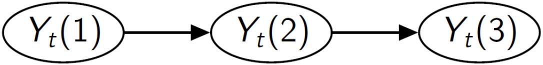
As the number of time series increases modelling becomes challenging.

Approach: graphical modelling to enable a complex problem to be split up into simpler ones.

Road layout example:



Directed acyclic graph (DAG):



DAG defines a set of conditional distributions: child | parents

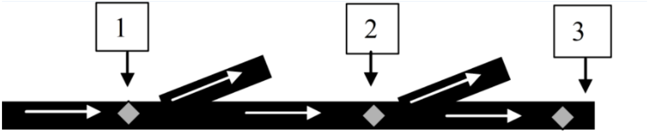
Breaks multivariate problem into univariate conditionals:

$Y_t(1)$ $\qquad$ $Y_t(2) | Y_t(1)$ $\qquad$ $Y_t(3) | Y_t(2)$

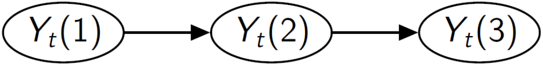
DAGs for traffic networks: Queen *et al.*, *ANZJS* (2007)

Representing a network by a graph

Road layout example:



Directed acyclic graph (DAG):

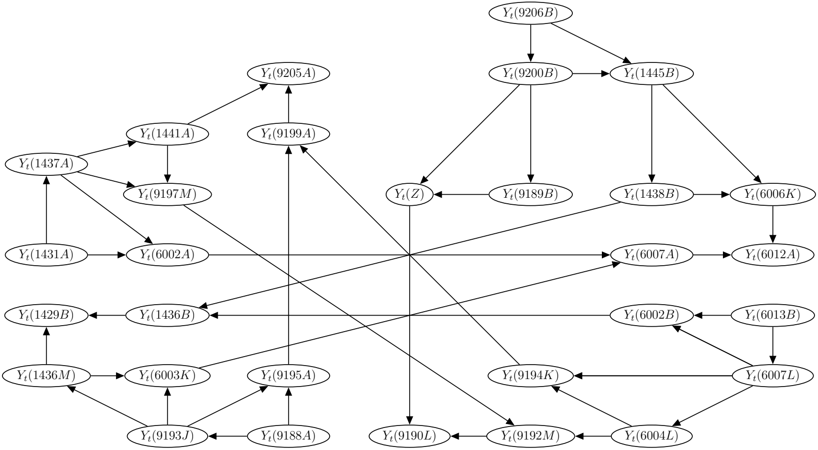


DAG defines a set of conditional distributions: child | parents

Breaks multivariate problem into univariate conditionals:
$Y_t(1)$ $\quad Y_t(2)|Y_t(1)$ $\quad Y_t(3)|Y_t(2)$

DAGs for traffic networks: Queen *et al.*, *ANZJS* (2007)

Manchester DAG



Main idea

Graphs allow local computations for joint probability distributions

graphical models theory + dynamic modelling approach = **dynamic graphical models**

Multiregression dynamic model (MDM)

- MDM is a dynamic Bayesian network.
- A DAG is used to decompose $f(\mathbf{y}_t)$ at each time t.
- Each $Y_t(i)|$parents is a Bayesian dynamic model, $i = 1, \ldots, n$.

MDM definition

Represent multivariate time series $\mathbf{Y}_t = (Y_t(1), \ldots, Y_t(n))^T$ by a DAG.
Observation equations:

$$Y_t(i) = \mathbf{F}_t(i)^\top \boldsymbol{\theta}_t(i) + v_t(i), \quad v_t(i) \sim N(0, V_t(i)), \quad i = 1, ..., n,$$

- $\mathbf{F}_t(i)$ is function of contemporaneous values of the parents of $Y_t(i)$,

Multiregression dynamic model (MDM)

- MDM is a dynamic Bayesian network.
- A DAG is used to decompose $f(\mathbf{y}_t)$ at each time t.
- Each $Y_t(i)|$parents is a Bayesian dynamic model, $i = 1, \ldots, n$.

