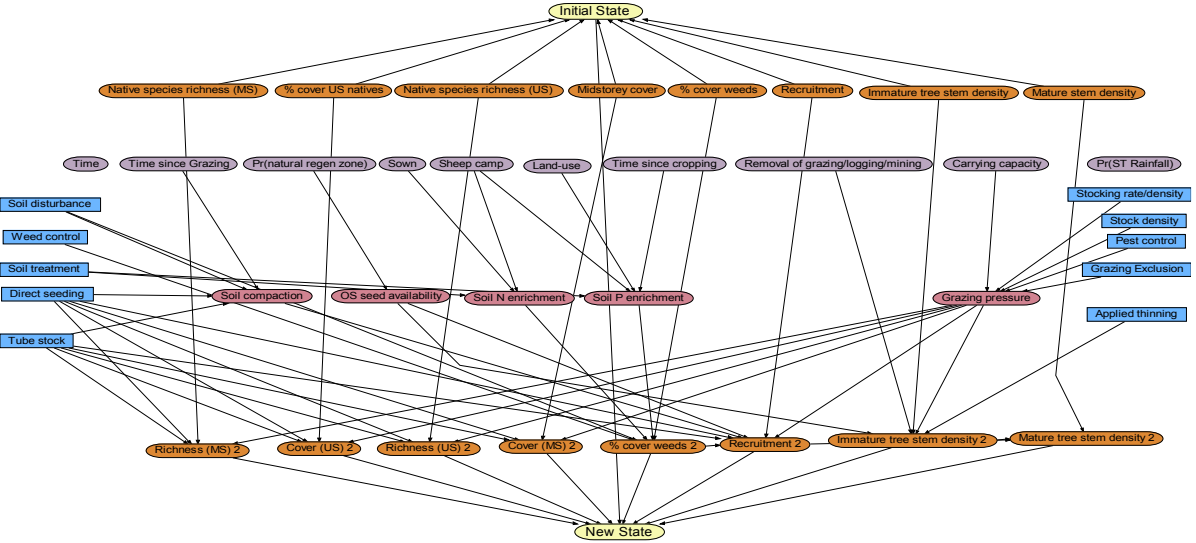
Spatialising a state-and-transition model for GIS integration

A modified woodland roadside in the Goulburn Broken catchment (Libby Rumpff)

Topics

- Background
- Work (in progress)
- Findings
- Discussion
- Future directions

Woodlands ST-BN



Woodlands ST-BN

- A site scale State-and-Transition BN for adaptive management
- No explicit spatial representation →
  - Can't explore management strategies at a large scale
- Alternative: model % of land in particular starting state?
  - Relies on experts/field workers to provide the spatial context “intuitively”

ST-BN + GIS?

Research question:

Can we overcome spatial limitation by

- coupling the State-and-Transition model to
- GIS
- via a object-oriented dynamic Bayesian network?

OOBNs

- Borrow terminology from Object-Oriented programming.
- Encourage Knowledge engineers (modellers) to think of models as made up of sub-models.
  - Focus on building each sub-model as a self contained model
  - Encourages re-use of sub-models
- Make developing and maintaining complex models more manageable (after learning curve at beginning!
- Help domain experts interact and understand the model by allowing them to drill down to the level they are interested in.

Method

Netica ST-BN

↓ (1. convert)

Hugin ST-OOBN (static) (2. regression testing)

↓ (3. convert)

ST-OODBN

+ (5. combine)

(4. develop) Prototype weed contagion model

+ (5. combine)

