

Dynamic Bayesian Networks for Decision Support: UK Vegetable Supply Model & GUI for Model Outputs

Lee Sheng Long (23318325)
sllee26@student.monash.edu
s.lee.21@warwick.ac.uk
<https://github.com/slee21>

19th March 2017

Contents

1	Introduction	1	A.1	Arrival Report	i
2	UK Vegetable Supply Model	1	A.2	Progress Report: 09/01/2017 to 20/01/2017	i
2.1	Motivation	1	A.3	Progress Report: 23/01/2017 to 03/02/2017	ii
2.2	Results	1	A.4	Progress Report: 06/02/2017 to 17/02/2017	iii
2.3	Limitations	1	A.5	Departure Report	iii
2.4	Future Work	1	B	Netica Networks	iv
3	GUI for Model Outputs	1	B.1	UK Vegetable Supply BN prior to 26/01/2017 meeting .	iv
3.1	Motivation	1	B.2	Revised UK Vegetable Supply BN prior to 15/02/2017 meeting	v
3.2	Results	1	B.2.1	Netica network	v
3.3	Limitations	2	B.2.2	Node and state defini- tions	v
3.4	Future Work	2	B.2.3	Conditional Probabil- ity Tables (CPTs) . . .	viii
4	Key Takeaways	2	C	Graphical User Interface (GUI)	
4.1	Insight into Academia	2		Screenshots	ix
4.2	Research Experience	2	C.1	GUI screenshots prior to 16/02/2017 meeting	ix
4.3	Development Experience	2	C.2	GUI screenshots after departure	x
5	Conclusion	2			
	Appendices	i			
A	Progress Reports	i			

1 Introduction

I spent 8 weeks in January and February of 2017 doing research projects under supervision by Dr. Martine J. Barons, Professor Ann E. Nicholson and Professor Jim Q. Smith.

The first project was to model UK vegetable supply with Bayesian Networks. It was more research-based.

The second project was to develop a simple Graphical User Interface (GUI) for model outputs. It was more application-based.

2 UK Vegetable Supply Model

2.1 Motivation

The recent UK vegetable shortage illustrated the utility and importance of the project (British Broadcasting Corporation, 2017).

Such a model could be useful to discover key factors behind the shortage and guide corrective measures to alleviate the shortage and prevent further recurrences. It could also predict price movements allowing time to prepare.

The resultant UK vegetable supply network would be integrated into a larger network based on the CPI food basket to model overall food supply in the UK.

It would ultimately contribute to a Decision Support System (DSS) for policy makers in the area of food security.

2.2 Results

The resultant partially parametrised Bayesian Network is described in B.2.1 and is available for download at <https://github.com/slee21/murpa>.

2.3 Limitations

Time constraints resulted in some limitations:

- Network is only partially parametrised with Conditional Probability Tables (CPT)
- Expert elicitations only involved 1 expert, Rob Lillywhite of Warwick Crop Centre

2.4 Future Work

Potential productive avenues for future work:

- Fully parametrize the network with CPTs
- Expert elicitations with more experts
- Sensitivity analysis on the network

3 GUI for Model Outputs

3.1 Motivation

The Netica software package interface is targeted at practitioners familiar with Bayesian Networks, statistics and probability.

Most policy makers likely lack the requisite knowledge and would be confused operating the Netica interface. Thus, a new interface targeted at policy makers would be useful.

The resultant GUI would provide a user-friendly interface for display and interaction with the model outputs aimed at policy makers with little background in statistics.

3.2 Results

The resultant frontend GUI is described in B.2.1 and is available for download at <https://github.com/slee21/bayeslayers>. The backend JSON API server is available for download at <https://github.com/slee21/gonetica>.

3.3 Limitations

Time constraints resulted in some limitations:

- Saving data to and loading data from local files remains unimplemented
- Interface look and feel is rudimentary and unpolished
- Hardcoded or unconfigurable options and settings

3.4 Future Work

Potential productive avenues for future work:

- Implement local file saving and loading
- Implement server side saving and loading
- Redesign and polish interface look and feel
- Expand configurable options and settings
- Implement distributed collaboration
- Redesign and optimize transfer protocol

4 Key Takeaways

4.1 Insight into Academia

I had the opportunity to speak with non-teaching academics which would otherwise be difficult as an undergraduate. I also met PhD candidates and other postgraduate students. They provided different perspectives of academia.