MDM definition

Represent multivariate time series $\mathbf{Y}_t = (Y_t(1), \ldots, Y_t(n))^T$ by a DAG.
Observation equations:

$$Y_t(i) = \mathbf{F}_t(i)^\top \boldsymbol{\theta}_t(i) + v_t(i), \quad v_t(i) \sim N(0, V_t(i)), \quad i = 1, ..., n,$$

System equation: $\boldsymbol{\theta}_t = \mathbf{G}_t \boldsymbol{\theta}_{t-1} + \mathbf{w}_t, \quad \mathbf{w}_t \sim N(0, \mathbf{W}_t)$

- $\mathbf{F}_t(i)$ is function of contemporaneous values of the parents of $Y_t(i)$,
- $\mathbf{G}_t$ and $\mathbf{W}_t$ are block diagonal matrices,

Multiregression dynamic model (MDM)

- MDM is a dynamic Bayesian network.
- A DAG is used to decompose $f(\mathbf{y}_t)$ at each time t.
- Each $Y_t(i)|$parents is a Bayesian dynamic model, $i = 1, \ldots, n$.

MDM definition

Represent multivariate time series $\mathbf{Y}_t = (Y_t(1), \ldots, Y_t(n))^T$ by a DAG.
Observation equations:

$$Y_t(i) = \mathbf{F}_t(i)^\top \boldsymbol{\theta}_t(i) + v_t(i), \quad v_t(i) \sim N(0, V_t(i)), \quad i = 1, ..., n,$$

System equation: $\boldsymbol{\theta}_t = \mathbf{G}_t \boldsymbol{\theta}_{t-1} + \mathbf{w}_t, \quad \mathbf{w}_t \sim N(0, \mathbf{W}_t)$

Initial information: $\theta_0 | D_0 \sim N(\mathbf{m}_0, \mathbf{C}_0)$

- $\mathbf{F}_t(i)$ is function of contemporaneous values of the parents of $Y_t(i)$,
- $\mathbf{G}_t$ and $\mathbf{W}_t$ are block diagonal matrices,
- $\mathbf{C}_0$ also block diagonal.

Multiregression dynamic model (MDM)

- MDM is a dynamic Bayesian network.
- A DAG is used to decompose $f(\mathbf{y}_t)$ at each time t.
- Each $Y_t(i)|\text{parents}$ is a Bayesian dynamic model, $i = 1, \ldots, n$.

MDM definition

Represent multivariate time series $\mathbf{Y}_t = (Y_t(1), \ldots, Y_t(n))^T$ by a DAG.
Observation equations:

$$Y_t(i) = \mathbf{F}_t(i)^\top \boldsymbol{\theta}_t(i) + v_t(i), \quad v_t(i) \sim N(0, V_t(i)), \quad i = 1, ..., n,$$

System equation: $\boldsymbol{\theta}_t = \mathbf{G}_t \boldsymbol{\theta}_{t-1} + \mathbf{w}_t, \quad \mathbf{w}_t \sim N(0, \mathbf{W}_t)$

Initial information: $\boldsymbol{\theta}_0 | D_0 \sim N(\mathbf{m}_0, \mathbf{C}_0)$

- $\mathbf{F}_t(i)$ is function of contemporaneous values of the parents of $Y_t(i)$,
- $\mathbf{G}_t$ and $\mathbf{W}_t$ are block diagonal matrices,
- $\mathbf{C}_0$ also block diagonal.

The MDM cont.

- $\boldsymbol{\theta}_0(1), \ldots, \boldsymbol{\theta}_0(n)$ independent $\Rightarrow$ $\boldsymbol{\theta}_t(1), \ldots, \boldsymbol{\theta}_t(n)$ independent after observing $\mathbf{y}_t$.
- Each $Y_t(i)$ modelled separately by (conditional) univariate dynamic linear model (DLM) with parents as regressors.
- Separate forecast distributions for $Y_t(i) \mid$ parents.
- Marginal forecast moments for $Y_t(i)$ calculated separately.
- After $\mathbf{Y}_t$ observed, each $\theta_t(i)$ updated separately.

