Understanding and predicting the influence of animal movement on the spread of transboundary animal diseases

Ben Madin

BSc (Vet Biol) BVMS MVPHMgmt

This thesis is presented for the degree of Doctor of Philosophy of Murdoch University

2011
Declaration

I declare that this thesis is my own account of my research and contains as its main content work which has not previously been submitted for a degree at any tertiary education institution.

Ben Madin

13 October 2011
Abstract

The aim of this thesis is to evaluate the potential of using existing and new data on disease outbreaks, livestock movements and prices to predict where outbreaks may occur.

An evaluation of the information on disease outbreaks stored in the regional animal health database was undertaken to determine if any relationship could be seen between outbreak locations over time and whether it would be possible to identify disease outbreaks early enough that they could be used in the prediction of disease spread. This work showed that disease reporting is incomplete and inconsistent, making it vital that increased effort is put into better outbreak investigation (including laboratory confirmation) and more timely reporting.

Information on the movement of animals through Cambodia and Laos was obtained to investigate patterns of movement. These data were incomplete, however application of network analysis techniques offered important insights into high risk areas for disease management. The full potential of this approach was established by applying it to Western Australian data from the National Livestock Information System.

Data on price were obtained in Cambodia and Laos to compare with the known movements to identify whether this could be used to predict animal movements. To overcome the complexity of collecting comprehensive data about different classes of animals, multilevel modelling was used to investigate the association of livestock movement with price difference between provinces.
Although the control of transboundary animal diseases is critically important for the economy of South East Asia, at the moment it is ineffective. It is unlikely that regional disease control programmes such as the South East Asia and China foot and mouth disease campaign will reach their potential until reliable, high-quality reports of disease are available to guide their design and implementation. Adding to the difficulty of this is the very sparse information available on the movement of animals across the region and the difficulty and expense involved in obtaining these data. Differences in market price between provinces may have a role in predicting animal movements; however, matching price to movements is difficult, and in this study appeared to only be useful over relatively short distances by regional standards.

Lack of a reliable means of identifying individual animals, regulatory and financial disincentives for using formal pathways for animal movement and a disregard for complying with government requirements to advise of animal movements make it difficult to follow animals as they move rapidly across the region. A new approach to animal movement management is required in which hazard reduction instead of revenue collection is the focus.
A Note on Style for Foot and Mouth Disease

There is little agreement on the correct syntax for foot and mouth disease (FMD). For this thesis I have adopted the same approach as the World Organisation for Animal Health (OIE), that is to use lower case for all words (except where the expression begins a sentence or is in a title) and not to hyphenate the first three words. This is consistent with the American Psychological Association Publication Manual (6th edition) recommendation of not using hyphens when the meaning is established\(^1\).

Contents

Abstract vi
A Note on Style for Foot and Mouth Disease xii
Table of Contents xiv
Tables xvii
List of Figures xix
Approvals xx
Animal Ethics Approval xix
FAO Statistics Division Approval xix
SaTScan xix
Glossary xx
Acknowledgements xxv

1 General Introduction 1
1.1 Introduction 3
1.2 Outline of Approach 4

2 Review of Relevant Literature 7
2.1 Livestock Production in a Rapidly Changing Society 9
2.2 Foot and Mouth Disease 13
2.2.1 Foot and Mouth Disease Virus 13
2.2.2 Vaccination 15
2.2.3 Impact 16
2.2.4 Spread of the Disease 18
2.2.5 International Context 21
2.3 Resources for Animal Disease Reporting 22
2.4 Sources of Outbreak Data 24
2.5 Livestock Movement 28
2.5.1 Comparison of Euclidean and Road Distances 31
2.6 Livestock Prices 33

3 Foot and Mouth Disease Reporting 37
3.1 Introduction 39
3.2 Materials and Methods 41
  3.2.1 Study Population 41
  3.2.2 Outbreak Data 41
  3.2.3 Phylogenetic Analysis 43
  3.2.4 Temporal Analysis 43
  3.2.5 Spatial Analysis 44
  3.2.6 Spatio-temporal analysis 46
3.3 Results 48
3.4 Discussion 58

4 Understanding Livestock Movement Patterns 63
4.1 Introduction 65
4.2 Materials and Methods 67
  4.2.1 Obtaining Movement Data 67
  4.2.2 Administrative Data 72
  4.2.3 Road Network Data 76
  4.2.4 Statistical Analysis 78
  4.2.5 Network Analysis 79
  4.2.6 A Quick Technology Summary 87
4.3 Results 88
  4.3.1 Western Australia 88
  4.3.2 Greater Mekong Subregion 98
4.4 Discussion 106
  4.4.1 Western Australia 106
  4.4.2 Greater Mekong Subregion 110
  4.4.3 Summary 113

5 Livestock Market Price 117
5.1 Introduction 119
  5.1.1 Overview of methods 120
5.2 Establishing Standard Price Categories 121
  5.2.1 Introduction 121
  5.2.2 Methodology 122
  5.2.3 Results—Consultation 123
  5.2.4 Results—Expert Opinion 126
  5.2.5 Discussion 129
CONTENTS

B Livestock Tracking and Tracing 209
B.1 Current Status ........................................ 211
B.1.1 Tracking Options ................................... 211
B.1.2 Tagging/Banding .................................... 212
B.1.3 Satellite .............................................. 212
B.1.4 Biometric Analysis .................................. 213
B.1.5 Genetic Fingerprinting ............................... 213
B.1.6 Better Resolution—Global Positioning System ....... 214
B.2 GSM Rumen Bolus ...................................... 215
B.2.1 GSM Rumen Bolus Concept .......................... 215
B.2.2 Rumen Boluses ....................................... 216
B.2.3 Core Design Requirements ............................ 217
B.2.4 Bolus Physical Limitations ............................ 219
B.3 Signal Strength Testing ................................. 221
B.3.1 Method ................................................ 221
B.3.2 Results ............................................... 224
B.3.3 Conclusions .......................................... 225

C Reporting Market Prices by Text Message 227
C.1 Sending Text Messages .................................. 229
C.1.1 Project Overview .................................... 229
C.1.2 The Importance of Market Price Data ................. 230
C.1.3 The Problem of Data Collection ..................... 230
C.1.4 Using the Mobile Telephone Short Messaging Service (SMS) 231
C.1.5 Limitations (and Solutions) to SMS ................... 231
C.1.6 What Information Do We Need? ...................... 232
C.1.7 Two Rules .......................................... 232
C.1.8 Message Format .................................... 233
C.2 Location Code Sheets .................................. 235
C.3 Commodity Code Sheets ............................... 235

D Code Examples .............................................. 237
D.1 Code for Standardised Morbidity Ratios ................. 239
D.2 Procedural language functions for routing ............... 239
D.2.1 Road distance calculations in South East Asia ...... 243
D.3 Livestock movement directions for online mapping ...... 244

Bibliography .................................................. 247
List of Tables

3.1 Reports of foot and mouth disease by serotype and country in South-East Asia from 2000–2010. 48

4.1 Data fields obtained from the NLIS database for animal movements in Western Australia. 68

4.2 Data fields obtained from the livestock movement permits database for animal movements in the Greater Mekong Subregion. 72

4.3 The weighting applied to different road segments of the Western Australia road network to enhance the probability that the routes chosen by the routing algorithms would match those chosen by livestock transporters. 77

4.4 A précis of network analysis terms used in this chapter. 81

4.5 Number of movements of cattle recorded in the NLIS database for WA from July 2005 to June 2008. 89

4.6 Number of cattle in WA from July 2005 to June 2008. 89

4.7 Matrix of total movements by origin and destination node type. Loops (movements between nodes of the same type) are shown in bold. 91

4.8 The number of movements within and between countries from January 2007 to December 2009 by species. 100

4.9 The number of animals moved within and between countries from January 2007 to December 2009 by species. 101

4.10 The number of buffalo, cattle, goats and pigs moving through a selection of checkpoints in Laos during 2009. 101

5.1 Definitions for price categories for cattle and buffalo. 124

5.2 Definitions for price categories for pigs. 124

5.3 The data fields associated with price data extracted from the database for analysis. 139

5.4 The number of reports of market prices received by species and country from January 2008 to February 2010. (Each report may contain more than one price record) 142
<table>
<thead>
<tr>
<th>Table</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>5.5</td>
<td>Statistical outputs from a multilevel linear model using the natural logarithm of price as an outcome for price variation for cattle in Cambodia and Laos.</td>
</tr>
<tr>
<td>5.6</td>
<td>Statistical outputs from a multilevel linear model using the natural logarithm of price as an outcome for price variation for pigs in Cambodia and Laos.</td>
</tr>
<tr>
<td>6.1</td>
<td>Provinces from which movements of pigs in Cambodia in 2009 originated and ended.</td>
</tr>
<tr>
<td>6.2</td>
<td>Statistical outputs from a logistic regression model using the occurrence of movement between two provinces as an outcome for cattle in Cambodia.</td>
</tr>
<tr>
<td>6.3</td>
<td>Statistical outputs from a logistic regression model using the occurrence of movement between two provinces as an outcome for pigs in Cambodia.</td>
</tr>
<tr>
<td>A.1</td>
<td>Details of ear tagging of cattle in six villages in three provinces in Northern Laos.</td>
</tr>
<tr>
<td>B.1</td>
<td>Mapping of received signal strength indication (RSSI) to signal strength (in decibels) for the Teltonika ModemCom-G10.</td>
</tr>
<tr>
<td>B.2</td>
<td>The number of text messages (SMS) received by a mobile phone at a given altitude/depth above the ocean surface at Cable Beach, Western Australia.</td>
</tr>
</tbody>
</table>
List of Figures

3.1 Temporal distribution of reported outbreaks of foot and mouth disease in South-East Asia per week for Type O serotype. 45
3.2 Report of foot and mouth disease virus serotypes in South-East Asia, by year from 2000 to 2010 49
3.3 Cumulative annual outbreaks reports of foot and mouth disease in South-East Asia normalised over the outbreak period 2000 to 2010. 50
3.4 Mid-point rooted neighbour joining tree showing phylogenetic relationships of Asia 1 foot and mouth disease virus in South-East Asia using VP1 sequence data from GenBank. 51
3.5 Association between outbreaks displayed using multi-dimensional scaling of the phylogenetic relationships of Asia 1 foot and mouth disease virus in South-East Asia using VP1 sequence data from GenBank. 52
3.6 Result of the point model of foot and mouth disease outbreak reports in South-East Asia showing increased intensity of reports in the East of the region. 53
3.7 Provinces in South-East Asia where reported outbreaks of foot and mouth disease empirical Bayes log-normal estimates were >1 for five or more years during the period 2000 to 2010. 54
3.8 Choropleth map showing the likelihood of underreporting of foot and mouth disease in South-East Asia during 2000 to 2010. 55
3.9 Distribution of reported foot and mouth disease outbreaks in South-East Asia during 2000 to 2010 57
4.1 A screenshot showing the capabilities of the online reporting interface to show real-time maps of animal movements based on permit data entered from provincial movement records. 71
4.2 Western Australia showing management regions used in this study. 74
4.3 An example of a network graph. 83
4.4 A screenshot of the web based National Livestock Identification System query interface developed for this research. 85
### LIST OF FIGURES

<table>
<thead>
<tr>
<th>Figure</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>4.5</td>
<td>A graphical output of the web based National Livestock Identification System query interface showing the network by node category.</td>
</tr>
<tr>
<td>4.6</td>
<td>Movement of three animals in Western Australia during the period July 2005–June 2008.</td>
</tr>
<tr>
<td>4.7</td>
<td>Seasonal decomposition of movements of cattle by month in Western Australia during the period July 2005 to June 2008.</td>
</tr>
<tr>
<td>4.8</td>
<td>Moves of cattle by season and management zone in Western Australia during the period July 2005–June 2008.</td>
</tr>
<tr>
<td>4.9</td>
<td>Centrality measures by node type in Western Australia during the period July 2005–June 2008.</td>
</tr>
<tr>
<td>4.10</td>
<td>Choropleth maps showing in-degree and out-degree by shire in Western Australia during the period July 2005–June 2008.</td>
</tr>
<tr>
<td>4.11</td>
<td>Choropleth map showing betweenness by shire in the south-west of Western Australia during the period July 2005–June 2008.</td>
</tr>
<tr>
<td>4.12</td>
<td>Number of Western Australia shires likely to become infected by season.</td>
</tr>
<tr>
<td>4.13</td>
<td>Plots of the seasonal decomposition of movements of cattle in Cambodia and Laos from January 2007 to December 2009.</td>
</tr>
<tr>
<td>4.14</td>
<td>Maps showing movements of cattle based on Cambodian movement records showing the direction of movement in 2004 and 2009.</td>
</tr>
<tr>
<td>4.15</td>
<td>Map showing the movements of cattle in 2009 based on the complete set of movement records collected during the research showing the direction of movement.</td>
</tr>
<tr>
<td>4.16</td>
<td>Network centrality of provinces in Cambodia and Laos during 2009.</td>
</tr>
<tr>
<td>5.1</td>
<td>A graphical representation of the complexity associated with categorisation of livestock by species, age, condition and sex (and breed for pigs).</td>
</tr>
<tr>
<td>5.2</td>
<td>Example decision tree for choice of parameters to collect when evaluating prices of cattle or buffalo.</td>
</tr>
<tr>
<td>5.3</td>
<td>A screenshot of the market price data entry form.</td>
</tr>
<tr>
<td>5.4</td>
<td>Number of market price reports by province and species.</td>
</tr>
<tr>
<td>5.5</td>
<td>Image plot showing the number of price reports submitted by each province in Cambodia.</td>
</tr>
<tr>
<td>5.6</td>
<td>Plots showing the variation in the value of regional currencies from January 2008 to February 2010.</td>
</tr>
<tr>
<td>5.7</td>
<td>The scale of the random effect for cattle for each province in Cambodia and Laos.</td>
</tr>
<tr>
<td>5.8</td>
<td>Semivariograms of price of cattle for each month in Cambodia and Laos.</td>
</tr>
</tbody>
</table>
5.9 A plot showing the anisotropy associated with the semivariogram for price in December for cattle in Cambodia and Laos. .................. 150
5.10 Isopleth map from kriging the price of an adult female cattle in medium condition in Cambodia and Laos in October. .................. 151
5.11 The scale of the random effect for each province for pigs in Cambodia and Laos. ................................................................. 153
5.12 Semivariograms of price of pigs for each month in Cambodia and Laos. 155
5.13 Anisotropy associated with the spherical variogram for price in December for pigs in Cambodia and Laos. .......................... 156
5.14 Isopleth map from kriging the price of an adult female pigs in medium condition in Cambodia and Laos in October. .............. 157
5.15 The range of predicted prices for each month in each province over 12 months in Cambodia for cattle and pigs. ......................... 158
5.16 Plots showing the change in price of pig meat, beef and mutton in two southern provinces in China from 2001 to 2008. .......... 160

6.1 Scatter graph showing association between the euclidean distance and the road network distance between province centres in Cambodia. 175
6.2 A map of Cambodian provinces with the ‘least cost’ path road network between provinces used in the calculation of distances shown in yellow. ................................................................. 176
6.3 Plots showing the relationship between the probability of a movement of cattle and pigs occurring between two provinces in Cambodia and the price, distance between provinces and season of travel. .... 180

B.1 A simplified schematic of a GSM radio transmitter within a rumen bolus. ................................................................. 215
B.2 Hypothetical animal tracing example .................................................. 216
B.3 Testing for signal strength when antenna is sited on top of a light gauge metal surface. ......................................................... 223
B.4 Testing for receipt of text messages at specified depths below sea level. 224
B.5 GSM signal strength recorded in various antenna positions and locations. ................................................................. 225

C.1 Map of Project Region ................................................................. 229
C.2 SMS codes for cattle prices ......................................................... 235
C.3 SMS codes for pig prices ............................................................ 236
Approvals

Animal Ethics Approval

The Murdoch University Animal Ethics Research Committee approved the ear-tagging trial in Laos, as the project “Livestock tracking in the Greater Mekong subregion”, Permit No. R2201/08. This approval was also recognised by the Department of Agriculture and Food WA Animal Ethics Research Committee.