4.2 Research Experience

I learned the dreadful practicality of Bayesian Networks. They were particularly useful when data was scarce. This was readily apparent in the modeling of the UK vegetable supply where often times the only sources were human wisdom.

4.3 Development Experience

I experienced the difficulties of working with a closed-source proprietary API (Netica API). I also gained the experience of developing a simple JSON API server and accompanying single page web application client.

5 Conclusion

MURPA was a valuable learning experience in many ways. It helped settle most of my doubts or worries about continuing on towards honours.

I want to thank my supervisors Dr. Martine J. Barons, Professor Ann E. Nicholson and Professor Jim Q. Smith for their kind guidance.

I also want to thank Dr. Arun Konagurthu and Caitlin Slattery for enabling and supporting the MURPA program.

Finally, I would like to apologise for any shortcomings on my part and wish any successor MURPA students reading this all the best with your projects.

References

- British Broadcasting Corporation. (2016, April 28). Banned pesticides 'not equally harmful' to bees. Retrieved January 5, 2017, from <http://www.bbc.com/news/science-environment-36154134>
- British Broadcasting Corporation. (2017, February 3). Why is there a vegetable shortage? Retrieved March 5, 2017, from <http://www.bbc.com/news/uk-38666752>
- Cain, J. (2001). *Planning improvements in natural resource management. guidelines for using bayesian networks to support the planning and management of development programmes in the water sector and beyond*. Wallingford: Centre for Ecology and Hydrology. Retrieved from <http://nora.nerc.ac.uk/9461/>
- Department for Environment, Food & Rural Affairs. (2016a, May 26). Agriculture in the united kingdom 2015. Retrieved January 27, 2017, from <http://www.gov.uk/government/statistics/agriculture-in-the-united-kingdom-2015>
- Department for Environment, Food & Rural Affairs. (2016b, July 22). Horticulture statistics 2015. Retrieved January 27, 2017, from <http://www.gov.uk/government/statistics/horticulture-statistics-2015>
- Nix, J. (2015). *John nix farm management pocketbook 2016*. Great Britain: Agro Business Consultants.
- Office for National Statistics. (2016, March 15). Consumer price inflation basket of goods and services: 2016. Retrieved January 5, 2017, from <http://www.ons.gov.uk/economy/inflationandpriceindices/articles/consumerpriceinflationbasketofgoodsandservices/2016>

Appendices

A Progress Reports

A.1 Arrival Report

First week was spent settling in. Check-in to accommodation, ID collection, visa confirmation and work space allocation went smoothly.

First meeting with supervisor Dr. Martine J. Barons determined the projects for the 8 week duration at University of Warwick:

- Model UK vegetable supply with Dynamic Bayesian Networks (DBN) in Netica
- Develop Graphical User Interface (GUI) to interact with and display model outputs

Expert meetings for the projects were scheduled:

- 12th January 2017
Meeting with Andy Davis of Warwickshire County Council to discuss the intended users and use cases of the GUI
- 26th January 2017
Meeting with Rob Lillywhite of Warwick Crop Centre to discuss the UK vegetable supply Bayesian Network structure

A.2 Progress Report: 09/01/2017 to 20/01/2017

The UK vegetable supply model was put on the backburner for most of the 2 week period. It was mainly due to the timing of the expert meetings.

Prior to meeting with Andy Davis, several broad possibilities were identified for the GUI:

- (A) Bespoke native application
- (B) Application plug-in or extension
- (C) In-browser web application

Of the three options, I most favoured (C) as it targets any platform with a web browser without much additional effort.

After meeting with Andy Davis, we decided to go with option (C) as it was the most similar to what was in use internally at the Warwickshire City Council¹. Familiarity reduces barriers to software adoption.

Further, a key element of the proposed GUI would be a geographical map to visualise model outputs at various governmental scales - region, county, district and so on. The niche is well-established with giants such as Google Maps.

I tentatively decided to use the same major components as the Warwickshire City Council². The intention was to reduce the maintenance burden as far as reasonably possible. It also helps reduce the scope of development.