The MDM cont.

- $\boldsymbol{\theta}_0(1), \ldots, \boldsymbol{\theta}_0(n)$ independent $\Rightarrow$ $\boldsymbol{\theta}_t(1), \ldots, \boldsymbol{\theta}_t(n)$ independent after observing $\mathbf{y}_t$.
- Each $Y_t(i)$ modelled separately by (conditional) univariate dynamic linear model (DLM) with parents as regressors.
- Separate forecast distributions for $Y_t(i) \mid$ parents.
- Marginal forecast moments for $Y_t(i)$ calculated separately.
- After $\mathbf{Y}_t$ observed, each $\boldsymbol{\theta}_t(i)$ updated separately.
- Model computations fast and relatively straightforward, no matter how complex the network.
- DLM-related techniques can be used.

MDM: Queen and Smith, *JRSSB* (1993)

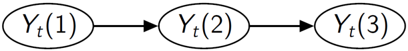
The MDM cont.

- $\boldsymbol{\theta}_0(1), \ldots, \boldsymbol{\theta}_0(n)$ independent $\Rightarrow$ $\boldsymbol{\theta}_t(1), \ldots, \boldsymbol{\theta}_t(n)$ independent after observing $\mathbf{y}_t$.
- Each $Y_t(i)$ modelled separately by (conditional) univariate dynamic linear model (DLM) with parents as regressors.
- Separate forecast distributions for $Y_t(i) \mid$ parents.
- Marginal forecast moments for $Y_t(i)$ calculated separately.
- After $\mathbf{Y}_t$ observed, each $\boldsymbol{\theta}_t(i)$ updated separately.
- Model computations fast and relatively straightforward, no matter how complex the network.
- DLM-related techniques can be used.

MDM: Queen and Smith, *JRSSB* (1993)

Example MDM

Directed acyclic graph:



Linear MDM consists of 3 separate univariate DLMs for:

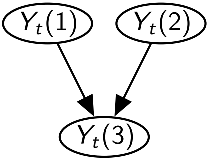
- $Y_t(1)$: any appropriate univariate DLM.
- $Y_t(2) \mid y_t(1)$: DLM for $Y_t(2)$ with $y_t(1)$ as linear regressor.
- $Y_t(3) \mid y_t(2)$: DLM for $Y_t(3)$ with $y_t(2)$ as linear regressor.

MDM for traffic modelling

MDM copes with traffic modelling problems in real-time

- Seasonality modelling in high-frequency traffic time series.
- Time-varying flow variances.
- Measurement errors.
- Use of extra variables as predictors.

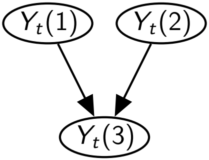
An MDM restriction



Under the MDM, forecast covariance of $Y_t(1)$ and $Y_t(2)$ is 0.

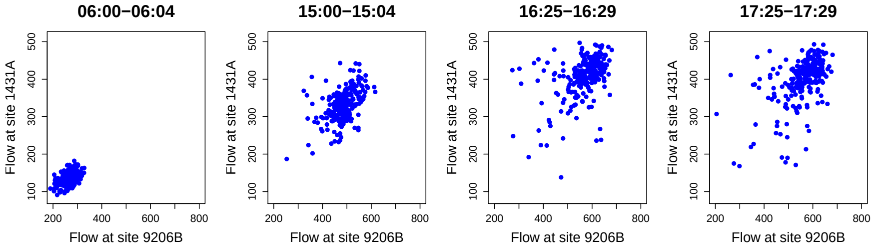
Not always true in practice.