FAO Statistics Division Approval

Permission was received from the FAO Statistics Division to use their data.

SaTScan

SaTScan™ is a trademark of Martin Kulldorff. The SaTScan software was developed under the joint auspices of (i) Martin Kulldorff, (ii) the National Cancer Institute, and (iii) Farzad Mostashari of the New York City Department of Health and Mental Hygiene.

SaTScan was used in the evaluation of the FMD reports to identify clusters of disease across the region during the 10 year period that was assessed.
Glossary

ACIAR  Australian Centre for International Agricultural Research
AIC  Akaike’s ‘An Information Criteria’
ARAHIS  ASEAN Regional Animal Health Information System
ASEAN  Association of South-East Asian Nations
AUD  Australian Dollar (currency)
CAMIS  Cambodian Agricultural Market Information System
CSF  Classical swine fever
DAFWA  Department of Agriculture and Food Western Australia
DAHP  Department of Animal Health and Production, Cambodia
DCW  Digital Chart of the World
DLD  Department of Livestock Development, Thailand
DLF  Department of Livestock and Fisheries, Lao PDR
EM  Expectation-Maximisation algorithm for Maximum Likelihood
FAO  Food and Agriculture Organisation of the United Nations
FMD  Foot and mouth disease
FMDV  Foot and mouth disease Virus
FOSS  Free and Open Source Software
GDP  Gross Domestic Product
GIS  Geographic Information System (sometimes Geographic Information Science)
GMS  Greater Mekong Subregion (includes Yunnan China, Myanmar, Lao PDR, Thailand, Cambodia and Vietnam)
<table>
<thead>
<tr>
<th>Acronym</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>GPS</td>
<td>Global Positioning System</td>
</tr>
<tr>
<td>GSM</td>
<td>Global System for Mobile Communications (Groupe Speciale Mobile)</td>
</tr>
<tr>
<td>HPAI</td>
<td>Highly pathogenic avian influenza</td>
</tr>
<tr>
<td>IMEI</td>
<td>International Mobile Equipment Identity</td>
</tr>
<tr>
<td>ISO</td>
<td>International Organisation for Standardisation</td>
</tr>
<tr>
<td>KHR</td>
<td>Cambodian Riel (currency)</td>
</tr>
<tr>
<td>LAK</td>
<td>Lao Kip (currency)</td>
</tr>
<tr>
<td>LGA</td>
<td>Local Government Area (terminology used in Australia)</td>
</tr>
<tr>
<td>LMCZ</td>
<td>Lower Mekong Control Zone</td>
</tr>
<tr>
<td>LMWG</td>
<td>Lower Mekong Working Group (representatives of the countries involved in the Lower Mekong Control Zone)</td>
</tr>
<tr>
<td>MCMC</td>
<td>Markov Chain Monte Carlo</td>
</tr>
<tr>
<td>MDS</td>
<td>Multi-Dimensional Scaling</td>
</tr>
<tr>
<td>ML</td>
<td>Maximum Likelihood</td>
</tr>
<tr>
<td>MLA</td>
<td>Meat and Livestock Australia</td>
</tr>
<tr>
<td>ND</td>
<td>Newcastle Disease</td>
</tr>
<tr>
<td>NLIS</td>
<td>National Livestock Identification System (Australia)</td>
</tr>
<tr>
<td>NLRS</td>
<td>National Livestock Reporting System (Australia)</td>
</tr>
<tr>
<td>OIE</td>
<td>World Organisation for Animal Health (Office International des Épizooties)</td>
</tr>
<tr>
<td>PIC</td>
<td>Property Identification Code</td>
</tr>
<tr>
<td>R</td>
<td>The R Statistical Environment and Language</td>
</tr>
<tr>
<td>RCU</td>
<td>Regional Coordinating Unit (of the OIE)</td>
</tr>
<tr>
<td>RDBMS</td>
<td>Relational Database Management System</td>
</tr>
<tr>
<td>REML</td>
<td>Restricted Maximum Likelihood</td>
</tr>
<tr>
<td>RFID</td>
<td>Radio Frequency Identification Device</td>
</tr>
<tr>
<td>RMB</td>
<td>Chinese Yuan Renminbi (currency)</td>
</tr>
</tbody>
</table>
GLOSSARY

RRL  Regional Reference Laboratory for foot and mouth disease (Pak Chong, Thailand)
SAHMBA  Strengthening Animal Health Management and Biosecurity in ASEAN
SEA  South-East Asia
SEAFMD  South-East Asia foot and mouth disease [Campaign / Programme] (until 2010)
SEACFMD  South-East Asia and China foot and mouth disease [Campaign / Programme] (from 2010)
Shapefile  Shapefile is a file format for GIS data developed by ESRI Pty Ltd. Although it is a de-facto standard for exchange of spatial data it suffers from a number of limitations which are now much better handled using spatially enabled relational databases.
SIR  Susceptible–Infected–Recovered (or Removed) a model where animals are either removed after infection (for example, they die) or recover but retain long term immunity.
SIRS  Susceptible–Infected–Recovered–Susceptible—a model with four states, representing a scenario where recovered animals become susceptible to the disease process again.
SMR  Standardised Morbidity Ratio
SMS  Short Message Service
SQL  Structured Query Language—a standard language for querying relational databases. SQL is used by most large databases (although each one may have some slight variations).
THB  Thai Baht (currency)
ULM  Understanding Livestock Movements—the short title of ACIAR project AH/2006/025
UMCZ  Upper Mekong Control Zone
UMWG  Upper Mekong Working Group (representatives of the countries involved in the Upper Mekong Control Zone)
Unicode  A standard for encoding text on computers that represents all known characters individually. It can also account for variations in character form when combined in Indic languages.
<table>
<thead>
<tr>
<th>Abbreviation</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>USD</td>
<td>United States Dollar (currency)</td>
</tr>
<tr>
<td>VAHW</td>
<td>Village Animal Health Worker (Cambodia)</td>
</tr>
<tr>
<td>VMAP</td>
<td>Vector Map Format</td>
</tr>
<tr>
<td>VND</td>
<td>Vietnam Dong (VND) (currency)</td>
</tr>
<tr>
<td>VVW</td>
<td>Village Veterinary Worker (Lao PDR)</td>
</tr>
<tr>
<td>WA</td>
<td>Western Australia</td>
</tr>
<tr>
<td>WHO</td>
<td>World Health Organisation</td>
</tr>
<tr>
<td>WRL</td>
<td>World Reference Laboratory for foot and mouth disease (Pirbright, United Kingdom)</td>
</tr>
</tbody>
</table>
Acknowledgements

The Big Picture

A large number of people were extremely generous with their time and expertise and I hope I have acknowledged them all below. If there are some I’ve missed, I hope I will be forgiven. To all those who have helped, I am truly grateful—thank you!

The support from a number of people kicked off this undertaking, including Drs Angus Cameron, Chris Baldock, Lisa Adams and John Edwards. The ongoing support of the ABCRC (especially Drs Peta Edwards and Stephen Prowse) actually made it possible. My colleagues at AusVet (Drs Angus Cameron, David Kennedy, Evan Sergeant, Jenny Hutchison and Nigel Perkins) were extremely generous with their professional skills and personal support over four long years, not to mention allowing me the time to complete writing up the thesis!

Foot and Mouth Disease

This research was done as a collaboration with the South East Asia and China Foot and Mouth Disease control programme. The support of Dr Ronello Abila in making this data available is greatly appreciated, as is the contribution of the National FMD Coordinators of the member countries. Drs Mark Stevenson and Sean Lafferty provided useful suggestions for further investigation and Dr Wilna Vosloo provided assistance with the phylogenetic analysis. Dr Michael Höhle was patient
with advice on the use of outbreak detection algorithms. Professor Michael Ward was extremely patient and supportive with advice on making the chapter better through his support in developing a journal article. Two anonymous reviewers had to suffer my first independent attempts at scientific writing, so I am humbled that they read to the end and provided such helpful and supportive advice.

**Movements - Western Australia**

Information on the structure of the NLIS database was provided by Dr Greg Hood from Bureau of Resource Sciences. Dr Bob Vassallo from DAFWA took the time to help with extracting and coding the raw data into a format suitable for analysis whilst ensuring privacy. Haydn Bufton from Main Roads Western Australia generously supplied statewide road data.

**Movements - Greater Mekong Subregion**

An enormous number of people contributed to the collection of data in this region. In Cambodia, Drs Sorn San, Holl Davun and Nget Kiry of the National Veterinary Research Institute, Phnom Penh, Dr Tum Sothyra of the Public Health Office and Dr Sieng Socheat provided an enormous amount of insight and assistance. Dr John Stratton was ever helpful and contributed to understanding of the Cambodian livestock production arrangements as well as stimulating thought on the rationale for livestock control. In Laos, Dr Phouth Inthavong was the keystone for the project in the Disease Investigation Centre in Vientiane, guided by Dr Syseng Khounsey and ably assisted by Anousone, Souksavanh, Phoutsakhon and Khamphanh. Dr Axelle Scoizec lent generously from her concurrent research working with livestock traders. Alison Wilson contributed illustrative data and insights from her work with trade development. The cooperation of Dr Luzia Rast and Professor Peter Windsor and their project staff in Northern Laos for the ear-tagging trial was vital.
Prices

Nancy Bourgeois-Lüthi was a goldmine for background on the pricing of livestock across the GMS. Dr Nigel Perkins cleared the way for dealing with an enormous pile of data as well as encouraging me to believe that I could ever understand hierarchical models. A number of people actually responded to my expert opinion survey, for which I am grateful, however as it was anonymous, I can’t identify them.

General support

All the data collection, analysis, presentation and the writing of the final document were written using free and open source software. This would not have been possible without the support of the open source community, those that not only code and provide for free this amazing software, but then patiently respond to emails from distressed users. This camaraderie was overwhelming when struggling from remote locations. Dr Petra Muellner receives a special mention for suggesting \LaTeX.

A smaller group went above and beyond the call of duty in their help with technical advice, rhetorical guidance, patient (and oft repeated) explanation and editing of inarticulate dribbling thoughts.

Dr Chris Hawkins from DAFWA provided invaluable feedback on content and structure of the movements paper in Western Australia and general advice on thesis content and appealing to journals. Professor Ian Robertson kept the university off my back and suffered through reading some of my early writing. Dr Angus Cameron set up the foundation for the thesis, an outline which I somehow managed to (mostly) follow, suffered endless interruptions via Skype, email, phone and retained his good humour throughout, even when I turned up on his doorstep in Sydney and invited myself to stay; Catriona Mackenzie, having suffered through all the above also deserves a thank you. Dr Simon Reid laid out a framework for writing a thesis that actually allowed me to get writing, while I sat in his dining room and drank his coffee and interrupted his attempts to pack up and sell his house before moving overseas.
Drs Jan Lievaart and John Morton put an enormous effort into reviewing a number of chapters, and providing detailed feedback which enhanced the final work—I’d love to say significantly, but I can’t calculate the P-value! As well as the modelling, Dr Nigel Perkins provided endless support on the practicalities and realities of the thesis process and presentation.

My brother Josh gave up many hours to read my early attempts at writing, and patiently put me on track to actually write a document. He then rose to the challenge a second time and actually read the re-written document again. Thanks mate!

Finally, Jim Kerr was fantastic in his relentless work keeping the larger project moving forward and data collection activities on track. It is no doubt that without his support, the components on which I was engaged would not have succeeded. His open interest in all that was going on around him was inspiring.

The smaller picture

A smaller group again spent the best part of four years living with a frequently absent father, both physically and at times mentally while he was struggling with this undertaking. Eloise (now 9), Cooper (now 7) and Meg (now 2) were very accepting of this, although Cooper did (not unreasonably) assert towards the end that he would never put his family through such an undertaking! I can only hope that one day he understands the motivation.

One person made a sacrifice of a magnitude that was unanticipated when drafting up the marriage vows. Indeed, who could anticipate evenings of patiently sitting through hilarious anecdotes about database indexing and hierarchical regression and countless mornings, days and evenings spent alone. Nevertheless, she displayed a commendable commitment to proofreading the worst of the scribblings before anyone else was subjected to them, and sharpening up my writing and working hard on curtailing my best efforts at procrastination. It is of course to my wife and best friend Rachael that I dedicate this work.
Chapter 1

General Introduction

“Judge a man by his questions rather than by his answers” (François-Marie Arouet, 1694–1778)
1.1 Introduction

Livestock movement is a daily fact of life in the Greater Mekong Subregion (GMS). Moving by foot, on motorbikes, in cars and trucks livestock undertake journeys ranging from a few to thousands of kilometres, in response to a steadily growing hunger for animal protein across the GMS and beyond. This kind of activity is not without its risks. While profits and losses are made every day by those directly involved in the trade of livestock, there is also an inherent risk to the broader livestock industry from the spread of disease associated with the movement of these animals. A better understanding of the nature of and reasons behind these movements is vital in focussing limited resources into reducing the spread of disease, and making recommendations of suitable changes to policy and practices associated with making livestock movement safe and efficient.

This system of movements is highly complex, and developing a methodological framework to better understand this system and the risk of disease spread is the primary objective of the work described in this thesis. In order to achieve this, investigations were undertaken relating to: the availability of reliable disease information; capturing data on livestock movements; and prediction of movement patterns and risk of disease spread.

First, could disease occurrence be reliably detected rapidly enough to be used as an input into an early warning system. For the purpose of this work a single disease, foot and mouth disease (FMD), was studied as it is a priority disease for the region and subject to a control programme, the South-East Asia and China foot and mouth disease (SEACFMD) campaign. The SEACFMD campaign encourages countries to share information about outbreaks occurring within the region by reporting into a regional animal health information system.