However, no existing glue code between the modelling software and the GUI software stack was found. It was necessary to develop custom software using the available Netica APIs³. Of these, only the Java and C APIs were pragmatic.

The Java and C APIs were learned and test programs written sometimes with wrappers for comparison purposes. The programs performed batch inference given input Bayesian Network and cases. Interactive versions would follow after.

Scattered time was spent reading literature on UK vegetable supply.

¹<http://www.warwickshire.gov.uk/gis>

²<http://www.ordnancesurvey.co.uk/business-and-government/case-studies/warwickshire-county-council-new-web-gis.html>

³http://www.norsys.com/netica_api.html

A.3 Progress Report: 23/01/2017 to 03/02/2017

An initial Bayesian Network B.1 was drawn up in Netica in preparation for the meeting with Rob Lillywhite.

It follows the recommendations contained in the MURPA 2016 report by Joshua Collins. The final fruit network by Sophia Wright was used as a reference because fruit was the most similar of the mapped Consumer Price Index (CPI) classes.

After the meeting with Rob Lillywhite, it was clear that the initial model needed more precise definition and further adjustment.

First, vegetables as defined by the UK CPI basket of goods and services⁴ include a variety of crops with different properties. For example, some vegetables may be stored whereas others deteriorate quickly after harvesting.

Therefore, it may be appropriate to focus the model on a subset of exemplar vegetables that have similar properties. Also, not all vegetables may be relevant from a food security standpoint. Expensive luxury items may be excluded.

Following are comments on the structure of the initial model – the nodes and edges of the Bayesian Network.

The production cost of vegetables does not significantly affect the vegetable production volume. The production cost is passed on to the buyer by increasing the selling price. Thus, it affects vegetable price instead of production.

Weather has a widespread impact and does not only affect vegetable production. It likely affects more variables than in the initial model. Additionally, world climate change is not the sole determinant of UK weather conditions.

Pollinators are beneficial to some vegetables but their impact on vegetables in aggregate is not significant. Similarly, organic methods yield significantly less produce but are a niche market likely irrelevant to food security.

Biofuel production has no significant impact on vegetable production. vegetables are not used for biofuel unlike cereals. Vegetables are also quite profitable and land for planting vegetables is unlikely to be repurposed for biofuels.

Vegetables are not attractive for research and development (R&D) spending. The market is too small. Instead most research is done for arable crops as a whole. Even then, most R&D spent tends to have impact only after 10 years or more.

Crop disease and pests are only a concern if available mitigation tools such as pesticides are limited. Government regulation may prohibit the use of certain active ingredients in pesticides. An example of this would be the ban on neonics⁵.

Absent in the initial model is the significant waste (up to 30%) that occurs in the supply chain between farm production and retailer supply.

Retailer or more specifically supermarket policy also has a significant impact on the supply of vegetables in the UK.

Labour cost and availability is also a concern. Agriculture is reliant on migrant labour (up to 90%) which is determined by government policy.

I decided to use Go⁶ to wrap the Netica C API and develop the backend for the GUI. The reasons were mostly pragmatic.

Although I can write scripts in Java and C, I am not confident I can deliver an application in the available time. I have more experience in Go and the compiled binaries are easy to maintain. Go is also a relatively simple language.

⁴Office for National Statistics, 2016

⁵British Broadcasting Corporation, 2016

⁶<http://golang.org/cmd/cgo/>

A.4 Progress Report: 06/02/2017 to 17/02/2017

Expert meetings for the projects were scheduled:

- 15th February 2017
Meeting with Rob Lillywhite to discuss the revised Bayesian Network structure and parameters
- 16th February 2017
Meeting with Andy Davis to demonstrate the GUI prototype and discuss issues that arose during development of the prototype

A revised Bayesian Network B.2.1 was drawn up after the first meeting with Rob Lillywhite. It attempts to incorporate the points discussed during the meeting. The nodes and states were specified in more detail in a table B.2.2.

The meeting confirmed that the structure of the revised network was reasonable. Rob also suggested more appropriate information sources for some of the nodes. Finally, he specified the Conditional Probability Table (CPT) for the “UK Vegetable Price” node B.2.3.

It was clarified that DEFRA horticulture production statistics ⁷ only account for “UK marketed produce” which excludes supply chain waste. Thus, the “UK Vegetable Production” node required a different information source that included farm waste in its statistics.