An MDM restriction

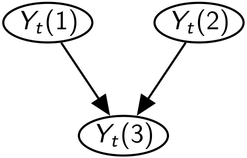


Under the MDM, forecast covariance of $Y_t(1)$ and $Y_t(2)$ is 0.

Not always true in practice.

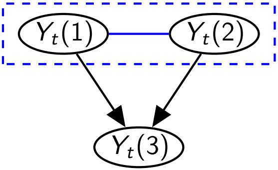


Motivation for a new model



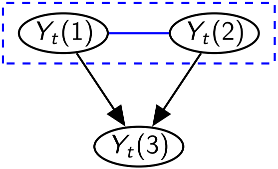
- DAG assumes a complete ordering of the time series components.
- In many cases only a partial ordering is available.
- A chain graph allows for partial orderings between time series components.

Motivation for a new model



- DAG assumes a complete ordering of the time series components.
- In many cases only a partial ordering is available.
- A chain graph allows for partial orderings between time series components.

Motivation for a new model

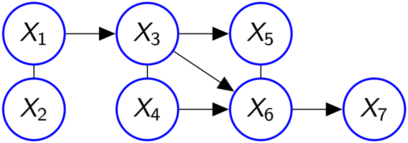


- DAG assumes a complete ordering of the time series components.
- In many cases only a partial ordering is available.
- A chain graph allows for partial orderings between time series components.

MDM ideas + chain graph = **dynamic chain graph model (DCGM)**

The dynamic chain graph model (DCGM)

Example chain graph:



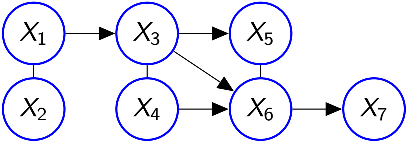
Chain components: $\{X_1, X_2\}$, $\{X_3, X_4\}$, $\{X_5, X_6\}$, $\{X_7\}$.

The DCGM:

- Each chain component modelled separately by (conditional) multivariate dynamic model.
- Each $Y_t(i)$ has its parents as regressors.

The dynamic chain graph model (DCGM)

Example chain graph:



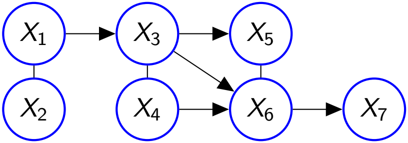
Chain components: $\{X_1, X_2\}$, $\{X_3, X_4\}$, $\{X_5, X_6\}$, $\{X_7\}$.

The DCGM:

- Each chain component modelled separately by (conditional) multivariate dynamic model.
- Each $Y_t(i)$ has its parents as regressors.
- n-dimensional problem broken into N separate multivariate problems of smaller dimensions.

The dynamic chain graph model (DCGM)

Example chain graph:



Chain components: $\{X_1, X_2\}$, $\{X_3, X_4\}$, $\{X_5, X_6\}$, $\{X_7\}$.

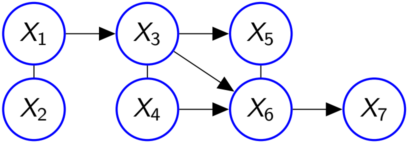
The DCGM:

- Each chain component modelled separately by (conditional) multivariate dynamic model.
- Each $Y_t(i)$ has its parents as regressors.
- n-dimensional problem broken into N separate multivariate problems of smaller dimensions.
- MDM is special case in which all chain components are single variables.

DCGM: Anacleto and Queen, OU Statistics Group Technical Report 13/02

The dynamic chain graph model (DCGM)

Example chain graph:



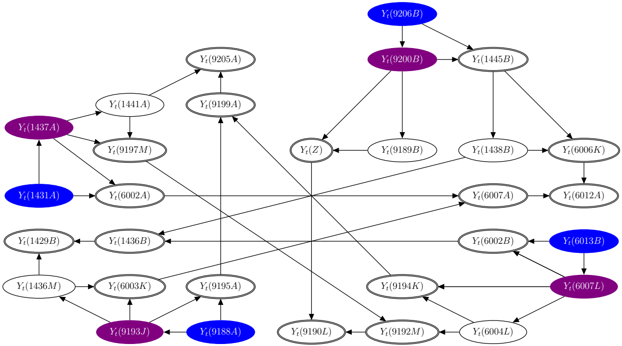
Chain components: $\{X_1, X_2\}$, $\{X_3, X_4\}$, $\{X_5, X_6\}$, $\{X_7\}$.