Second, could animal movements be recorded and analysed to describe the movement patterns and use them to identify at-risk areas for the introduction of disease as a result of an outbreak in another part of the region. In contrast to the disease data,
no central repository of information on animal movements exists in the region, so a number of different approaches to investigate and analyse patterns of movement were used.

Finally, to avoid the expensive ongoing collection of animal movement data in the region, the potential of using differences in livestock price to predict movement at a province level was investigated.

Key components of this research relating to livestock movement and prices in the GMS were conducted as part of a larger project being conducted in Cambodia and Laos at the time. This project was funded by the Australian Centre for International Agricultural Research (ACIAR) and was titled “Understanding Livestock Movement and the Risk of Spread of Transboundary Animal Diseases”\(^1\). This project was also referred to as “Understanding Livestock Movements” (ULM) by project partners in the GMS. The FMD data were made available for analysis by the SEACFMD campaign. The data used for movement investigations in Western Australia (WA) were made available by the Department of Agriculture and Food Western Australia.

### 1.2 Outline of Approach

In order to understand better the issues associated with this research, Chapter 2 provides an overview of the relevant literature. This includes literature related to FMD, disease reporting, livestock movements and livestock prices. In addition, to improve the understanding of the context under which this work was undertaken, this review begins with a description of some changes occurring that are affecting the risk of livestock disease spread in the GMS through their impact on society and in particular the livestock production system.

In Chapter 3 an evaluation of disease reporting based on the data submitted to the ASEAN\(^2\) Regional Animal Health Information System (ARAHIS) is summarised. The objective of this work was to develop estimates of the probability of FMD oc-

---

\(^1\)ACIAR Project AH/2006/025  
\(^2\)Association of South East Asian Nations
1.2. OUTLINE OF APPROACH

occurring in provinces in the GMS. As well as identifying foci of disease, identifying those provinces likely to be underreporting FMD was also attempted. This was probably the first large scale assessment of the ten years worth of data contained in the ARAHIS. The analysis was constrained by the quality of the data and the subsequent limitations on results due to questions about the sensitivity and specificity are discussed. The large number of different topotypes of FMD virus circulating in the region complicates the impression of disease spreading when SEA is treated in isolation. Using log-normal estimates of provincial disease risk provides a reasonable basis for developing a probability of a province being infected in a given month. Estimates of the probability of provinces having disease outbreaks, but not reporting it, are also developed. This provides an important mechanism for focusing activities designed to improve reporting at the provinces that most need it. Chapter 3 was originally published in Preventive Veterinary Medicine (Madin, 2011).

Chapter 4 focuses on the animal movement pathways, identifying and describing the actual routes along which the trade occurs. This work was conducted to develop two main outputs. One was a suitable set of clean data and appropriate statistics to provide some objective details and guidance to a predictive model about movements. The second was the development of suitable outputs (graphical and tabular) to provide policy-makers, industry and the broader public with a clear and unambiguous display of what trade was occurring. In order to evaluate techniques for analysing the GMS data the Western Australian cattle movement records for the period 2005–2008 were analysed as an example of a complete network prior to an evaluation of the situation in the GMS. The same approaches were then applied to the GMS data focussing on relative measures of network connectivity in order to provide estimates of risk for disease introduction at a province level.

As collection of movement data is difficult and expensive, in Chapter 5 the use of postulated key drivers for this trade is investigated, with a focus on using price as a proxy for the balance between supply and demand. To determine if reported variations in animal movement patterns (volume and direction) were correlated with
changes in price at different locations the prices of three FMD susceptible livestock species (cattle, buffalo and pigs) were collected across all provinces in Cambodia and Laos. Price data may be collected much more cheaply than the direct measurements on animal movement, but teasing out the values which allow comparison of price gradients between provinces is a far more complex issue. In this case the use of a multilevel linear model provided a solution.

In Chapter 6 the relationship between the movements analysed in Chapter 4 and the price values generated in Chapter 5 was modelled to determine the role of price in predicting animal movements, and therefore its utility as a key element of a stochastic model for predicting risk of animal movements and potential disease spread.

Attached to this thesis are two appendices detailing investigations that were undertaken to explore prospective animal movements and better tracking of individual animals. Both were unable to contribute to the thesis however they offer some insights into the difficulty in investigating animal movements and are as such discussed as appropriate.

Overall, the thesis integrates the information on disease occurrence, animal movements and animal price in order to allow the development of predictive variables for a modelling framework for the risk of the introduction of disease associated with livestock trade. The characterisation of the relationship between animal movements and animal price will support the development a model without the requirement for the ongoing (and expensive) capture of animal movement information.
Chapter 2

Review of Relevant Literature

“It is important to understand what you can do before you learn to measure how well you seem to have done it.” (Tukey, 1977)
2.1 Livestock Production in a Rapidly Changing Society

Major demographic changes are occurring in the Greater Mekong Subregion (GMS). A previously agrarian society is becoming increasingly industrialised, wealthy and urbanised (Knips, 2004). The increases in industrial processing and manufacturing have been particularly marked in the powerhouses of Vietnam (Hanoi and Ho Chi Minh) and China (more generally), as well as Cambodia. While animal proteins are considered luxuries (compared with staple crops like rice and wheat) the demand for them is normally associated with an income elasticity of one (Leslie and Upton, 1999), meaning that increasing income is linearly related to increases in demand for meat and other livestock products. Huang and Bouis (1996) also contend that the demand for protein is driven by lower demand for energy in the diet and wider exposure to a varied diet in urban environments. Although demand has remained basically static in developing countries, and even fallen in parts of Africa, in South-East Asia (SEA) demand has increased by 2.5–5% per annum since the 1980’s, with production generally keeping up (Delgado et al., 1997; Stür et al., 2002). Thus the major issue facing the livestock production chain is delivering product to urban markets from increasingly distant production areas. As a result, managing the risks associated with this movement, particularly the risk of spreading disease, is increasingly important for livestock producers.

Fifty-nine million people in SEA were described as livestock keepers living below the poverty line in 2000; this equate to about 12% of the population in the region (Perry et al., 2002b). These producers are typically involved in mixed practice agriculture with cropping and other activities providing household income, but a dependency on livestock for subsistence is considered a common feature of the poorer sectors, and a key issue for poverty reduction initiatives (Rapsomanikis and Maltzoglou, 2005). For these producers livestock production is not intensive. Although slow growth and low animal performance are expected, a serious transboundary animal disease (TAD) such as foot and mouth disease (FMD) or highly pathogenic avian influenza (HPAI) can be devastating economically (Khounsy et al., 2008; Stür et al.,...
Any movement of a live animal carries a risk of spreading disease from an infected population into naïve populations of animals thus creating a risk for an epidemic outbreak (Bessell et al., 2008; Savill et al., 2006). Understanding the drivers of this movement, including the seasonal and long term variation, is essential to being able to model potential disease spread and target the implementation of programmes to reduce the risk of the spread of disease through the GMS. The Food and Agriculture Organisation of the United Nations (FAO) definition of a TAD is:

“Those [diseases] that are of significant economic, trade and/or food security importance for a considerable number of countries; which can easily spread to other countries and reach epidemic proportions; and where control/management, including exclusion, requires co-operation between several countries.” (Otte et al., 2004)\(^1\)

This definition is still being used at an international level (Lubroth et al., 2011), and whilst not incorrect might be considered to be lacking with no explicit mention of zoonotic risk. As well as FMD and HPAI, other significant diseases in a regional context include classical swine fever, newcastle disease and rabies (World Organisation for Animal Health, 2011). Since 2006, porcine reproductive and respiratory syndrome has increasingly been causing concern in the GMS (Dietze et al., 2011).

Increasing wealth and focal urbanisation are clearly important drivers of livestock trade, but other factors exist and include climatic and topographic suitability for livestock production and varying cultural and religious beliefs. A number of considerations affect, or are likely to affect, the value of livestock in the GMS (Knips, 2004; Osborne, 2009):

- Agricultural intensification leading to pressure for higher value production than is possible with rice or other field crops;

\(^1\)Essentially the same definition also appears from 1996 at [http://www.fao.org/DOCREP/004/W3737E/W3737E04.htm](http://www.fao.org/DOCREP/004/W3737E/W3737E04.htm).
2.1. LIVESTOCK PRODUCTION IN A RAPIDLY CHANGING SOCIETY

- the opportunity for improved health associated with higher protein consumption by those in rural areas;

- the opportunity to store capital by investment in livestock (particularly cattle and buffalo) which can be liquidated relatively easily if cash is required;

- the transfer away from aquatic protein due to a potential decline in wild fish populations associated with reduced flows in the Mekong due to dam building projects.

The widespread distribution of the population, dependency on agrarian production, limited public resources, low levels of income and education and an institutionalised distrust of authority have previously made the control of animal movements by regulatory means impracticable. As a result efforts to enforce controls have inspired the emergence of a number of practices focussed on actively avoiding any contact with regulatory authorities—or expediting the contacts that do occur (Cocks et al., 2009).

A blanket ban on animal movements is impracticable and it would result in significant hardship for both livestock producers and consumers. Neither is it legal under the rules of the World Trade Organisation (World Trade Organisation, 1994). In any case, many countries do not have the legal frameworks or resources to implement animal movement controls, a matter about which they may be relatively candid.

“We are aware of the illegal import trade, but we can not enforce the ban to prevent it happening, as people smuggle birds across the border in remote areas where there are less enforcement officers.” Livestock and Fisheries Department Director General, Dr Bounkhuang Khamboungheuang, attributed the latest bird flu outbreak in Xayaboury province to illegal imports from Thailand (Vientiane Times, Thursday November 13, 2008).

Long borders and limited regulatory resources are often cited as impediments to effective control of livestock movements across borders (Ansell et al., 1994; Perry
et al., 2002b). In reality, most movements occur along existing roads, and border crossings are either through or—perhaps more importantly—near\(^2\) official border checkpoints, using generally well known specific locations for the crossing especially where mountainous terrain or rivers are a feature of the local geography (Cocks et al., 2009). During the course of this research, on a number of occasions animals were observed being unloaded from trucks just prior to reaching the official border checkpoint. Local villagers were then paid to walk these animals along tracks away from the road and across the border, before meeting the road again some kilometres past the border to be loaded onto other trucks for the onward journey. Discussions with villagers indicate that this was a major source of income for the village, and it was not uncommon for one family member to work at the checkpoint in an official capacity, while other family members walked livestock around the same checkpoint. Based on this information, relying on the detection of disease at the point of entry is an error-prone strategy as many animals will never pass through an inspection checkpoint.

Recently, there have been repeated calls for improved understanding of trade routes, trade volumes and the risks associated with livestock movements over larger areas than individual countries where consistent standards may theoretically be applied (Abila, 2008; Fèvre et al., 2006). This understanding would imply looking at regional spread across a number of neighbouring countries, or even global movements (for high value livestock).

Exactly how much illegal trade in livestock occurs is unknown. While there are some resources which present data on international trade that has been collated through formal channels (such as the United Nations Food and Agriculture Organisation’s FAOSTAT\(^3\)), there is much less certainty about the illegal trade, which in some cases has been shown to be of higher risk for the spread of disease due to the impact of measures taken to avoid official checkpoints or inspections. Recognising

\(^2\)There are reports of outbreaks of foot and mouth disease occurring in villages near checkpoints after animals have been walked through to avoid the checkpoint (Mohamed Naheed bin Mohd Hussein, personal communication, 2009).

\(^3\)http://faostat.fao.org/
that the enactment of legislation at a national level is only a first step, and in many countries is impacted upon by a lack of resolve and resources, the World Organisation for Animal Health (OIE) is leading the development of programmes\(^4\) to enable members to implement sanitary measures to help to reduce the risk of disease spread (Fèvre et al., 2006).

### 2.2 Foot and Mouth Disease

#### 2.2.1 Foot and Mouth Disease Virus

Foot and mouth disease virus (FMDV) is an aphthovirus which affects artiodactylid species including cattle, sheep, pigs, goats and buffalo (Geering et al., 1995; Knowles et al., 2005; Radostits et al., 2000; Samuel and Knowles, 2001). The virus causes fever and targets the epithelium of high wear parts of the body resulting in characteristic vesicular lesions in the mouth, on the snout, around the feet and occasionally teats. Although the morbidity rate in naïve animals may approach 100%, a number of authors suggest the case fatality rate in adult cattle is typically 2% or up to 20% for neonates (El-Kholy et al., 2007; Radostits et al., 2000). For lambs, Ryan et al. (2008) reported a mortality rate ranging from 5 to 95%, and Kitching and Hughes (2002) reported a mortality rate of up to 90%. The contagious nature of the virus was well understood prior to its identification in 1898 by Loeffler and Frosch\(^5\). The corresponding high prevalence of intra-herd infection allows villages to be used as a unit in analysis and modelling (Brown, 2003; Ferguson et al., 2001; Keeling et al., 2001).

During the 1920’s researchers recognised that failure in vaccine protection was due to FMDV having more than one serotype with no cross protection. Vallée and Carré initially identified two serotypes (Oise and Allemagne) in France in 1922. In 1926,
Waldmann and Trautwein confirmed these two serotypes\(^6\) and added a third which they called type C. Since 1926 by agreement with the OIE they are now known as O, A\(^7\) and C (Brown, 1986; Knowles and Samuel, 2003). Subsequently, four other serotypes have been identified, three from the South African Territories (SAT1, SAT2 and SAT3), and Asia 1 (Brown, 2003). On a number of occasions during the 1950’s and 60’s vaccination failure was attributed to antigenic variation within serotypes. As a result the importance of identifying the topotype circulating in the field prior to embarking on a vaccination campaign is now widely understood (Brown, 2003).

Less clearly understood is the enormous impact advances in typing virus strains could have on advancing our epidemiological understanding of the disease. Recent work has clearly shown that strain variation within serotypes can be as significant in understanding the epidemiology of the disease as basic serotype information (King D, personal communication, 2010). What is clearly understood is that the paradigm of thinking of FMD as a single virus, and even the concept of serotype, is now insufficient to manage the development of a clear epidemiological picture of the disease (Rweyemamu et al., 2008b).

According to the World Organisation for Animal Health (2010b), FMD is endemic throughout the ‘mainland’ nations of SEA and their continental neighbours. Although Indonesia has been free FMD since 1983, it is only in the last few years that the Philippines has approached FMD freedom, with no outbreaks reported since 2005. Since June 2011 the Philippines has been recognised by the OIE as ‘Free from FMD without Zoning’. There have been no reports of FMD to the OIE by any of the three countries on Borneo (Malaysia, Indonesia and Brunei Darusalam) (Rweyemamu et al., 2008b). In accord with this, Gleeson (2002) reported that it was apparent that the prevalence of FMD in eastern Malaysia (Sabah and Sarawak) was likely to be zero.

Only three serotypes of FMD appear currently to be endemic in mainland SEA and

\(^{6}\)Waldmann and Trautwein called them Types A and B

\(^{7}\)O from Oise and A from Allemagne
contiguous countries; types O, A and to a lesser extent Asia 1 (World Organisation for Animal Health, 2010b). Type C was historically reported from the Philippines but has not been detected for over 15 years\(^8\) (Gleeson, 2002).