DEFRA agriculture statistics ⁸ only provide total production cost figures. It was suggested that vegetable specific production costs could be estimated based on cost per yield figures from literature. The “John Nix Farm Management Pocketbook”⁹ was suggested as a source.

Kantar¹⁰ was suggested as an authoritative information source for “UK Vegetable Demand” and “UK Retailer Policy” nodes.

Fresh Produce Journal ¹¹ was suggested as an information source for “UK Vegetable Price” and “World Vegetable Price” nodes.

The CPTs for the network were elicited. However, time constraints resulted in only the partial CPT for the central “UK Vegetable Price” node. The remaining entries in the CPT may be interpolated¹² from the entries available in the partial CPT.

Prior to second the meeting with Andy Davis, a rough prototype GUI C.1 was developed to provide a visual frontend for the Go command line interface (CLI) backend. The GUI was a single page webapp using HTML and CSS for display and Javascript for logic.

There was a tentatively positive response to the rough prototype GUI at the meeting. Minor tweaks were suggested before a final meeting was planned for the departure week. The desired map boundaries were confirmed. Also, inference was confirmed to only produce expected values.

A.5 Departure Report

Unfortunately, progress on the GUI was blocked by some bugs. There was no significant progress to show at the final meeting.

The rest of the week was spent preparing for departure and attempting to debug the GUI Javascript frontend. Accommodation check-out went without issue. Final report was written-up.