GIS layers

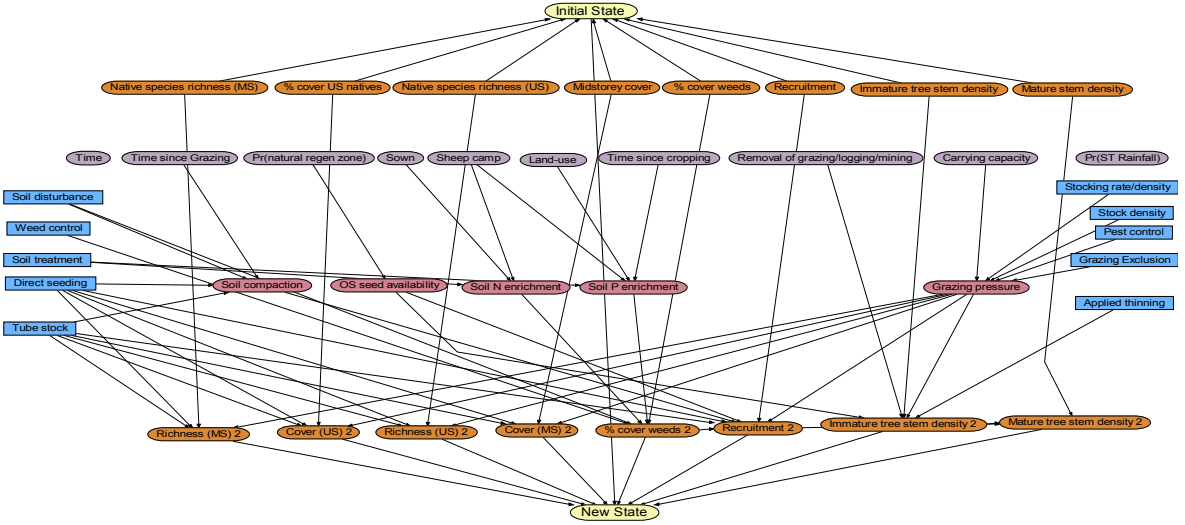
ST-BN to ST-OOBN

- Followed a 7 step process including understanding the intent of the model so the translation is faithful.
- Used regression testing to ensure that the output of each object was as close as possible as the original network fragment.
- Found bugs and inconsistencies in original ST-BN
- Small differences were found based on the sampling method used during the initial development.

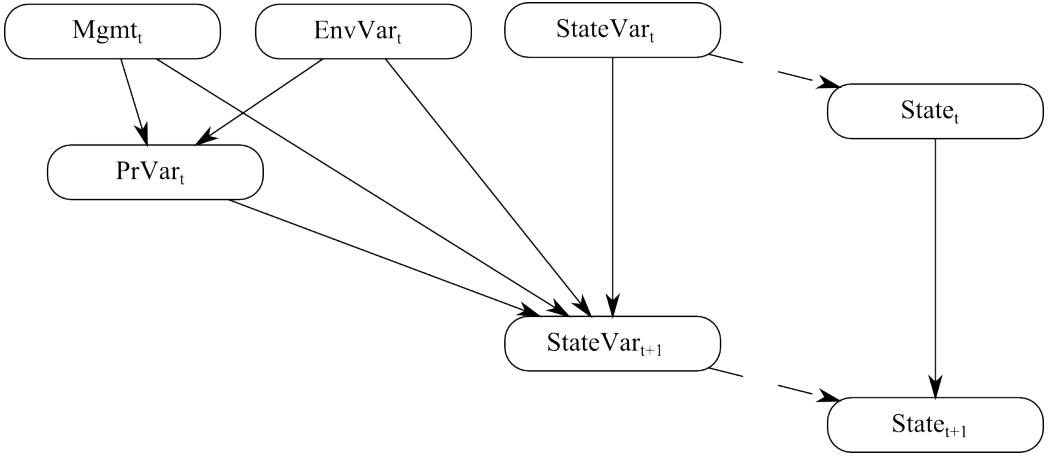
ST-OOBN to ST-OODBN

- Nicholson et al. (2011) showed that a static BN could be internally inconsistent when time as represented using a ranged node.
- Removing time and converting to an (OO)DBN encourages thinking about incremental causal changes over a time period.
- Longer time periods are modelled by rolling out the network by N time slices.
- Rolling out using a sliding window can keep computation feasible.

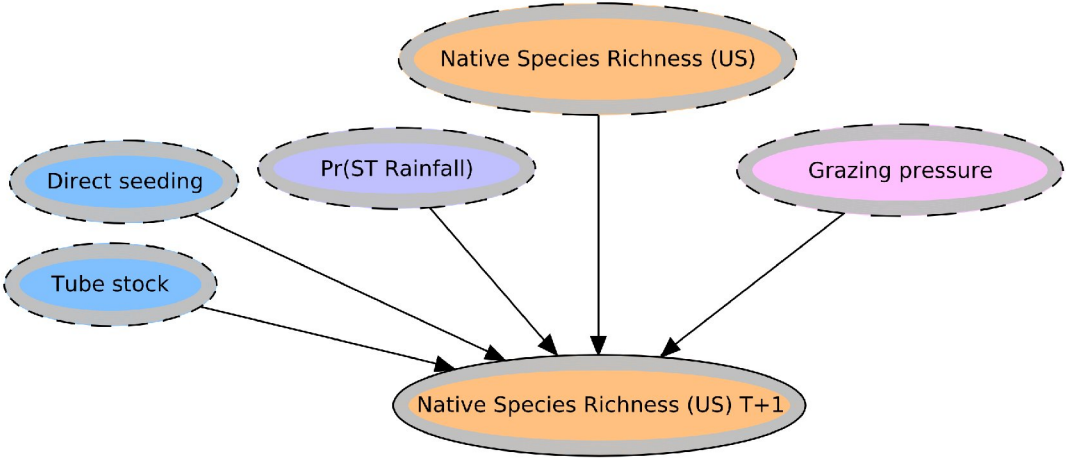
ST-BN (Before)