The O SEA Topotype has always been present throughout the GMS—recently it has been almost exclusively the Myanmar 98 strain (The previously ubiquitous Cambodia 94 strain has not been reported in the last five years). Since the late 1990’s the type O Middle East–South Asia (PanAsia) topotype has been identified in most countries (Khounsy et al., 2009). More recently Vietnam has reported the O Cathay topotype (Knowles et al., 2005) and it was also reported in Thailand in 2007. The only FMD type A topotype identified has been ASIA\(^9\).

In the GMS, Asia 1 has only been reported from Vietnam and Myanmar during the 21st Century, most recently in 2005 (Linchongsubongkoch et al., 2010). During this time outbreaks of Asia 1 in China, Russia, India and part of the Middle East were also detected, however they all appear to have significant variation from the local outbreaks (Valarcher et al., 2005).

### 2.2.2 Vaccination

Vaccination is currently used in SEA in a number of scenarios. In ‘Region 2’ in south-east Thailand there is a programme of vaccination to try to create a disease free area (Chotiyaputta, 2009). There is also a program of providing small amounts of vaccine to GMS countries to allow the implementation of ring vaccination during outbreaks. Current annual vaccine use ranges from 25 million doses (in Thailand) down to 10–20 000 doses in Cambodia and Laos (Stratton, 2010). The cost of large-scale vaccination is likely to be unaffordable for countries such as Cambodia and Laos (Gleeson, 2002).

Although Forman and Leslie (1997) suggested a range of ways in which vaccine could be used, they did not foresee a time when mass vaccination programmes were

---

\(^8\)Globally Type C has not been reported since 2004 (Hammond, 2011).

likely to be a viable solution. Some farmers view vaccination with distrust and Perry et al. (2002b) discussed how such distrust has been shown to inhibit the success of vaccination campaigns (in their case in Uganda with Contagious Bovine Pleuropneumonia).

A range of vaccines is available for FMD control, although limited information is available on the antigens included beyond the serotype.

2.2.3 Impact

Foot and mouth disease is arguably the most important disease of livestock from a global trade perspective (James and Rushton, 2002; Rweyemamu et al., 2008a). The economic impacts of FMD in the GMS have been subject to a number of publicly recorded recent studies, some as a part of a larger technical report (Perry et al., 2002a; Perry and Grace, 2009; Ramsay et al., 1999; Randolph et al., 2002; Thompson et al., 2002; Tung and Thuy, 2007). Critically, many of these studies differ in their approach to methodology, geographic region, industry sector(s) assessed and motivation for undertaking the study, making it difficult to compare the findings. This disparity is not unusual and conflicting reports are in the literature. Rushton (2008) discussed an example in Bolivia where two separate economic analyses were conducted within a two year period—the first suggested a benefit-cost ratio of 0.6, the second 1.38. A valuable contribution provided by Rushton’s investigation is that the underlying reason for the difference was a lack of understanding of the disease epidemiology and particularly the impact of livestock trade on FMD spread.

While FMD can cause significant production losses in intensive agricultural settings, the direct impact is far less clear where the raising of livestock is carried out under suboptimal conditions. The common sequelæ to FMD infection—delayed growth rates, increased calving/farrowing intervals—are subclinical and not necessarily obvious or well understood by poor livestock producers, and have been termed ‘invisible losses’ by Rushton et al. (1999). For example, unpublished data collected on produc-
tion in two villages in Cambodia as part of an ACIAR project looking at population dynamics identified calving rates (calves born per female of reproductive age per annum) in the order of 7–11%. Not surprisingly then, Perry et al. (2002b) identified that FMD in SEA was considered a relatively low priority by poor producers. FMD is rarely fatal, so other diseases such as haemorrhagic septicaemia, clostridial diseases, anthrax and parasitism are likely to be rated much more of a problem for producers in the GMS, along with low feed quality and availability during the dry season and inherently low productivity from native breeds (Forman and Leslie, 1997; Stür et al., 2002). Nonetheless, a common approach amongst smallholders to reduce the impact of livestock disease is to diversify their species holding, choosing between a number of the common livestock species, for example cattle, buffalo, pigs, goats and chickens. In this context FMD is significant as nearly all of the common livestock species (except chickens) are susceptible to the disease, thus it represents a much broader risk than other non-zoonotic diseases such as classical swine fever or newcastle disease (Perry et al., 2002a; Perry and Rich, 2007; Perry and Grace, 2009).

Epidemic outbreaks of FMD through the introduction of shedding animals into immunologically naïve populations are a cause of considerable morbidity and occasionally mortality (Khounsy et al., 2008). These outbreaks cause significant impacts on livelihoods of smallholders\textsuperscript{10}. While some of these impacts are direct outcomes of disease, such as reduced growth rates, diminished milk production, loss of draught power and neonatal mortality, they may not be the most severe impacts at a farm level. Measures used to control the disease leading to destruction of affected and unaffected livestock (usually without full market value compensation in SEA), welfare slaughter in the face of movement restrictions and reduction in market potential and limited trade opportunities are often more damaging financially and personally (Tung and Thuy, 2007). This is especially important in regions where large volumes

\textsuperscript{10}In fact, not just smallholders. Globally, a number of recent outbreaks have occurred in previously FMD free countries and areas leading to severe widespread disease in naïve animals in smallholder and intensive farming situations, adversely affecting livestock producers (ProMED-mail, 2011).
CHAPTER 2. REVIEW OF RELEVANT LITERATURE

of livestock trade are associated with regular movement of livestock, such as the GMS (Gleeson, 2002; Perry et al., 2002a).

FMD control has traditionally been viewed from the Western context of managing a highly contagious disease with severe production impacts in intensive animal farming enterprises, particularly dairy and swine (James and Rushton, 2002). This production environment is quite different from that in the GMS, which is characterised by non-intensive production in ‘pastoral’ conditions (low input) where livestock ownership is often about storing capital rather than active production (Perry et al., 2002a). From a national perspective, Rushton et al. (2002) questioned the incentive to control FMD in countries where the opportunity to export meat to high value markets would not be an automatic outcome of the control programme. To illustrate this, Randolph et al. (2002) provide a useful example (using economic modelling) of the importance of the export opportunity. Within their model, eradication in the Philippines would realise a positive benefit-cost ratio of 1.63–2.41 under a ‘no exports’ assumption—the benefit-cost ratio improved fivefold to 7.41–11.95 if ‘low- and high-value exports’ were included. Rather than focussing solely on national or regional eradication, a more realistic solution may be a mix of better understanding of movements, controls based on this knowledge, compartmentalisation of intensive high-value industries (Taylor et al., 2003) and “strategic vaccination campaigns of valuable animals” (James and Rushton, 2002).

2.2.4 Spread of the Disease

Local spread is the primary mechanism by which FMD is maintained in populations where it is endemic, and for the development of epidemics when disease is introduced into naïve populations. Local spread is understood in this context to refer to spread between locations over relatively short distances in the absence of any other specific mechanism (Sanson, 1994).

The spread of viral diseases is normally characterised by factors such as the pres-
ence of a susceptible host population, whether or not direct inoculation occurs, virus survival outside the host and distance between hosts for airborne or fomite spread, or whether the virus can survive in a vector and the presence and range of the vector (Fèvre et al., 2006; Kitching et al., 2005; Thrusfield, 1997). Infection with FMD is normally a result of inhalation of aerosolised virus or ingestion of contaminated material. Different species vary in their susceptibility to infection. Cattle are highly susceptible to infection via inhalation, but much less to infection after oral consumption (Kitching, 2002). Conversely, pigs are more susceptible to infection via ingestion than inhalation (Kitching and Hughes, 2002). Cleland et al. (1995) investigated outbreaks of FMD in 11 villages in Northern Thailand. Of the 5 306 cattle and buffalo at risk (no pigs were reported to be affected in any villages), 960 showed clinical signs of FMD (18%). Of these, adult beef cattle were the largest group and the worst affected class of animals with an attack rate of 24% (771 of 3218 animals showing signs of FMD), while other groups such as younger cattle, buffalo and ‘work cattle’ typically showed an attack rate in the order of 4–11% . This may reflect some pre-existing immunity within the animal groups. The clinical presentation may not be as severe in some indigenous species of cattle in Africa and Asia, and may be slight or sub-clinical in sheep and goats (Kitching, 2002; Kitching and Hughes, 2002; Thomson, 1994). Even within serotypes, some topotypes may only cause clinical disease in a limited range of species (ie Hong Kong type O in pigs, but not in cattle) (Kitching and Alexandersen, 2002).

A population of susceptible animals and the Euclidean (straight-line) distance between infectious animals and the susceptible population are generally regarded as the most important factors in local spread. Climatic conditions and the number and type of infected animals in the population are also important considerations (Bessell et al., 2008; Keeling et al., 2001; Savill et al., 2006). Eblé et al. (2006) demonstrated a significant difference in transmission rates between animals sharing a pen and animals in adjacent pens. Cleland et al. (1996a) suggested that direct transmission is likely to be most important in hot climates with virus survival times on fomites or aerosolised particles so short that they are unlikely to play a role in virus spread.
Under this scenario of limited survival time outside the host, the longer distance spread of disease is probably almost completely dependent on the movement of infected animals or where virus is protected from the environment, for example in meat.

It has been shown that spread through the movement of the host animals is often responsible for epidemics of disease (Food and Agriculture Organisation and World Health Organisation, 2004; Sutmoller et al., 2003). This role is also important where disease is endemic (Cleland et al., 1996b). Perhaps more complex is the situation experienced in SEA where there appears to be a conflict between the reports of endemic disease and outbreaks of the same. Stratton et al. (2010) surveyed over 400 village animal health workers who almost universally reported seeing a disease consistent with FMD\(^{11}\) on a regular basis yet ‘outbreaks’ are still considered to occur on 3–5 yearly cycles (Abila, 2011). A possible hypothesis is that the regular cycling of endemic FMD is responsible for the regular occurrences, and introduction of ‘exotic’ topotypes of FMD drives the majority of large scale outbreaks (King, 2010).

One of the highest risks for epidemic spread is the mixing of livestock from different origins, and this was well documented in the United Kingdom FMD outbreak in 2001 where it occurred in markets (Fèvre et al., 2006; Gibbens et al., 2001). The practice of selling animals displaying clinical signs of disease for a salvage price (Rushton et al., 1999) exacerbates this by providing an opportunity for a trader to realise a greater profit by taking on the risk of a sick animal. Mixing may also occur during transport and in quarantine stations where complete separation is not possible and batch handling (ie all in/all out) is not utilised, such as those observed in 2009 in quarantine stations being constructed in Vietnam during the course of this research (and also reported in Cocks et al., 2009). Cleland et al. (1996b) suggested that simple quarantine measures such as isolating newly arrived animals may be the most effective way of reducing spread of FMD. In western countries this is common practice (Ferguson et al., 2001), but in the context of the resource constraints which will be discussed in section 2.3 this is not always easy to enforce.

\(^{11}\)This was based only on recall of clinical signs, not laboratory testing—see Section 2.3.
2.2. FOOT AND MOUTH DISEASE

2.2.5 International Context

Not surprisingly, FMD causes concern for livestock owners. Media involvement in recent epidemic outbreaks (such as in the United Kingdom in 2001 and 2007, Taiwan in 2009, South Korea and Japan in 2010) has led to a heightened awareness of and interest in how the disease should be managed by the general public and policy makers. This has led to substantial investment by FMD free countries into research and control programmes in infected areas (Escolar, 2011; Morrisey, 2011; Vandersmissen, 2011). These programmes are working hard to raise the awareness by farmers of FMD and promoting it as a very important disease issue. Important or not, there appears to be a noteworthy downside to this level of external interest. During the course of this research it became apparent that the functioning of the animal health services in Cambodia and Laos is heavily dependant on donor funding. Very few staff have government salaries, and many are only paid out of externally-funded projects. In this context the prospect of donor funding may be an incentive to fabricate reports of diseases of interest to overseas funding agencies. Some reported cases of FMD do not have laboratory confirmation. In some cases, further probing may reveal a description of clinical signs that are not consistent with FMD—an outcome which could critically affect the value of a reporting system and thus the design and implementation of a control programme.

Increased market access is a fundamental objective of many developing countries in order to repay bond loans and enhance their economies (Moyo, 2009). Although livestock could contribute to this it is argued that FMD is the most important factor restricting trade in livestock (Knowles et al., 2005), and the implications of this demonstrated by Randolph et al. (2002) have been discussed on page 18. However, the fear of FMD mentioned above is so great in disease free countries that even where suitable protocols exist for accepting livestock products from infected countries (and in some cases even free zones within those countries) it is unlikely that they will be approved (Perry and Grace, 2009) and Taylor et al. (2003). In this context these authors caution that export should not be a primary motivation for
development of disease free zones.

2.3 Resources for Animal Disease Reporting

Many developing nations suffer from a scarcity of resources that impacts on the ability of government to deliver services. Numerous reports from organisations such as the FAO and the OIE have discussed the issue of the limited resources available for animal health in the GMS (Forman and Leslie, 1997) and other developing countries (Perry et al., 2003; Rushton et al., 2002), and have identified issues for priority action.

The scarcest resource in many GMS countries is skilled people. In many cases competent people are recruited into less operational positions or promoted and unable to continue to provide the information and skills the animal health programmes require. In other situations inconsistent and rudimentary use of skills leads to decay of capacity. As a result it is likely that the true presence of disease is unknown and underreporting is highly likely. Recognising this, Sumption et al. (2008) attempted to compensate by using indicators of available resources (disease reporting efficiency and the likely effectiveness of control on movement over borders) to enhance classification of disease status. While this is a valuable contribution to understanding of disease status, it does not overcome the basic issue of limited resources.

A number of issues associated with disease reporting often create problems for research and disease investigation programmes. This includes inappropriate reporting processes, lack of clarity in priorities, poor data collection and a lack of timely analysis. Lines of communication are often established to follow existing bureaucratic processes. In some instances this requires data collected at the village or district level to be directed through the provincial office. Any reports may be delayed or summarised at this level and only further submitted as part of a monthly report (Si-eng S, personal communication, 2010). In other instances reports only come to the attention of national authorities through informal communications at higher govern-
Compounding this are differing sets of priorities about the need to report, or when to report. In much of the GMS FMD appears to be endemic and little or no funding is made available to prevent or manage the disease. This presents a clear message that FMD is not considered important and there is little or no priority to report. This is consistent with a perception of ‘normal’ disease occurrence identified by Hickler (2007) when investigating poor adoption of measures to prevent HPAI. In part this stems from the cultural understanding of disease described as naturalistic—disease occurs and treatment is the only consideration—that is common throughout the GMS. Hickler notes that increasing understanding about the contagion model of disease will be important but provides limited advice on how best to increase understanding.