⁷Department for Environment, Food & Rural Affairs, 2016b

⁸Department for Environment, Food & Rural Affairs, 2016a

⁹Nix, 2015

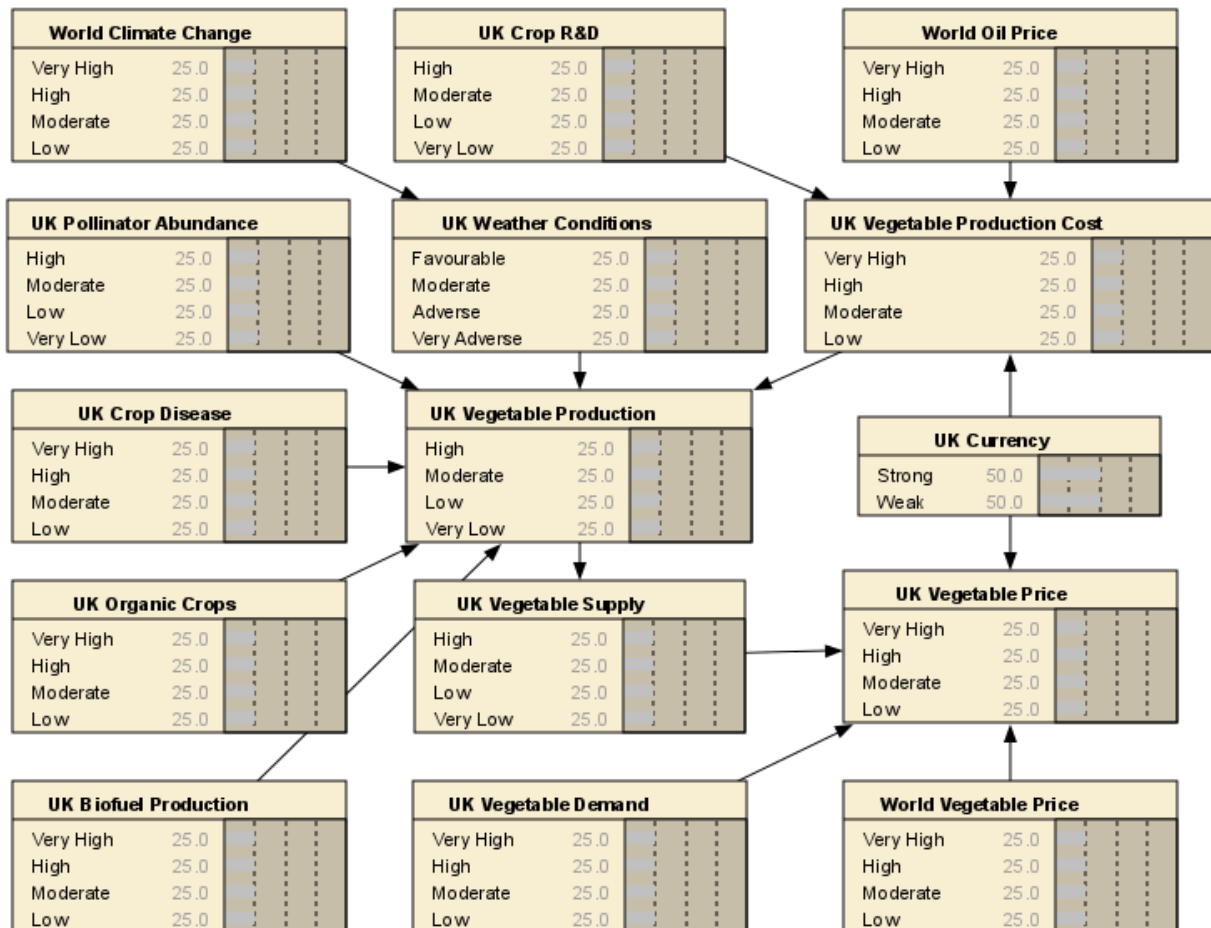
¹⁰<http://www.kantar.com/>

¹¹<http://www.fruitnet.com/fpj>

¹²Cain, 2001

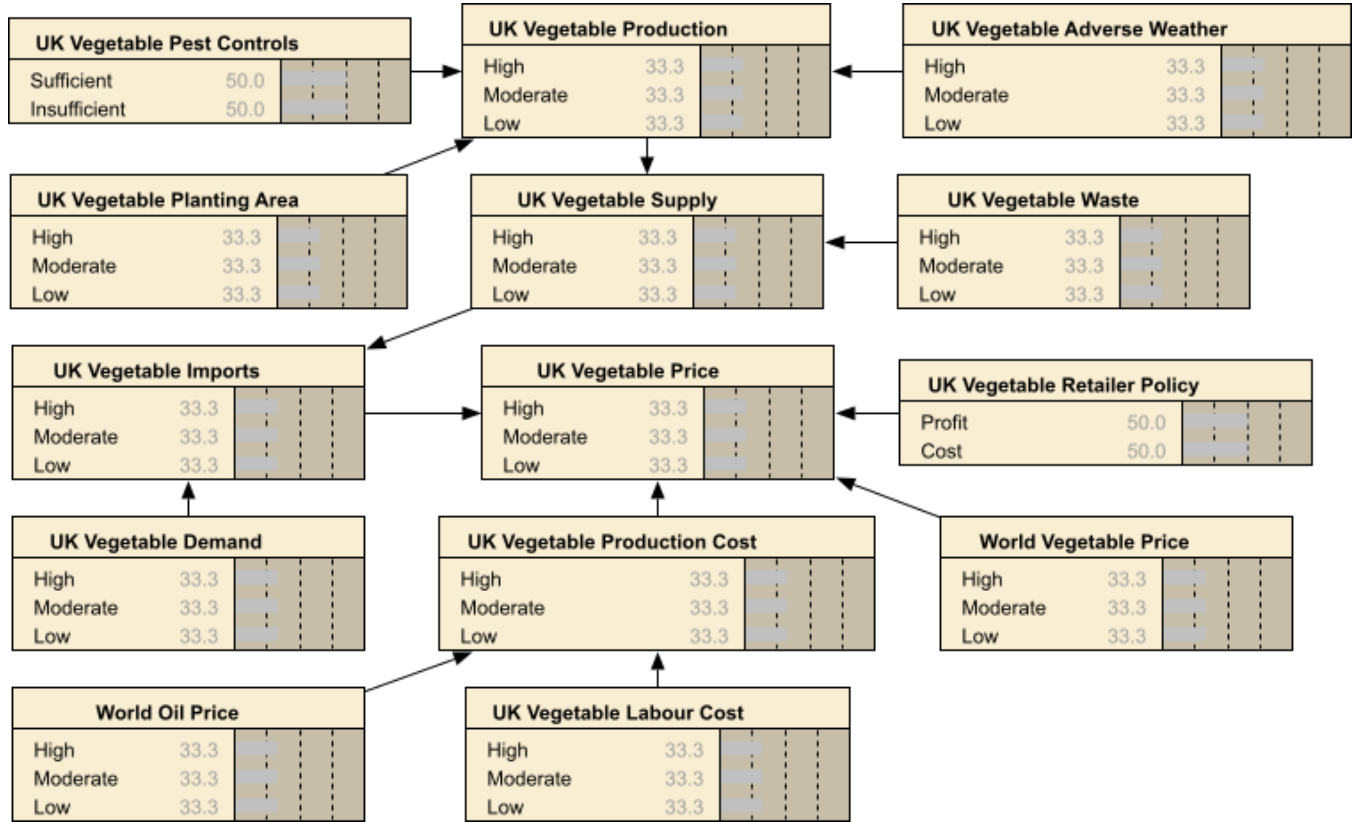
B Netica Networks

B.1 UK Vegetable Supply BN prior to 26/01/2017 meeting



B.2 Revised UK Vegetable Supply BN prior to 15/02/2017 meeting

B.2.1 Netica network



B.2.2 Node and state definitions

Node Details		State Details	
UK Vegetable Pest Controls	Availability of pest controls such as pesticides.	Sufficient	Any banned control has reasonable alternatives.
		Insufficient	Exists banned controls without any reasonable alternatives.
Parents	None		
Info sources	?		
UK Vegetable Production	Total production of vegetables in tonnes.	High Moderate Low	% change in total vegetable production relative to previous period.
Parents	UK Vegetable Pest Controls, UK Weather Conditions, UK Vegetable Planting Area		
Info sources	Horticulture Statistics by DEFRA		

Node Details		State Details	
UK Vegetable Adverse Weather	Risk of adverse weather affecting vegetable crop such as floods.	High Moderate Low	Average % risk of adverse weather over the time period weighted by planting area affected.
Parents	None		
Info sources	Met Office		
UK Vegetable Planting Area	Total land area used for planting vegetables in hectares.	High Moderate Low	% change in land area used for vegetables relative to previous period.
Parents	None		
Info sources	Horticulture Statistics by DEFRA		
UK Vegetable Supply	Total domestic supply of vegetables in tonnes.	High Moderate Low	% change in total vegetable supply less imports relative to previous period.
Parents	UK Vegetable Production, UK Vegetable Waste		
Info sources	Horticulture Statistics by DEFRA		
UK Vegetable Waste	Produce lost to waste along the supply chain before reaching end consumer.	High Moderate Low	% proportion of vegetable waste along the supply chain before reaching end consumer over total vegetable production for the period.
Parents	None		
Info sources	Quantification of food surplus waste and related materials in the supply chain by WRAP		