The DCGM:

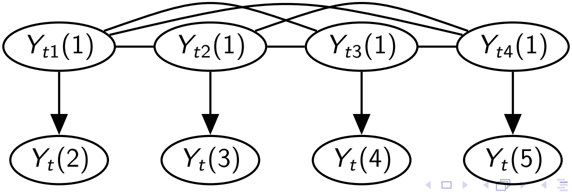
- Each chain component modelled separately by (conditional) multivariate dynamic model.
- Each $Y_t(i)$ has its parents as regressors.
- n-dimensional problem broken into N separate multivariate problems of smaller dimensions.
- MDM is special case in which all chain components are single variables.

DCGM: Anacleto and Queen, OU Statistics Group Technical Report 13/02

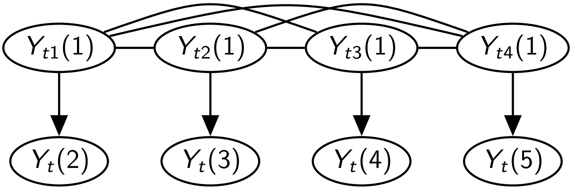
Chain graph for the Manchester network



Form chain graph of 4 root nodes and 1 child each (and relabel):



DCGM for subnetwork

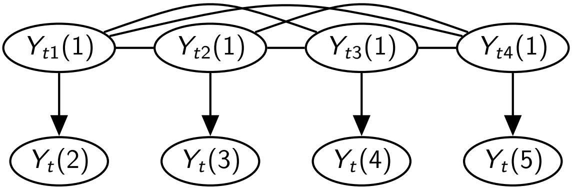


DCGM consists of 5 separate models:

- Multivariate DLM for $Y_{t1}(1), \ldots, Y_{t4}(1)$ (matrix-normal DLM),
- Univariate DLM for $Y_t(2)$ with $y_{t1}(1)$ as regressor,
- Univariate DLM for $Y_t(3)$ with $y_{t2}(1)$ as regressor,
- Univariate DLM for $Y_t(4)$ with $y_{t3}(1)$ as regressor,
- Univariate DLM for $Y_t(5)$ with $y_{t4}(1)$ as regressor.

Problem: Occupancy, Headway and Speed cannot be incorporated into model for $Y_{t1}(1), \ldots, Y_{t4}(1)$!

DCGM for subnetwork



DCGM consists of 5 separate models:

- Multivariate DLM for $Y_{t1}(1), \ldots, Y_{t4}(1)$ (matrix-normal DLM),
- Univariate DLM for $Y_t(2)$ with $y_{t1}(1)$ as regressor,
- Univariate DLM for $Y_t(3)$ with $y_{t2}(1)$ as regressor,
- Univariate DLM for $Y_t(4)$ with $y_{t3}(1)$ as regressor,
- Univariate DLM for $Y_t(5)$ with $y_{t4}(1)$ as regressor.

Problem: Occupancy, Headway and Speed cannot be incorporated into model for $Y_{t1}(1), \ldots, Y_{t4}(1)$!

Comparison: DCGM with LMDM

One-step forecasts obtained for flows in December 2010.

Big freeze returns: Drivers face new chaos as snow on way back

Richard Wheatstone
December 13, 2010

Submit Comments | Comments (25) | Printable Version   Previous | Next

AdChoices

Motorists may have to endure more snow later this week

Manchester Evening News headline, December 2010

DCGM gives better forecasts based on the joint log-predictive likelihood of the 8 series

Comparison: DCGM with LMDM

One-step forecasts obtained for flows in December 2010.