One of the most common issues is difficulty in reporting due to inappropriate data collection tools and methods. Currently available tools appear to be either highly specific to a particular disease or circumstance or completely generic in such a way that important disease related information is not captured. However, not all staff or all locations have ready access to the appropriate tools—it may only be a small subset of national and provincial staff who have had training and been provided with the appropriate tools.

Although the internet has revolutionised data collection from remote sites, where the site is so remote that the internet is not actually available alternative methods for collation of data are still required. When this research commenced internet was available in major centres in the GMS, but was virtually unheard of outside major towns\textsuperscript{12}. In the absence of the internet, data collection is paper based or uses local computer databases with export of data which is then transferred to a central database, normally on a CD or USB flash drive/memory stick. When combined with the lack of updating anti-virus databases, this creates a nightmare of computer

\textsuperscript{12}Major towns includes capital cities and major tourist destinations such as Siem Reap and Luang Prabang.
virus circulation, computer downtime and a general reluctance to participate—the FAO TADInfo system lies unused in Cambodia in part due to these difficulties (Hoogendijk, 2009).

Inappropriate tools or tools that simply do not work are also major causes of failed disease reporting. Data collection forms and devices that are overly complex may cause frustration with users. Conversely, overly simple approaches may result in poor quality data unsuitable for further analysis. Where staff are trained in using a particular system and told to use it they often will, however Cameron et al. (2004) note that when a system repeatedly fails to work and data entry becomes difficult users will eventually stop trying to enter data and not report. Failure to add value to the data and ensure that the results of analyses are fed back to users will suggest that the information being collected is not particularly important and discourage ongoing contributions. Both of these scenarios exacerbate the under-reporting of disease.

The limited human resources available in many countries make it difficult to analyse and interpret data in a timely fashion. Development of tools to allow for automated analysis of repeat data and aid interpretation are vital to ensure the maximum return is obtained from data collected. Repeated analysis of outbreak information is best managed by automated systems so as not to create an unnecessary burden on staff.

2.4 Available Sources of Outbreak Data

When undertaking a multi-country analysis it is generally not practical to rely on visiting local animal health offices for data collection. A number of potential sources exist for outbreak data, including local disease reports, national animal health information systems, disease specific systems and regional or global systems. For example, Ankers and Harris (2011) discuss three global systems for animal health and zoonotic disease (and briefly mention a number of others):
• The World Animal Health Information System\textsuperscript{13} (WAHIS) is an internet based system operated by the OIE. This system receives official reports of notifiable diseases from member countries, and makes this information available for other countries.

• The FAO’s Emergency Prevention System (EMPRES) programme operates the Global Animal Disease Information System\textsuperscript{14} (EMPRES-i), another web based information system. Unlike WAHIS, EMPRES-i receives data from a number of sources, including official reports, local feedback from non-government organisations and personal contacts.

• A more recent initiative, the Global Early Warning System\textsuperscript{15} (GLEWS) combines information from the two above mentioned systems with information on human disease sourced through the World Health Organisation—this is valuable for animal health as for some zoonoses human disease may be detected earlier than animal cases, thus improving the sensitivity of the detection system.

• Disease bioportal \textsuperscript{16} is a website that aggregates information from a range of sources, allowing greater sensitivity and offers a range of tools for analysing selected data.

These systems provide a great deal of data on animal diseases, and provide a tool for monitoring at a general level. However, for this research a number of key issues limit their usefulness. First, under section 1.1.3 of the Terrestrial Animal Health Code (World Organisation for Animal Health, 2010a), countries only report endemic diseases to WAHIS under specific conditions such as change in epidemiology. For most of the countries in the GMS FMD is assumed to be a constant presence, and apart from the occasional change of serotype or topotype does not warrant reporting to WAHIS. Second, information collected from non-official sources provides difficulties

\textsuperscript{13}Public interface available at http://www.oie.int/wahid
\textsuperscript{14}http://empres-i.fao.org/
\textsuperscript{15}http://www.glews.net/
\textsuperscript{16}http://fmdbioportal.ucdavis.edu/
in interpretation, and is often likely to be incomplete. This limits the usefulness of value-adding systems such as EMPRES-i for statistical analysis. Third, these systems are in general terms part of a monitoring system—they are not directly associated with a disease control programme—and so standards of data collection and verification may be inconsistent between countries. Salman (2003) cautions against using such data for comparisons between countries. Instead, he recommends data sourced as part of a surveillance system associated with a disease control programme. Luckily, for FMD such a disease control programme exists.

FMD was identified as the number one priority disease at a symposium coordinated by OIE and the Federation of Asian Veterinary Associations in 1990. Subsequent to this, with the support of ASEAN and the governments of Switzerland, Australia and Japan, in 1997 a Regional Coordination Unit (RCU) for the South-East Asia foot and mouth disease (SEAFMD) campaign was established (Gleeson, 2002). The OIE coordinates the campaign from the Bangkok sub-regional office. The RCU undertakes a number of activities (grouped into eight core programme areas) aimed at supporting and facilitating the advancement of a control programme in SEA. The progress of the RCU has led to it being hailed as a successful model for the implementation of cooperative regional disease control programs, including strong support from all OIE member nations.

It was intended that control of this programme would be assumed by ASEAN. Although this was intended to have happened by 2004 (Gleeson, 2002) at the time of writing this thesis arrangements for the ongoing funding have not been satisfactorily established (a trust fund has been created to fund the ongoing operations however at this stage there is insufficient capital in the fund to generate adequate financial return). The programme continues to rely heavily on cash support and donations from Australia and New Zealand (Abila, 2010), and in-kind and staff support from the OIE, European Union and France, amongst other donors.

\[17^\text{In 2010 China officially became a part of the programme, and the name has subsequently become the South-East Asia and China foot and mouth disease campaign (SEACFMD)}\]
2.4. SOURCES OF OUTBREAK DATA

Forman and Leslie (1997) identified clearly that the implementation of widespread control programmes was of doubtful value in light of the then poor understanding of the disease prevalence and economic impacts. Through an investigation in Laos, Cambodia and Vietnam they identified the following core activities as a priority for regional coordination:

- Epidemiologic and economic studies
- Diagnostic laboratory support
- Training and education
- Disease control activities
- Development of detailed national plans

Although recognising that all three countries would require support in the form of external funds, the report did not discuss the importance of high-level political will to commit to large scale disease control programmes. As well as these, a major focus of RCU activities has been on strengthening disease surveillance and reporting from member countries, including a focus on transparency and consistency. Underreporting is a well recognised problem for analysts when it comes to reporting endemic disease. Internationally there is no requirement to report routine outbreaks of endemic disease in any way other than a single six monthly or annual figure to the OIE\(^\text{18}\) (Ben Jebara and Shimshony, 2006; Sumption et al., 2008; World Organisation for Animal Health, 2010a).

The RCU is improving consistency by maintaining a database of FMD outbreaks reported by participating countries. Since 2007 the management of this database has been integrated with ARAHIS. As mentioned above, the RCU has invested substantial effort in improving reporting consistency. To effect this they have developed a case definition for an outbreak of FMD to be reported to the SEAFMD programme office (Bouchot, 2010). The definition is paraphrased below as the analysis of data collected by SEACFMD forms a major part of this work.

\(^{18}\)In fact, the OIE encourages but does not require monthly reports.
“A foot and mouth disease outbreak is the occurrence of FMD in one or more animals in a farm, or village, or group sharing a common area (e.g. pastureland, watering point, slaughterhouse, market etc.). All cases occurring within 2 weeks of the previous case are considered as part of the same outbreak.”

This definition has attracted some criticism for lacking specificity, and most certainly would not be suitable for a disease eradication campaign. For example, it should contain more detail about what clinical signs or laboratory confirmation is required to define ‘the occurrence of FMD’. However, it is suitable for managing a control programme over the wide range of cultures, animal husbandry styles and resource capacities of the programme member nations.

The data from the ARAHIS is the only coherent and consistent data on FMD outbreaks available across the GMS. For the purpose of this thesis, it will form the basis for identifying the risk of animal infection prior to undertaking a movement.

2.5 Livestock Movement

During the 1990’s Thailand and Malaysia were the major consumers of beef in mainland SEA, with consistent flows of cattle into Thailand, and south to Malaysia from Cambodia, Laos and Myanmar (Forman and Leslie, 1997; Gleeson, 2002). Various reports demonstrated likely spread of disease along these lines, including Perry et al. (2002a) who reported the movement of disease from the Eastern Districts of Savannakhet province in Laos in September 1999 into the Western districts by the end of the same year. This trade pattern was believed to be fairly constant and contributed to the proposal for multi-country zones of FMD control in the Upper and Lower Mekong areas and the Myanmar-Thailand-Malaysia region (Edwards J, personal communication, 2006).

One of the greatest limitations to understanding contemporary movement is the acquisition of recent data on which movements have occurred. To overcome this
2.5. **LIVESTOCK MOVEMENT**

limitation, a number of different types of movement data, and a number of different approaches to collection of movement data were implemented during this work. A number of countries (and regions) have implemented animal tracing systems; The United Kingdom, Switzerland the the European Union (Dubé et al., 2009; Gilbert et al., 2005; Mitchell et al., 2005), Italy (Natale et al., 2011) and a number of South American countries (Brazil, Argentina, Chile and Uruguay) (Stevenson et al., 2007). A comprehensive, efficient and cost-effective example is Australia’s National Livestock Identification System (NLIS) (Meat and Livestock Australia, 2010). Since July 2005 it has been mandatory in Australia to record the movement of cattle using the NLIS (Department of Agriculture and Food Western Australia, 2005). In WA there is an exception for animals moving directly from their property of birth to live export or slaughter. Development of the NLIS began in response to increasing pressure to be able to certify beef for import to the European Market as free from chemical residues. This was dependent on the ability to implement a whole of life tracing mechanism. The system for tracing animal movements at the time relied on wrap around tail-tags, and the paper trail of National Vendor Declarations, which may not have been complete, may have been lost or may have been time consuming to trace (Department of Agriculture and Food Western Australia, 2006). The system in its current form was introduced in 1999 (Meat and Livestock Australia, 2010). The system retains the concept of a Property Identification Code (PIC), but each animal carries a unique code in the form of a tamper-proof ear tag or rumen bolus with an embedded radio frequency identification device (RFID). This code is used to link the animal to premises during its lifetime in a national database. This database is managed on behalf of industry and government by Meat and Livestock Australia. State and territory governments have enacted legislation requiring livestock producers to record movements on the database. Access to the database is limited to registered users, normally the managers of premises involved in cattle production, such as farms, saleyards, abattoirs and export facilities. The use of electronic identifiers at an animal level allows analysis of the full number of movements at a very high level of resolution, both spatially (between premises) and
temporally (one week).

In the GMS, there is no NLIS to use. In both Cambodia and Laos permits are required to move livestock outside their current province\textsuperscript{19}. For domestic export (other provinces in the same country) these are issued by the provincial animal health office, normally based on the presentation of a request and suitable documentation, and the payment of a fee. For international export (other countries) an export permit may be the responsibility of the national level animal health office. These may be issued in advance to exporters and nominate a number of livestock that may be exported. The exporter/trader is then free to collect animals and export up to the number on the permit within the period specified on the permit. Copies of movement permits issued by offices are normally retained in the office for variable periods (Nget K, personal communication, 2009).

There are a number of different ways of approaching the description of movements and all can contribute to presenting a picture of the nature of livestock trade. The individual records of movements between provinces does not provide a great deal of information about the role of the province in livestock movements across the GMS. Although demographic information is available on the provinces (such as population, urbanisation etc), of more interest is the relationship between provinces. Network analysis provides a formal framework to aid and test understanding of the relationships in complex systems (Martínez-López et al., 2009b; Wasserman and Faust, 1994). The units of interest are normally individuals, groups or locations of interest (known as actors or nodes), although in some cases information about the behaviour and structure of the network is more important (Martínez-López et al., 2009b). Although increasingly used in the social sciences, mathematics and statistics since Moreno described the sociogram during the 1930’s (Wasserman and Faust, 1994), network analysis has only relatively recently been applied to veterinary epidemiology (Bigras-Poulin et al., 2006; Dubé et al., 2009; Robinson and Christley,

\textsuperscript{19}There are a substantial number of other permissions that may be required to sell animals in each country, including permission from the village or commune chief and sanitary certificates from a district veterinary office.
2.5. LIVESTOCK MOVEMENT

A number of excellent recent reviews (Dubé et al., 2009; Martínez-López et al., 2009b) provide information on network analysis terminology and specifically its application to epidemiology.

By specifying that a disease is spread by contact, useful insights have been obtained from these studies (Ortiz-Pelaez et al., 2006). Network analysis may be particularly valuable for identifying critical transmission points (and ultimately critical control points) in the animal movement pathway by not only characterising the source, spread and ultimate range of a disease outbreak, but also identifying nodes with particularly high connectivity (Dubé et al., 2011).

Another option being used for understanding livestock movements, but not utilised here is working with traders to obtain the information directly. This approach has some limitations as commercial information is rarely shared readily with outsiders, and it takes time to locate traders and obtain their trust.

2.5.1 Comparison of Euclidean and Road Distances

An important aspect of all spatial analysis is the distance between any two or more events of interest. Also important in spatial analysis, perhaps more than in other statistical fields, is the issue of independent observations, often summarised as “Everything is related to everything else, but near things are more related than far things” (Tobler, 1970 cited in Waller and Gotway, 2004). Many of the principles associated with spatial statistics have developed in the field of geology, and are particularly targeted at measuring the distance between primarily stationary events, for example varying concentrations of mineral deposits. These principles and associated techniques have been extended into ecological, economic and health analyses.

While there are robust statistics to analyse these distances, there is less questioning associated with how to measure the distance between these events. Most standard spatial analyses used in geology, health sciences or economics rely on either the euclidean distance between events—being the straight line between two points—or the
CHAPTER 2. REVIEW OF RELEVANT LITERATURE

contiguity—the comparison of factors in adjacent areas.

From a health perspective these approaches are generally adequate for investigations of disease spread and modelling between wild or free living animals such as birds or ocean fish (Doran and Laffan, 2005; Gilbert et al., 2008). When applied to human endeavours, they can suffer some limitations. Firstly, very few human directed activities involve travelling in a straight line (apart from small-scale contacts, such as may occur in school playgrounds, sporting fields or paddocks). Medium-scale movements almost invariably rely on using transport networks such as road and rail with the distance to be travelled often considerably greater than that of the straight line between the start and end points. Large-scale movements (for example, airplane and shipping) often follow a geodesic—part of a great circle (the straightest line possible on the surface of sphere), which is rarely accounted for during measurement and analysis (Zuur et al., 2010). Use of a straight line, other than in short distance movements may have significant implications for misclassification bias and the reduction in power of the statistical test (Berke and Waller, 2010; Waller et al., 2006). Nonetheless, the use of euclidean distance is commonplace in epidemiological studies (Bigras-Poulin et al., 2006; Lindström et al., 2009; Robinson and Christley, 2007), although a number of studies have recognised that factors other than straight lines need to be taken into account, for example rivers and mountains (Bessell et al., 2008; Savill et al., 2006) and transport of infected animals (Farnsworth and Ward, 2009). Ward et al. (2008) suggested a comparison using road distances instead of euclidean would be valuable.