UK Vegetable Imports	Total imports of vegetables in tonnes.	High Moderate Low	% change in imported vegetable supply relative to previous period.
Parents	UK Vegetable Supply, UK Vegetable Demand		
Info sources	Horticulture Statistics by DEFRA		
UK Vegetable Price	Domestic retail price of vegetables to end consumer in £.	High Moderate Low	% change in Consumer Price Index (CPI) for the vegetables CPI class in the CPI basket of goods relative to previous period.
Parents	UK Vegetable Imports, UK Retailer Policy, UK Vegetable Production Cost, World Vegetable Price		
Info sources	UK Consumer Price Inflation by ONS		

Node Details		State Details	
UK Vegetable Retailer Policy	Policy or treatment of vegetables as a product by retailers such as supermarkets.	Profit	Retailer treats vegetables as products to profit from directly.
		Cost	Retailer treats vegetables as a cost to attract consumers so they purchase other products and profit indirectly.
Parents	None		
Info sources	?		
UK Vegetable Demand	Total domestic demand of vegetables in tonnes.	High	% change in total vegetable supply relative to previous period.
		Moderate	ASSUMPTION:
		Low	Total demand = Total supply
Parents	None		
Info sources	?Horticulture Statistics by DEFRA		
UK Vegetable Production Cost	Total domestic production cost of vegetables in £.	High	% change in total production cost of vegetables relative to previous period.
		Moderate	
		Low	
Parents	None		
Info sources	?Agriculture in the United Kingdom by DEFRA		
World Vegetable Price	Global price of vegetables adjusted to £.	High	% change in world price of vegetables relative to previous period.
		Moderate	ASSUMPTION:
		Low	World price = Import price
Parents	None		
Info sources	?Horticulture Statistics by DEFRA		
World Oil Price	Global price of a barrel of crude oil adjusted to £.	High	% change in average crude oil price over the period relative to previous period average.
		Moderate	ASSUMPTION:
		Low	Crude oil price is representative of world oil price.
Parents	None		
Info sources	Crude Oil Prices at NASDAQ		
UK Vegetable Labour Cost	Total labour cost of domestic vegetable production in £.	High	% change in total labour cost of domestic vegetable production relative to previous period.
		Moderate	
		Low	
Parents	None		
Info sources	?Agriculture in the United Kingdom by DEFRA		