Big freeze returns: Drivers face new chaos as snow on way back

Richard Wheatstone
December 13, 2010

Submit Comments | Comments (25) | Printable Version   Previous | Next

Motorists may have to endure more snow later this week

AdChoices

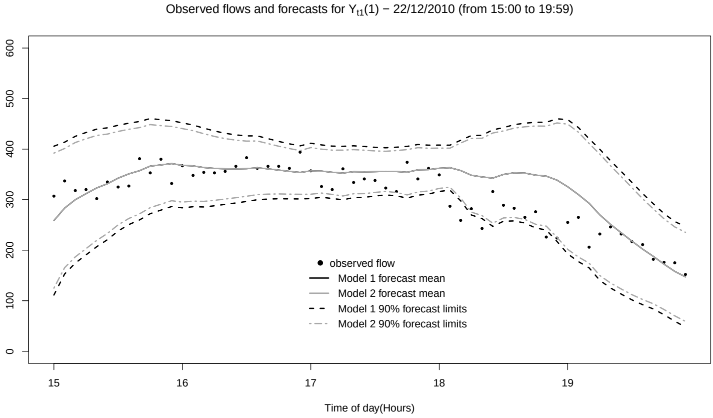
Manchester Data Centre
New Greater Manchester 1350m2 data centre. Book a tour.
www.zen.co.uk/datacentre

Solar Electric Systems
Professional and Competitive. Efficient Solar Energy Services.
rr-ltd.com

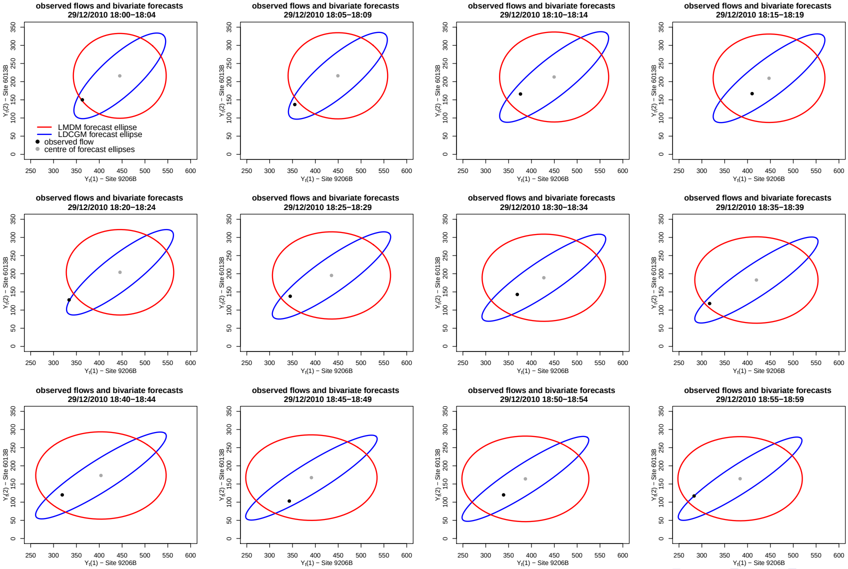
Manchester Evening News headline, December 2010

DCGM gives better forecasts based on the joint log-predictive likelihood of the 8 series

Marginal univariate forecasts



Marginal bivariate forecasts



Work in progress...

In traffic modelling: direction of the flow induces a graph for traffic networks.

But...

structural learning is a key problem in many applications.