Often, particularly with aggregated data, rather than the geographic distance between events, it is the notion of adjacency that is more appropriate. Although adjacency is normally referring to contiguous borders between neighbours, Pfeiffer et al. (2008b) discuss its extension to neighbours of neighbours (second order adjacency). An alternative approach uses distance to select the $n$-nearest neighbours (Bivand et al., 2008; Waller and Gotway, 2004). Disease spread is related to the probability of contact occurring, and so does not only occur at local scales, but also due to long
2.6. LIVESTOCK PRICES

distance movements. Network analysis provides a framework to extend the concept of adjacency to include the probability of contact between susceptible animals that are not geographically adjacent.

There is also an increasing interest in how far apart events are, not only whether susceptible animals come into contact with infectious ones. This may better be measured by the use of network analysis, using concepts such as latent space (Handcock et al., 2007), where the relatedness of nodes is represented in the estimation of the location of the node in euclidean space. Some recent work has demonstrated the valuable contribution of such methods to understanding outbreaks in both endemic mixing and epidemic scenarios (Firestone et al., 2011; Martínez-López et al., 2009a).

2.6 Livestock Prices

The ongoing collection of animal movement information may be difficult and potentially expensive. Additionally, where a large number of movements may be undertaken that do not comply with the requirement to notify the government authorities and obtain permits it is unlikely to provide a high quality set of data. In order to overcome a lack of data the use of a proxy source of information is a common solution.

The livestock market in the GMS satisfies many of the criteria for a market in which perfect (or total) competition exists—large numbers of producers and consumers providing product with little differentiation into a market with little or no restriction on entry\(^{20}\) (Sloman and Norris, 1999). The widespread adoption of the mobile telephone has even allowed for rapid sharing of information on livestock prices. In this scenario the value of an animal will closely reflect the cost of production plus the cost of transporting animals to market and is referred to as its natural price (Smith,

\(^{20}\)In some aspects of the livestock industry in some countries there is not unrestricted entry due to regulatory constraints—but for the large majority of participants (farmers and consumers) this is an unimportant impediment.)
According to Smith, any person putting that effort in must at least cover his or her own costs in terms of wages and subsistence, often termed profit. Where the market price allows no opportunity for covering this expense it is unlikely that the person will continue to supply stock to the market.

This incentive to make a profit, in the context of perfect competition, allow us to postulate that livestock trade will follow the price gradient from lower priced areas to higher priced ones.

Livestock valuation is a highly complex process taking into account a large number of variables. For example, the National Livestock Reporting Service (NLRS) in Australia employs 21 trained livestock market officers on a full time basis to attend regular cattle markets (Meat & Livestock Australia Limited, 2005b). These officers record the price of animals being sold and a range of other parameters classifying the animals. These classifications include total number, age, sex, type (dairy or beef), liveweight (actual or estimated), body condition, muscling, estimated dressing percentage. They also record details such as market destination, feeding background and pregnancy status (if appropriate) (Meat & Livestock Australia Limited, 2005a).

The methods used for valuing livestock in the GMS vary widely across the region. A per head price for an animal or group of animals is often quoted but this often masks the large range of factors being considered in arriving at this price. Per head prices are normally quoted where the animal is likely to be valued as a whole, such as young animals for growing, breeding, draught and fighting animals—it is also commonly used for piglets for consumption. These prices are often based on local market prices. For slaughter animals the situation is more complex, and although prices may be quoted for the entire animal (per head) they are normally a reflection of the price per kilogram. These may be quoted as either per kilogram liveweight (more common in Malaysia) or per kilogram expected meat yield (more common in Cambodia, Laos and Vietnam). These prices per kilogram tend to vary by area, but may be negotiated locally under certain circumstances (Scoizec, 2009).
important in determining the price of an animal is the evaluation of yield. This requires the ability to evaluate potential yields of meat, second meat (non-premium cuts), offal, hide and bones and knowing what prices they can expect to receive for each commodity (Bourgeois-Lüthi N, personal communication, 2009).

It may be this difficulty in describing the commodity that has limited the development of market information systems for livestock. Agricultural market information systems have been under development in both Cambodia and Laos during the last few years (Food and Agriculture Organisation of the United Nations, 2009). Similar examples exist for other commodities globally\(^{21}\). However the sustainability of the systems is yet to be established, with systems such as the LaoTrade Market Information System, which was established in 2009, no longer available\(^{22}\). On a more positive note, the Cambodian Agricultural Market Information System\(^{23}\) (CAMIS) appears to have evolved and provides a comprehensive service with data collected by dedicated marketing officers in 13 provinces. From a livestock perspective CAMIS is limited to pigs and poultry, but it also provides information on a wide range of horticultural produce.

A number of other variables may be assumed to be likely to influence price but are difficult to manage objectively when assessing costs. These include direct monetary costs such as official government permits and charges, unofficial fees (bribes) and cost of transport and quarantine holding charges; and indirect costs such as interest on money borrowed (often at high short term rates); and overhead costs such as trading licenses. Wilson and Xayaleth (2009) have done some work on quantifying these charges in Laos—many of them appear to be highly variable both by location and over time. Trade is often conducted using money borrowed at very high short-term rates (Sieng S., personal communication, 2009). These rates are high enough that delays of more than a few days could turn a profit into a loss. This has important implications when attempting to evaluate the opportunity costs associated
with regulatory delays such as 10–14 day mandatory quarantine for animals moving through official import channels.
“We cannot emphasise too strongly the need for farmers to report suspi-
cions of foot-and-mouth disease immediately. We realise that many of
these reports would not in the event be confirmed . . . . On the other
hand during the epidemic there were several cases where suspicions
should have been aroused earlier and where prompt reporting might have
limited the spread of disease. When the county is free of foot-and-mouth
disease for long periods there is a danger that farmers, and those veter-
inarians who have had little or no experience of foot-and-mouth disease,
may be slow to recognize the early signs of the disease.” (Northumber-
land et al., 1969)
3.1 Introduction

The under-reporting of infectious disease outbreaks is a serious problem for both human and animal health worldwide—for example, Rabies (Curtis, 1999); Highly Pathogenic Avian Influenza (Gilbert et al., 2008; Balcan et al., 2009); West Nile Virus (Sugumaran et al., 2009) and Q-Fever (Schrödle and Held, 2010). Unsurprisingly then, this is also true for foot and mouth disease (FMD) (Barnett and Cox, 1999; Ekron and Weaver, 1999). For a disease event to be captured in a reporting system requires a sequence of events to occur (Cameron, 1999, p. 15), and failure of any of these will mean no report is recorded. The current situation with FMD in South East Asia (SEA) is characterised by an endemic virus causing sporadic outbreaks of disease often associated with the movement of subclinically infected livestock (Perry et al., 2002a). These conditions create a situation where disease reporting is unlikely to occur, as although disease may be inapparent or is less likely to cause severe disease (especially when compared to other endemic diseases such as Haemorrhagic Septicaemia), its regularity may make owner’s unlikely to bother reporting. As it is unlikely that there will be any benefit to reporting, there is little incentive for the owner, or the village animal health worker or the veterinary staff to report. In this case, the sensitivity of the reporting system is likely to be less than optimal.

As well as sensitivity, system accuracy in the context of a diagnostic system also requires an understanding of specificity. Confirmation of disease agent by a diagnostic test (usually in a laboratory) is a vital component of an accurate reporting
system, as clinical signs do not have the degree of specificity required for many diseases.

In the case of animal diseases, individual country status is based on disease reporting to the World Organisation for Animal Health (OIE) by member countries that is not (and never has been) formally evaluated for reporting accuracy (Garabed et al., 2009). As previously discussed, in SEA outbreaks of FMD are reported by countries to the regional animal health information system (ARAHIS). This process is managed by the South East Asia and China foot and mouth disease programme (SEACFMD), a programme of the OIE Regional Coordination Unit (RCU) based in Bangkok.

To better understand the risks associated with the movement of livestock due to trade, historical data on outbreak locations are needed to inform a risk model of disease spread in SEA. As well as information on the movements of animals, the model required information on the risk of animals being infected either prior to movement or during the course of movement. The objective of this work was to evaluate the reporting system for FMD in SEA to provide some indication of the accuracy of the data as an input for the model.

Given the incomplete nature of the data, a number of different techniques were used to evaluate the quality of the reporting system and identify locations where data was of questionable quality, as well as build a picture of the epidemiology of FMD in SEA. The resulting information can be used to develop policy and guide activities targeted at improving the reporting system by identifying weakness in both the process and identifying locations where under-reporting is likely to be inhibiting efforts to reduce the impact of disease across the region.
3.2 Materials and Methods

3.2.1 Study Population

The SEA mainland region comprises seven countries: Cambodia, Laos, Malaysia, Myanmar, Singapore, Thailand and Vietnam. During the study period, no outbreaks of FMD were reported from Singapore. All of the other countries reported outbreaks. In all cases the location of the outbreak could be identified to at least the first administrative level within the country (province—or in the case of Myanmar—state or division). In many cases only the number of affected animals was available, and there was little information at the village level on the population at-risk. This is important because outbreaks are—by the SEACFMD definition—reported at the village level. Data on the number and location of villages was obtained from local sources, including gazetteers where available and affordable\(^1\). However, because there was no data available for some countries, for consistency it was decided to use a single source covering the entire region (United States Board on Geographic Names, 2010).

3.2.2 Outbreak Data

Since 2007, SEACFMD members have entered their outbreak data directly into ARAHIS. Prior to this, reports were submitted to the RCU in a number of formats and entered into a Microsoft Access database by RCU staff. To facilitate the use of the online reporting systems available through ARAHIS some existing data from the Access database was also imported into ARAHIS when this new system was commissioned. For the purposes of this investigation, the data from ARAHIS was downloaded from the SEACFMD website\(^2\) and merged with data available in the SEACFMD database. The data were imported into a relational database (PostgreSQL 8.4) with a spatial extension (PostGIS 1.5) and a procedural language (PL/R)

\(^1\)Not all countries were prepared to make such details available to non-citizens.

allowing the use of the R statistical environment (R Development Core Team, 2011) from within the database.

This study only evaluated mainland outbreaks, so the Philippines outbreak data were removed. Any duplicate outbreak reports (often a function of ‘follow-up’ reports relating to a previously notified outbreak) were also removed.

Where administrative boundaries had changed during the study period a customised matching tool was used to recode the likely outbreak location to best match the current administrative location\(^3\). Although ARAHIS allows accurate recording of outbreak longitude and latitude, these data were not available for 4 098 (78%) of the reports. To overcome this lack of data the centroid of the smallest known administrative division (district or province) was used. These data were matched with a list of currently known locations from the same database. Although data were stored in English in many cases it was difficult to match names of provinces and districts using simple matching (due to transliteration issues). To overcome this a phonetic matching technique known as Metaphone (Philips, 1990, 2000) was used. In the majority of cases it was possible using this technique to identify at least the district (second administrative level) referred to in the report. The remainder were manually identified to the district level by individual matching, and using tools such as web-based mapping systems\(^4\).

Where longitude and latitude were reported, but inconsistent with the reported administrative district (for example, a point north of the Maldives that should have been in Myanmar) the centroid of the district was used in preference. Using the centroid as an approximation may have caused some nondifferential misclassification bias. Based on district diameters this is not likely to be very serious—for 90% of reports the centroid will be less than 39 km from the actual location.

Serotype information was not available for all reports. Where outbreak serotype was recorded as ‘not sampled’, ‘not typed’ or ‘pending’ it was coded as ‘unknown’.\(^3\) The derivation of the spatial data used in this work are describe in Section 4.2.2.\(^4\) http://maps.google.com/
Where information on affected species was available and included number affected, number of deaths and population at risk then morbidity rates and case fatality rates were evaluated.

### 3.2.3 Phylogenetic Analysis

Information on disease outbreak topotypes was accessed from the World Reference Laboratory (WRL) website\textsuperscript{5}. Complete (633 base pairs) VP1 sequences were obtained from GenBank. The information available from this resource was detailed, high quality information about the virus and entries included information on country of origin, date collected, species, VP1 sequence and homology with other known viruses. However, it lacked any kind of identifier which could readily be used to link the laboratory analysis results with the epidemiological information available through the SEACFMD and ARAHIS databases.

Only a small number of disease reports resulted in samples being submitted for typing and so there was a limited amount of information that could be gained using the available data and there was a substantial risk of selection bias. To illustrate this a simple assessment of homology between countries was carried out by assessment of the genetic difference between outbreaks. A midpoint rooted neighbour joining tree was created from the VP1 sequence using Kimura’s 2-parameter distance, as implemented in the R package ape (Paradis et al., 2004). The distance matrix was also used to display isolate locations using multi-dimensional scaling (MDS) to provide a clear indication of the distance between groups.

### 3.2.4 Temporal Analysis

The temporal pattern of the reports was assessed using a number of methods to determine if there was a seasonal component and to evaluate if it were possible to identify larger than normal numbers of concurrent outbreaks. This was done using

\textsuperscript{5}\url{http://www.wrlfmd.org/}, accessed 2009-10-09, 2010-03-18, 2010-11-03.
simple graphical techniques (normalising the number of outbreaks by the day of the year and multiplying up to 500 outbreaks per year to allow comparison across the years) and then seasonal decomposition by fitting a locally weighted least squares model using monthly outbreak totals (using loess implemented in the stl function in R) to look for seasonal trends (Cleveland et al., 1990).

An implementation of cluster detection methods was developed for the SEACFMD programme. This was a web based system using routines from the R surveillance package (Höhle, 2007; Höhle et al., 2009) linked to a copy of the SEACFMD database using the PL/R procedural language. The output included a straightforward weekly outbreak graph (an example output is shown in Figure 3.1) as well as graphical output from a number of different detection algorithms; methods utilised included CUSUM (Rossi et al., 1999), log-linear regression (Farrington et al., 1996), generalised likelihood ratio (Höhle et al., 2009) and Bayesian (Reibler, 2004 in German, cited in Höhle, 2007) algorithms provided through the R package Surveillance (Höhle, 2007). Although these were pre-populated with sensible default values, the user was able to adjust various parameters and repeat the analysis and reproduce the outputs. The sensitivity and specificity of these algorithms have not been formally evaluated for their usage in the FMD programme, and will not be discussed further here.