B.2.3 Conditional Probability Tables (CPTs)

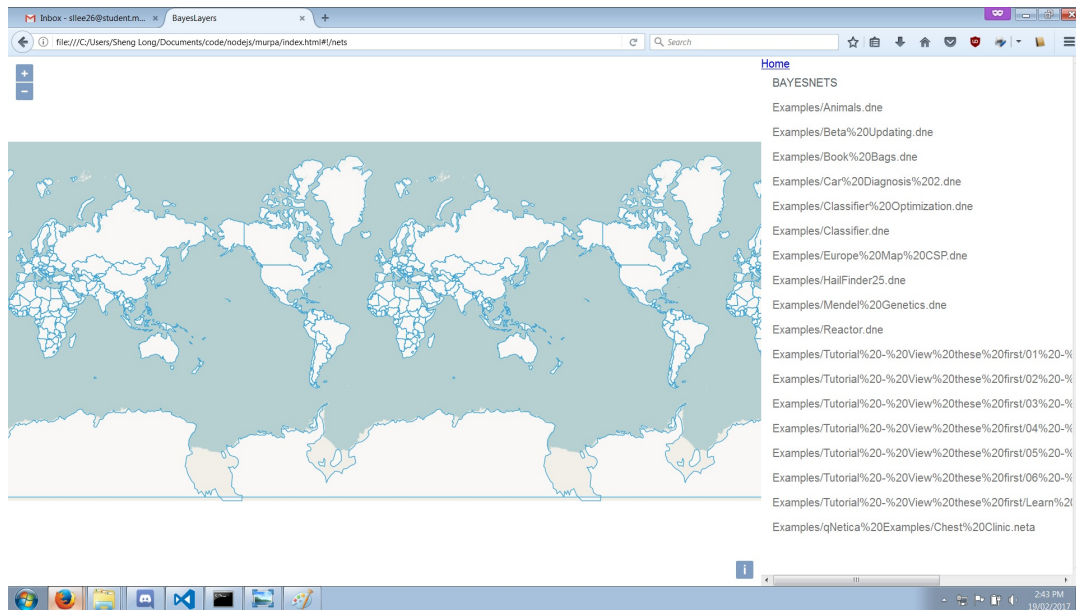
Transcribed below is the partial CPT for the “UK Vegetable Price” node elicited from Rob Lillywhite before interpolation.

Imports	Parents			UK Price		
	Retailer Policy	Production Cost	World Price	High	Moderate	Low
High	Profit	High	High	90%	8%	2%
High	Profit	High	Moderate	80%	15%	5%
High	Profit	High	Low	50%	25%	25%
High	Profit	Moderate	High	90%	8%	2%
High	Profit	Moderate	Moderate	80%	15%	5%
High	Profit	Moderate	Low	70%	20%	10%
Moderate	Profit	High	High	90%	8%	2%
Low	Profit	High	High	70%	20%	10%
High	Cost	High	High	80%	15%	5%
High	Cost	High	Moderate	70%	20%	10%
High	Cost	High	Low	60%	30%	10%
Low	Cost	Low	Low	30%	40%	30%