Initial State
Native species richness (MS)
% cover US natives
Native species richness (US)
Midstorey cover
% cover weeds
Recruitment
Immature tree stem density
Mature stem density
Time
Time since Grazing
Pr(natural regen zone)
Sown
Sheep camp
Land-use
Time since cropping
Removal of grazing/logging/mining
Carrying capacity
Pr(ST Rainfall)
Soil disturbance
Weed control
Soil treatment
Direct seeding
Tube stock
Soil compaction
OS seed availability
Soil N enrichment
Soil P enrichment
Grazing pressure
Stocking rate/density
Stock density
Pest control
Grazing Exclusion
Applied thinning
Richness (MS) 2
Cover (US) 2
Richness (US) 2
Cover (MS) 2
% cover weeds 2
Recruitment 2
Immature tree stem density 2
Mature tree stem density 2
New State

ST-OODBN (after)



Drill down to details



Map of Study Area

Not sure if I can show you this so I've left it out of this talk. Sorry.

Study Area

- Chosen because:
  - Live study area.
  - Independently negotiable management programs therefore neighbour effects may be large.
  - Good chance of some areas being in close to reference state.
- Inputs derived from native vegetation model (DSE unpublished data).
- Additional data - tree count data.
- Cell size (resolution) impacts on how you scale input values.

Prototype models

- Weed dispersions model
- Initial state
- No next state layers produced yet as we ran out of time to:
  - identify appropriate input layers for environment variables.
  - develop management scenarios.
  - elicit updated node values for OODBN
- The model does compile and run, however the layer results are meaningless at this stage.

Findings

- Converting the model to an ST-OODBN reduces clique table* size
  - Is the difference between being able to compile & run the model on commodity desktop PCs and not being able to.

Model	Total Clique Table Size	As percentage of original clique table
ST-BN	396808130	100%
ST-OOBN	619102853	156%
ST-OODBN (width 1)	1733360	0.4%
OODBN without State Object	10372	0.002%

*A clique table can be thought of representing all the possible combinations of inputs and outputs of a model.

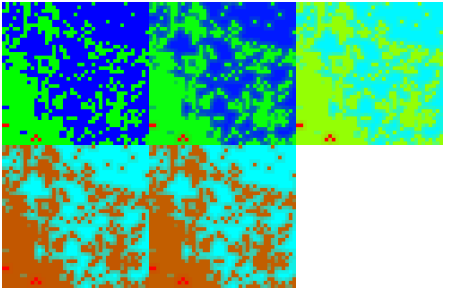
Findings

- Converting the model to an ST-OODBN reduces clique table size
  - A 10 time step sliding window (50 years) ST-OODBN is still an order of magnitude smaller than the equivalent ST-BN

Model	Total Clique Table Size	As percentage of original clique table
ST-BN	396808130	100%
ST-OOBN	619102853	156%
ST-OODBN (width 10)	1733360*10	40%
OODBN without State Object	10372*10	0.02%

Findings

- An ST-OODBN can model state transitions in a spatial context.
- No usable layer for whole state but weeds sub-model has produced some interesting layers.
- Shown: Weed model at t0, t+1, t+2 and t+10 showing a weak neighbour effect.



Discussion

- Converting the model to an ST-OODBN encourages good Knowledge Engineering practices
  - Model reuse: objects from one model could be used in another model, just change the parameterisation
  - Information hiding: Hide details about the internal workings of the model and concentrate on the interconnections of sub models
  - Modularity: focuses attention on sub-problems, models as part of a larger model
- Even if not interested in linking to GIS OODBN these points hold true.

Discussion

- Given the current structure how big could the model be scaled?
- If we had:
  - Input data
  - Models for different EVC groupings
  - Enough CPU cores
- Technically, most probably state-wide.
- It becomes an expert elicitation problem again.

Future Directions

- Conversion of structure suitable for distributed computing
- Improvements to caching algorithm
- Improvements to network structure to reduce clique table sizes
- Use of different BN libraries or platforms
- Investigating further BN variants to improve computational efficiency
- Integration with alternate models to provide hybrid models (e.g. BN+stochastic simulation)
- Extension of the model to cover multiple EVCs

Thanks also to...

- Dr Steve Sinclair for help with data and understanding the study area.
- Dr Peter Griffioen for help with getting me acquainted with just a small part of the huge amount of data the DSE has available.
- Matt White & Dr Graham Newell for helping me understand the domain

Lauchlin Wilkinson[1,3], Dr. Libby Rumpff[1], Dr. David Duncan[2] and Assoc. Prof Ann Nicholson[3]

[1] **School of Botany, The University of Melbourne**
Parkville Campus

[2] **The Arthur Rylah Institute for Environmental Research**
Department of Sustainability and Environment
Heidelberg VIC 3084

[3] **School of Information Technology, Monash University**
Clayton Campus