### 3.2.5 Spatial Analysis

A number of methods for spatial clustering were performed to examine whether or not disease reporting was clustered. Although it is generally accepted that a highly infectious disease such as FMD will exhibit clustering, it may not be sensible to extend this assumption to the reporting of disease, particularly across a large multi-jurisdictional area such as SEA. The presence of clustering by serotype under the assumption of an isotropic and stationary environment was assessed using summary statistical functions. Tests used included a nearest neighbour distance function (G), the empty space function (F), a combination of these (J) (van Lieshout and
Figure 3.1: Temporal distribution of reported outbreaks of foot and mouth disease in South-East Asia per week for Type O serotype. The figure is derived from the online outbreak detection system and shows the analysis range for the cluster detection algorithms, and indicates where the number of reports a week exceed a user defined number (green crosses).

Baddeley, 1996), as well as Diggle & Chetwynd’s second order K test (Waller and Gotway, 2004).

To identify the location of disease point clusters it was necessary to take into account the heterogeneity of the village population. Baddeley and Turner (2005) describe an implementation of a maximum pseudo-likelihood method to allow the evaluation of a parametric model for point processes. Using this technique, an inhomogeneous Poisson model was fitted using the average number of villages per unit area—known as village intensity.

The impact of disease at a provincial level was described by generating standardised morbidity ratios (SMR) for reported FMD outbreaks (Thrusfield, 1997; Waller and Gotway, 2004). As the FMD outbreak reports were reported by groups in many cases approximating a village, outbreak reports were standardised for each province.
using the number of villages in each province. To do this, first the expected num-
ber of outbreaks in each province was calculated by dividing the total number of
outbreaks in a year by the total number of villages in the region, and then mul-
tiplying the result by the number of villages in each province. The SMR was then
the observed number of outbreaks for the year divided by the expected number for
the same year. An example of the R code for this process is in Appendix D.1. The
confidence intervals of the SMR for each year were obtained using a Poisson model
(Aragon, 2010; Stevenson et al., 2010). When displayed on a map, SMRs may give
a distorted impression of disease occurrence due to small numbers of observations or
large populations (Pfeiffer et al., 2008b). To obtain improved estimates of the SMRs
and take into account overdispersion and autocorrelation, an alternative estimator
using a log-normal model of relative risk was used. This produces empirical Bayes
estimates using an approximation of the expected values (Clayton and Kaldor, 1987;
Bivand et al., 2008). The number of years each province had a SMR >1 was also
calculated and displayed as a choropleth map as these provinces are likely to be
important contributors to disease spread through livestock movement.

Choynowski (1959) demonstrates a method where instead of displaying the actual
incidence of disease (in his case Brain Tumours), the probability of the result being
correct is displayed. An assessment of reporting was undertaken by using a max-
imum likelihood estimate to parameterise a negative binomial model. A probability
map was generated (Fig. 3.8) to highlight areas where disease reporting was less than
would be anticipated if disease were evenly distributed across all provinces.

3.2.6 Spatio-temporal analysis

Spatial and temporal cluster detection was attempted using a discrete poisson spatio-
temporal scan (SaTScan™) (Kulldorff, 1997).

The contagious nature of FMD means the likelihood of clustering and the identi-
fication of cluster locations is important. The SaTScan test allows detection of the
3.2. MATERIALS AND METHODS

location of the most-likely cluster and any other significant clusters—each cluster is
determined independently (Kulldorff, 1997; Kulldorff et al., 1998).

Reports are based on the SEACFMD definition of an outbreak, so it was inferred that
the unit at-risk is a village. A retrospective space-time analysis using the discrete
Poisson model was undertaken. The case file (province, date and number or reports)
was generated from the ARAHIS database. The population file (province, date and
population) was created using the Postgis ST_Within() function to determine the
number of villages within each province. The coordinates file (province, latitude and
longitude) was obtained by calculating the centroid of the province polygon. Each
serotype group was analysed individually. The untyped reports were also analysed
as a single group. All data sets were analysed using a three week aggregation.

This scale was chosen because it reflects the median time spent infected based on
Carpenter et al. (2004). Due to the highly contagious nature of FMD and the low
likelihood of survival on fomites in the study region a relatively small cluster size was
chosen. Standard parameters for the analysis included a maximum spatial cluster
size of 10% of the population (or a 250km radius). Purely temporal clusters or
those that covered the entire region were ignored as they had already been assessed
using other methods. The maximum temporal cluster size was set at 90 days, being
the maximum modelled length of animal infectiousness, again based on Carpenter
et al. (2004). Under the assumption of local spread by animal contact, no particular
benefit was anticipated using elliptical clusters, so circular clusters were specified.
Clusters were allowed to overlap, but not such that both clusters contained the
other’s centres. No cluster was allowed to contain >10% of the total population.

These constraints may produce the counterintuitive visual impression of clusters not
being centred on a group of outbreaks. A significance level of 0.05 was used for all
clusters, and 9999 iterations were used to calculate P-values. Results files were saved
as ASCII text files and imported into the database using Sed (McMahon, 1974). The
process of geocoding the original data resulted in some clusters being reported where
a number of outbreaks all occurred at exactly the same point location (the district
or province centroid). This meant the cluster had a radius of 0, and would not be
plotted. To overcome this anomaly, the geometry of the appropriate location (for example, the province boundary) was used instead of a circle. Across SEA this allows the display of these clusters, albeit with the risk of the visual perception of a non-circular scanning window. The outbreaks and significant clusters were then clipped where they extended over the ocean and displayed using Quantum GIS (Quantum GIS Development Team, 2009).

3.3 Results

From the start of January 2000 to October 2010, 5,237 outbreaks of FMD were reported to the SEACFMD RCU or to the ARAHIS. The breakdown by serotype is shown in Table 3.1. Of these reports, the most prevalent serotype was Type O (48%), then Type A (9%). A small number of Type Asia 1 cases were also seen (<1%). Perhaps the most notable outcome is the large proportion of reports where no serotype result was reported (43%). This has important implications for the specificity of the disease reporting system. The number of outbreaks by year for each serotype is shown in Figure 3.2. The number of cases involving Asia 1 was very small and this serotype has not been reported to SEACFMD since 2005. Serotype Asia 1 was reported to the World Organisation for Animal Health (OIE) by Vietnam in 2006–07 because it involved a change in the epidemiology of an endemic disease, but it was not submitted to the regional system.

<table>
<thead>
<tr>
<th>Serotype</th>
<th>Cambodia</th>
<th>Laos</th>
<th>Myanmar</th>
<th>Malaysia</th>
<th>Thailand</th>
<th>Vietnam</th>
<th>Total</th>
</tr>
</thead>
<tbody>
<tr>
<td>A</td>
<td>0</td>
<td>12</td>
<td>1</td>
<td>18</td>
<td>334</td>
<td>92</td>
<td>457</td>
</tr>
<tr>
<td>Asia1</td>
<td>0</td>
<td>0</td>
<td>6</td>
<td>0</td>
<td>0</td>
<td>2</td>
<td>8</td>
</tr>
<tr>
<td>O</td>
<td>3</td>
<td>103</td>
<td>119</td>
<td>394</td>
<td>353</td>
<td>1,536</td>
<td>2,508</td>
</tr>
<tr>
<td>Unknown</td>
<td>411</td>
<td>618</td>
<td>129</td>
<td>118</td>
<td>716</td>
<td>272</td>
<td>2,264</td>
</tr>
<tr>
<td>Total</td>
<td>414</td>
<td>733</td>
<td>255</td>
<td>530</td>
<td>1,403</td>
<td>1,902</td>
<td>5,237</td>
</tr>
</tbody>
</table>

| (%)     | 99       | 84    | 51      | 22       | 51       | 14     | 43    |

Table 3.1: Reports of foot and mouth disease by serotype and country in South-East Asia from 2000–2010.
3.3. RESULTS

Figure 3.2: Report of foot and mouth disease virus serotypes in South-East Asia, by year from 2000 to 2010. Asia 1 outbreaks occurred in 2000, 2001 and 2005. The horizontal line shows the median number of outbreaks during the period.

The number of outbreak reports per week is shown in Figure 3.1. The large number of outbreak reports (1 782) during 2006 were mainly from Vietnam (1 326) and related to a large number of cases of type O FMD.

Seasonal decomposition of the outbreak times suggested a seasonal pattern with peaks in May (decreasing in June and July) and December. There was strong influence in the trend corresponding with the very high number of reports from Vietnam during mid-2006. Removing the extreme monthly totals recorded during this period reduced the effect on the seasonal data and the model residuals. To reduce the impact of the single group of outbreaks, reports to SEACFMD from 2000–2009 were normalised based on 500 outbreaks per year for each year of the study period. When plotted (Figure 3.3) they showed no apparent seasonal trend in reporting.

Details of the number of animals affected and number of deaths was investigated. Only 1 368 of the 3 139 (44%) reports since 2005 had information on the population
at-risk, the number of cases and the number that died. The age of affected animals was not available. As a result of the missing data the opportunity for robust statistical analysis was limited. The mean reported morbidity rate was 17%, and mean case fatality rate (CFR) of 1.4%. Although FMD is normally associated with a low CFR, in the data available there was a relatively high CFR (>5%) in 40 reports. In 23 of these reports, the serotype was unknown, including the highest CFR (44%).

The VP1 sequence data from GenBank were used to create both a dendrogram (Figure 3.4) and a network diagram using MDS (Figure 3.5) showing the relatedness of genotypes associated with different outbreaks of Asia 1 virus in the region. Only limited inference may be made due to the constraints associated with the selection of isolates. From this analysis it appears that separate outbreaks (homology <95%) occurred in Myanmar in 2005 and 2006. The 2005 Myanmar outbreaks may have been more closely related to the Asia 1 virus reported in China during 2005. The
2006 outbreaks were more closely related to the outbreaks in Vietnam in 2005–06.

Tests for spatial clustering of disease reports were consistent with clustering occurring within the region. The best fit for the inhomogenous Poisson model included the village intensity as an offset, and the longitude and latitude of the reports. With the village intensity accounted for, there was still increased reporting in the eastern part of the region (Vietnam) and the Southern Thailand/Malaysia border (Fig 3.6). As with the seasonal decomposition, it is likely that the large number of reports from Vietnam during 2006 would have impacted the model output.
Standardised Morbidity Ratios were calculated using the empirical Bayes method and the provinces which had an SMR >1 in 5 or more years during the study period are shown in Figure 3.7.

A choropleth map displaying the likelihood of a province not reporting disease is shown in Figure 3.8.

Spatio-temporal clusters were detected using the SaTScan discrete Poisson method (Kulldorff, 1997). SaTScan identified significant (p < 0.05) clusters for Type A and O serotypes during the period investigated. A further 55 significant clusters of
Of the 43 clusters detected for Type O (Figure 3.9a), 21 were associated with increased reporting during 2006 and were detected in all countries except Thailand. Eleven clusters included only one province, however the 13 clusters associated with the large number of reports from Vietnam in 2006 included more than one province. There were three clusters of reports from northern and eastern Thailand in 2001–2002, then 13 from Laos (Vientiane and Bokeo), Malaysia, Myanmar and Thailand in 2003. In 2004 only one cluster on the eastern Thailand/south Laos border was
Figure 3.7: Provinces in South-East Asia where reported outbreaks of foot and mouth disease empirical Bayes log-normal estimates were >1 for five or more years during the period 2000 to 2010.
Figure 3.8: Choropleth map showing the likelihood of underreporting of foot and mouth disease in South-East Asia during 2000 to 2010.
detected. Since 2006 the only clusters identified were associated with reports in late 2007 along the Thailand/Malaysia border (Kelantan). Type A FMD showed slightly less tendency to clustering: 33 clusters were detected, three of which included only one province. The Type A clusters were distributed earlier in the study period, with 5 in mid-2002 in northern Thailand, 17 during the second half of 2003 covering much of the region, three in late 2004 and one in mid-2005 in South Vietnam (north of the Mekong Delta), one in late 2006 around Vientiane in Laos and 6 in Central Vietnam in 2009. Clusters of reports in which no serotype was recorded were also examined. A number of these clusters are of interest because of temporal and spatial overlap, including 6 in southern Laos in 2003 and three on the Myanmar/Thailand border, 5 in late 2004 in southern Thailand and northern Malaysia and 7 in north Vietnam during the second half of 2009. A final observation was that the numbers of clusters of unknown serotype was in most years not dissimilar to the numbers of clusters focussed on type O or type A. To illustrate, a group of three clusters (involving 107 reports) in Cambodia during 2006 is of special interest, as all three overlap with Type O clusters centred in Vietnam during the same time (Figure 3.9b).
Figure 3.9: Distribution of reported foot and mouth disease outbreaks in South-East Asia during 2000 to 2010 showing (a) locations of significant clusters of type O FMD ($p < 0.05$) detected using SaTScan, and (b) where clusters of unknown disease occurred at the same time as—and within or immediately adjacent to—a larger group of type O clusters during 2006.
3.4 Discussion

The need for this investigation was driven by a project to develop a risk model of disease (FMD) spread associated with livestock trade in the Greater Mekong Subregion. Very few published analyses exist on FMD with a regional focus on SEA, thus it was important to describe the recent disease situation and identify where (and when) outbreaks might be occurring. Reports from the regional system were analysed and the quality of the reporting system critically evaluated. The sensitivity and specificity of the reporting system is an important consideration when evaluating disease reports. In total, 5,237 outbreak reports submitted to the SEACFMD programme over 10 years were analysed. The finding that in 43% of reports serotype was unknown has important implications for the quality of the overall system. In particular, the specificity of the disease reporting system becomes suspect and raises broader questions about the overall data quality. Some countries had very few laboratory confirmed cases. Because FMD is rarely associated with high levels of mortality, the apparent high case fatality rates estimated in situations in which the serotype was not confirmed raises questions about the validity of the diagnosis. Discussion with provincial staff during this study raised concerns about the validity of clinical diagnostic skills. The personal experience of the author has included hearing reports such as “the bad FMD—where the cattle get sick and die within 24 hours with swelling underneath [ventral oedema], and the meat smells bad when you eat it”.

Under these circumstances, ARAHIS reports may be regarded as indicators of disease, but not necessarily FMD. Advances in decentralised testing techniques using Lateral Flow Devices and portable RT-PCR offer hope for improved field verification of FMD as the cause of disease outbreaks (King et al., 2008; Sammin et al., 2010).

The tendency to describe FMD as endemic in mainland SEA misrepresents the sporadic nature of the disease across the region (Perry et al., 2002a). This may
create some problems with interpreting reports of FMD in the classical framework of ‘outbreaks’. In this context, the case definition used by the SEACFMD programme provides a useful (although arguably vague) approach to interpreting the basis on which outbreaks should be reported. However, in many provinces endemic disease may result in under- or over-reporting depending on the level of enthusiasm of the responsible authority. Similarly if disease is reported to the global database and not the regional one it may not be included in analyses such as the present one. The degree of selection bias is very difficult to quantify using the data in the reporting system. Thus absence of reports may be an indication of system insensitivity rather than a lack of disease occurrence. To overcome this, the SEACFMD has recently undertaken an initiative for national coordinators to review all reports provided to the database and fill in missing details. In some situation (such as Cambodia in 2006) where reports of unknown serotype were occurring within the same temporal and spatial windows as other reports, it may be possible to tentatively classify the reports based on this association. A robust means of doing this would be useful, but should not detract from a focus on improved recording, sampling and testing during field investigations.