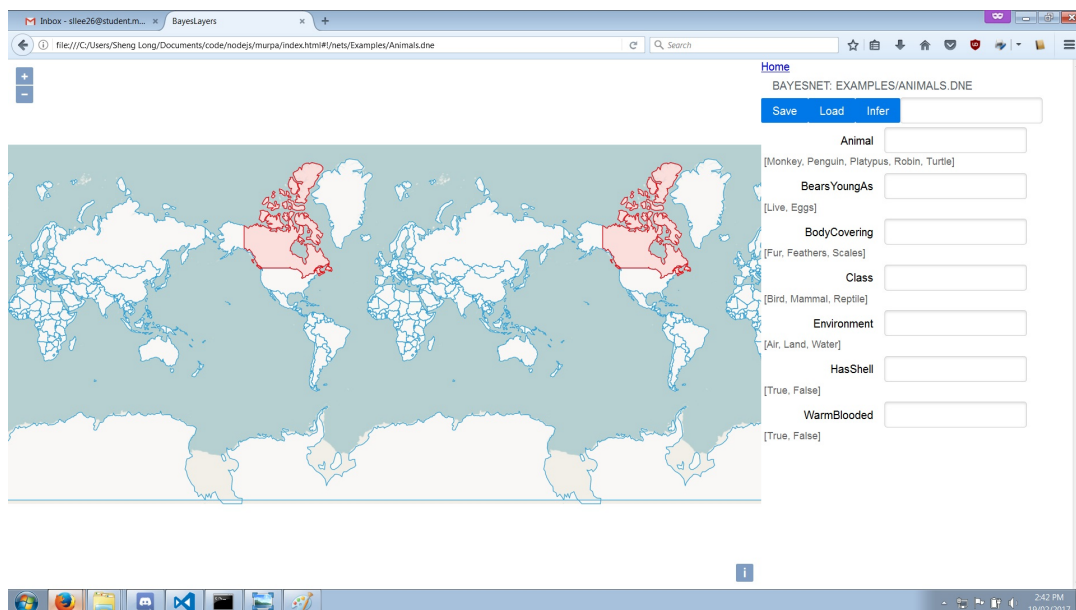
C Graphical User Interface (GUI) Screenshots

C.1 GUI screenshots prior to 16/02/2017 meeting

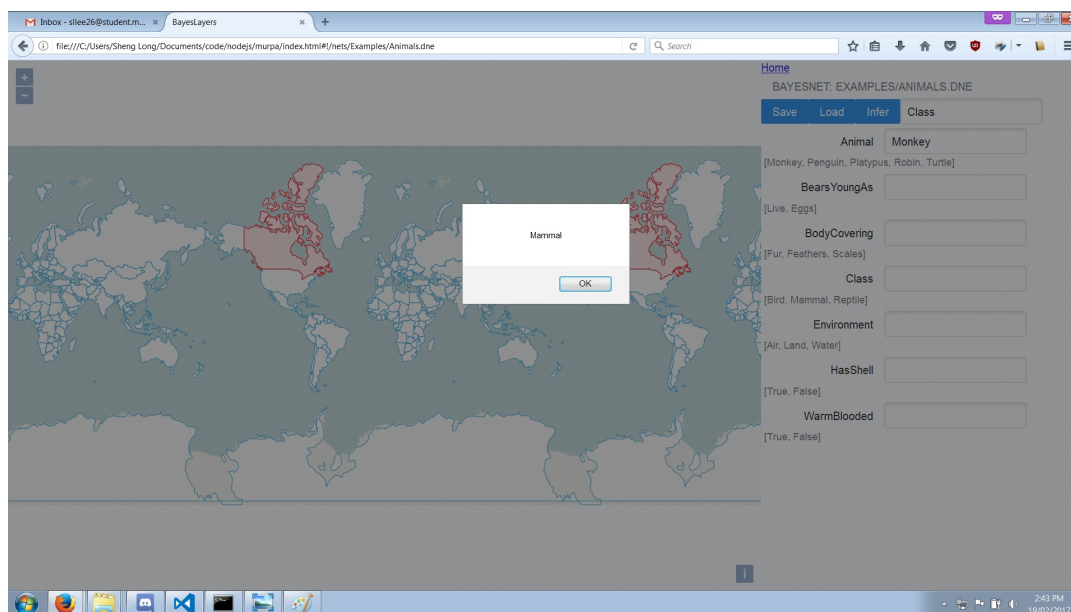
The user may select from the list of available Bayesian Networks.



Once a Bayesian Network is selected, the user may modify the node states or values associated with regions on the map.



The user may then get the result of Bayesian inference of a desired node for the map region with entered node states and values.



C.2 GUI screenshots after departure

The text inputs have been replaced with dropdown menus where it made sense.