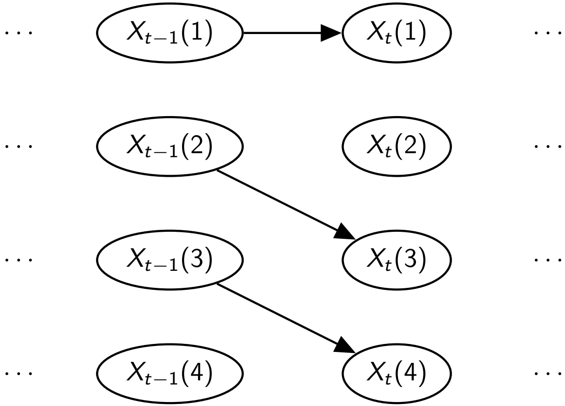
Example: gene expression modelling

- Nodes in the graph represent genes.
- Data are high-dimensional time series of gene expression levels ($n << p$).
- **Goal: estimate relationships between different genes.**

Work in progress...

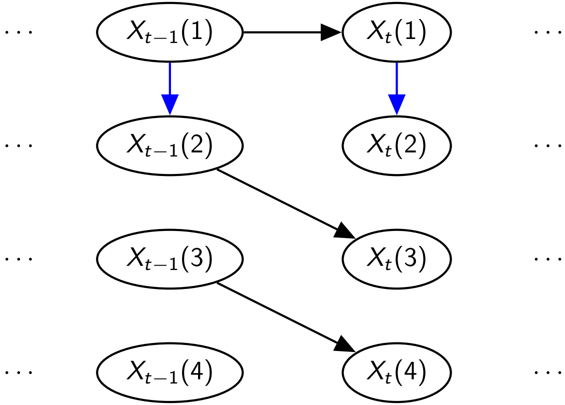
$\mathbf{X}_t = [X_t(1), X_t(2), X_t(3), X_t(4)]^\top$: vector of gene expression levels.

Dynamic Bayesian networks seem to be useful tool.



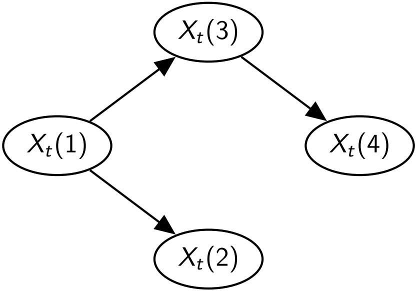
Work in progress...

Data observed on different time scales ⇒ contemporaneous relationships.



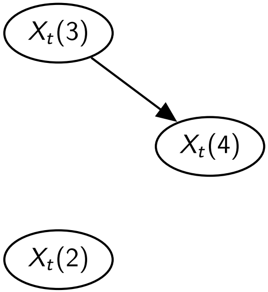
MDM possible model.

Work in progress...



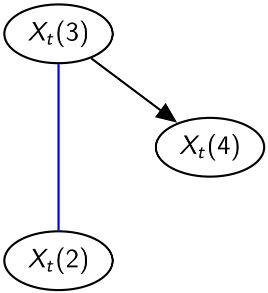
Gene $X_t(1)$ regulates $X_t(2)$ and $X_t(3)$.

Work in progress...



$X_t(1)$ may be missing from the data set.

Work in progress...



Missing common regulator can be accommodated with DCGM.

Summary

- Graphs can be used to simplify multivariate time series modelling.
- Multiregression dynamic model can provide on-line traffic forecasts — series represented by a DAG.
- Dynamic chain graphical models can improve forecasts by allowing partial orderings in a graph — series represented by a chain graph.
- Use of the MDM and DCGM for structural learning for gene expression data seems promising.

More details:

Anacleto, O., Queen, C.M. Dynamic chain graph models for multivariate time series (osvaldoanacleto.com/DCGmodel.pdf).

Anacleto, O., Queen, C.M, Albers, C.J. (2013). Multivariate forecasting of road traffic flows in the presence of heteroscedasticity and measurement errors. *JRSS-C 62(2), March 2013.*

Anacleto, O., Queen, C.M., Albers, C.J. (2013). Forecasting multivariate road traffic flows using Bayesian dynamic graphical models, splines and other traffic variables. *Australian & New Zealand Journal of Statistics, 62(2), June 2013.*