In this study, outbreak location was not always known by point coordinates—often it was limited to province and district. Due to an inability to obtain data on village locations from some countries, a standard source covering the whole region was used. It is possible that this source was not an accurate representation of the real situation, and may have created a misclassification bias if there is a trend towards a lesser number of villages in countries where data is not readily available. A spatio-temporal cluster detection method (SaTScan) was applied in this study and identified significant clusters of disease for Type O, Type A, Type Asia 1 and the reports where no serotype was known. There were a number of limitations that affected this approach. The first is the need to specify a cluster shape (either circular or elliptical), which may not match well the distribution of susceptible units (villages) and may affect the power of the test to detect clusters. In this case the maximum cluster size was set to 10% of the population to account for the char-
acteristics of the disease in the region and to improve the probability of detecting adjacent smaller clusters. The second is the supposition that the clusters detected are circular. Changes in geography and cultural practices will in many cases invalidate this. Alternative tests such as that proposed by Tango and Takahashi (2005) may help to improve the detection of non-circular clusters, although currently computational requirements would not allow it’s use on such a large dataset. Other tests have also been reviewed recently, including a Bayesian modelling approach (Berke and Waller, 2010). A further concern is the performance of such tests when the report specificity is low. Berke and Waller (2010) concluded that the effect of a misclassification bias was limited when the spatial distribution of the misclassification was unbiased. However, the proportion of outbreaks of unknown serotype varied widely between countries supporting further effort to confirm serotype to improve the validity of the results.

An extension of investigating outbreaks at the province level was the development of a probability map (Figure 3.8) to identify provinces where reporting may be below that which would be expected for an endemic disease. As is the case with all choropleth maps, there may be a tendency for users to infer greater importance to larger geographic areas. However, this offers an important insight into the sensitivity of the disease reporting system. Where historically effort may have focussed on provinces with high levels of FMD, this map highlights provinces where more effort may be required to ensure that disease events are being reported.

The ability to detect clusters of outbreaks as soon as possible was investigated during this work. Cluster detection methods were coded as procedural language functions in the database and made available through a web interface. Methods available through this interface include CUSUM (Rossi et al., 1999), log-linear regression (Farrington et al., 1996), generalised likelihood ratio (Höhle et al., 2009) and Bayesian (Reibler, 2004 in German, cited in Höhle, 2007) algorithms provided through the R package Surveillance (Höhle, 2007). These provided insights into the likely timing

---

of temporal clusters, and although they were not formally assessed, many of these clusters were temporally consistent with clusters identified by SaTScan.

Missing data were a problem in this study. Improvements in technology now offer systems where data can be inputted directly from the field. The increasing coverage of mobile data networks (Banks and Burge, 2004), use of digital pens (Fong P, personal communication, 2009, http://www.xcallibre.com/clients/foa.html, accessed 12 October 2009) for form entry, and smart phone based systems (Aanensen et al., 2009) offers an enormous opportunity to improve the quality of data collected. This is achieved by validation at the time of entry (for example, are the numbers at risk greater than the number affected—the longitude and latitude of the report realistic?), reducing the time from field investigation to data analysis, and especially avoiding transcription errors.

Recent work on the molecular epidemiology of FMD suggests that the understanding of serotypes is not sufficient to fully understand the epidemiology of FMD. In SEA over a five-year period there has been limited typing of outbreaks, particularly in Cambodia. This leaves a gap in our knowledge of serotype dependent spread in the centre of the region, and even raises questions about the validity of the diagnosis of FMD using clinical methods, an issue which has been previously investigated in the context of the 2001 outbreak in the United Kingdom (Ferris et al., 2006; King et al., 2008). The lack of a consistent identifier to link reference laboratory results to the reports in the regional database is a major impediment to the ability to use the database as a central component of a disease control programme.

Advances in molecular typing may provide a substantial leap forward in understanding the circulation of FMD in SEA. Most of the available sequence data is limited to the VP1 capsid protein, the high antigenicity of which makes it important for vaccine development (Mohapatra et al., 2002). Recent work utilising whole genome investigations (around 8 192 nucleotides) has been undertaken in the UK (Cottam et al., 2008). This has shown enormous potential for high resolution understanding of the spread of FMD, however this work is only just beginning in the SEA
region (King D, personal communication, 2010). The system is still very reliant on high quality virus samples being collected and submitted to a laboratory capable of conducting such testing.

In summary, both the ability to conduct comprehensive investigations, submit appropriate samples and report in a timely manner remains a constraint to the understanding of the FMD situation across mainland South-East Asia.
Chapter 4

Understanding Livestock Movement Patterns

“Logic will get you from A to B. Imagination will take you everywhere.”

Attributed to Albert Einstein (1879-1955)
4.1 Introduction

Trade in livestock is a vital part of the economy and the Greater Mekong Subregion (GMS) is no exception. Any movement of animals carries a risk of disease spread and long distance transport by truck or rail is a well known risk for the rapid geographic dispersal of disease (Fèvre et al., 2006; Gilbert et al., 2005; Keeling et al., 2001; Rweyemamu et al., 2008b). Better understanding of livestock movements is vital for the prevention of disease spread without unnecessarily restricting economically valuable trade. This understanding is particularly important for evaluating consequences and the planning and implementation of control programmes (Dubè et al., 2009; Keeling et al., 2010). A key component of planning for outbreaks is the use of models to evaluate different outbreak scenarios and the effect of alternative control strategies (Garner and Beckett, 2005). These models require input parameters such as information about the nature of the livestock industry, risk of exposure and contact structure (Garner and Beckett, 2005).

The purpose of this investigation is to answer the question of whether animal movement data could be recorded and analysed to help with understanding the contact structure and patterns associated with livestock trade in the GMS. In order to achieve this it was intended to capture information on the movement of animals within two countries in the region (Cambodia and Laos) including imports and exports. Neither country had established systems for collecting this data and data collection was thus limited to the use of movement permits issued at the provincial level for movements between provinces. As a result, the resolution of the data was likely to be limited to groups of animals, and spatially to the provincial level. It was also realised early in the research that there were likely to be insurmountable issues with collecting data from all provinces, due to the decentralised nature of the government in each country.

As a result, a concurrent research activity involving interviewing Cambodian and Laos traders was undertaken to better understand the sociological profile of these
traders, and the impact that their networks and actions might have on better understanding animal movements. While this work is highly complementary to the approaches described here, for practical reasons, this work was undertaken by speakers of Khmer and Lao languages. Analysis and reporting on this work will be done by those involved.

Due to the incomplete nature of the movement data collected in the GMS, in order to conduct the proposed investigation into the patterns of animal movements, the contact structure of animal movements in Western Australia (WA) during the first three years of mandatory notification of cattle movements was also investigated. This was done to develop and validate appropriate analytical techniques using an unusually complete set of data—virtually a census\(^1\). The GMS and WA data were then analysed using a number of methods, including classical statistical methods, displaying of the movements on a map and network analysis. Some aspects of the analysis of each set of data were integrated into a web based query system, allowing repeated analysis and display of results. The use of relative measures of network centrality allowed the identification of a number of key locations for their role in supplying and receiving livestock.

A related investigation, a trial of prospective animal tracking was done using cattle being enrolled in a population dynamics study as part of concurrent research in Northern Laos. The aim of this trial was to determine if the animals could be detected at road checkpoints or abattoirs after leaving the village in which the trials were being undertaken. If it was possible to detect these animals, it would also have contributed valuable information about the movement patterns and it may have been possible to validate the numbers of animals moved as recorded on movement permits. This investigation did not detect any tagged animals, and is reported in Appendix A.

\(^1\)For historical/political reasons some classes of cattle movements (typically direct to slaughter or live export from the property of origin) are not required to be reported to the NLIS in WA.
4.2 Materials and Methods

In order to learn more from the existing animal movement data, analysis of two different sets of movement data was undertaken. The first set was cattle movement information from WA from July 2006 through to June 2008 taken from the National Livestock Identification System (NLIS). The second set was movement data from Cambodia and Laos taken from official movement permits issued by provincial offices for interprovincial movements.

As well as data on animal movements, this study also required information on the spatial attributes of the shires from which animals were moving from and to, and the road network on which they were travelling. Source and destination were handled at the level of the shire (local government area) in WA (rather than the property code) and the province in the GMS. A number of factors influenced this decision, including privacy concerns in WA, spatial resolution of collected data in the GMS, and computational pressure in both locations.

4.2.1 Obtaining Movement Data

Western Australia—NLIS

The NLIS is managed by Meat and Livestock Australia. Data can be obtained only with appropriate permissions from the data owner, which is the state (or territory) in which the movement occurred. Some states (WA, Victoria and Queensland) maintain their own database mirrors (Hood and Martin, 2008). Because of this it was possible to obtain the data directly from the Department of Agriculture and Food Western Australia (DAFWA). A database query was developed using structured query language (SQL) to extract suitable data fields from the database (Table 4.1). The origin and destination properties are identified in the NLIS by a Property Identification Code (PIC). This code is typically structured as an eight character code, where the first four characters represents the state or territory and
Table 4.1: Data fields obtained from the NLIS database for animal movements in Western Australia.

<table>
<thead>
<tr>
<th>Data field</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>Tag ID</td>
<td>To protect privacy this number was actually the NLIS database unique record identifier, which is not available through any public interface, rather than the animals tag number or RFID chip number.</td>
</tr>
<tr>
<td>Source</td>
<td>The shire (or state) where the move originated.</td>
</tr>
<tr>
<td>Source type</td>
<td>The establishment type of the source, such as producer, abattoir, export depot etc.</td>
</tr>
<tr>
<td>Destination</td>
<td>The shire (or state) where the move ended.</td>
</tr>
<tr>
<td>Destination type</td>
<td>As for source type</td>
</tr>
<tr>
<td>Date of transfer</td>
<td>The date that the transaction was recorded in the database—not necessarily the date that the animal moved, but due to legislative requirements likely to be with a week of the movement.</td>
</tr>
<tr>
<td>Deceased</td>
<td>Recorded whether an animal’s death had been recorded in the system, either by the producer or subsequent to an abattoir visit.</td>
</tr>
</tbody>
</table>

the next level of administrative region, and the last four numbers represent the property number within this administrative region. To maintain privacy, the final four digits of the PIC were dropped during the extraction process, and for this exercise each animal was identified using a unique number (with no relation to the original code) rather than one associated with either the device or the property. Some establishments (abattoirs, lairages, saleyards, showgrounds, feedlots, export depots and agents) have a PIC in which the first four characters relate to the state and type of establishment not the next level of administrative region. In this case the type of establishment was recorded and the location manually matched to the appropriate shire. It may be possible to manually identify these establishments (few shires have more than one abattoir or showground). Movements to interstate locations were coded as a state of destination code, not to the level of the shire.

The data were then imported into a PostgreSQL 8.4 relational database (PostgreSQL Global Development Group, 2010) with the PostGIS 1.5 spatial extension (Ramsey, 2009). A database table linking shires with spatial units (further described in Section 4.2.2) was developed.
4.2. MATERIALS AND METHODS

A class of PIC called a Buyer PIC is for livestock buyers (for example, livestock agents or live export buyers) to use until destination is determined when buying from sales. These were accounted for by assigning the PIC to the buyer’s address destination, which may be a metropolitan address, not a property. This may cause some confusion as the animals were not likely to actually be moved to this location, but this only applied to a small percentage (0.093%) of the total set of those records retained for analysis.

Greater Mekong Subregion

Data on livestock movements were collected from a number of sources, including official movement records, import and export permits, checkpoint records, abattoir slaughter records, expert opinion and by discussion with traders.

A number of different methods were trialled to collect this data, including the use of local computer databases, entry into a central database and submission of original paper records (or copies) to a central office.

Initially, livestock movement permit data were entered into a customised local computer database by provincial or central office staff after the completion of a training course on the use of the database and the purpose of the activity. This database was developed using Microsoft Access 2003\(^2\) and was distributed to all staff at the training course. The database was translated into Khmer using old style (non-Unicode) Khmer font sets. Where possible the database provided pre-filled lists and tick boxes to try to reduce issues associated with free entry fields, however some fields (numbers of stock and trader name and contact details) were required to be entered as text or integers. A drop down menu offered the option to re-use traders whose details were already stored in the system.

With a view to a potentially region-wide benefit of understanding livestock movements, a web-based interface for animal movement recording was also created\(^3\) to

\(^2\)http://office.microsoft.com/en-au/access/

\(^3\)http://trade.animalhealthresearch.asia
allow the direct entry of animal movement records into an online database. This web application provided a movement entry interface based on the movement permits of both Cambodia and Laos, and included drop down lists of traders, provinces, districts, species and reasons for movement, as well as pop up calendars. The same information was collected as with the Access database, but this approach had a number of important benefits. The process of merging the individual local databases was very time consuming, and almost every individual database file came with its own unique set of computer viruses. The online approach also allowed easy bug-fixing and the ability to make modifications rapidly, such as improved data-validity checking, ensured that all users were using the appropriate version. As well, it was fully Unicode compliant, allowing the development of a simple translation system and the opportunity for users to record data such as trader names in the appropriate native script, thus avoiding transliteration errors (and subsequent ‘duplication’ of trader names). Another advantage of the centralised system was the immediate availability of the data, providing the ability to automatically generate reports and maps such as that shown in Figure 4.1.

Unfortunately, although the internet is available in much of Cambodia and Laos, it is much less available in government offices in regional areas. There was some speculation during the early phases of the research that this would change with the roll-out of the FAO TADInfo system, but in neither country did this proceed. This meant the use of the web based system was limited, although likely to improve over time. In the interim, in Laos the web system was installed onto memory sticks, and staff were trained in using the web based system from a local computer. An option in the web system interface when used from memory sticks allowed the staff to export the raw tables into a file in comma separated values (.csv) format that was relatively safe from computer viruses, which could then be submitted to the national office.

Finally, where the above mentioned approaches were ineffective, it was agreed that provinces would just send in their paper copies to the national office in each country
Figure 4.1: A screenshot showing the capabilities of the online reporting interface to show real-time maps of animal movements based on permit data entered from provincial movement records. In this example FMD outbreak locations during the corresponding period are also shown. Clicking on the arrows opens a further window with a listing of the movements that occurred along that route including date, species and number.

where the staff could enter them. Once these reports were received, they were entered into the online database. Offices that continued to fail to submit reports were individually visited and the available records copied and subsequently entered.

The data fields collected are shown in Table 4.2. At the time of data analysis, the reporting system also included reports of other species and commodities (owing to a generic design to allow for potential future expansion), including poultry and meat. As this research is focusing on FMD in livestock, for future analysis only the records pertaining to cattle, pigs, water buffalo and goats were included. The type of paper form from which the information was being transferred was not recorded in the original Access database. It was subsequently recognised that some provinces were using other sources of movement data (such as sanitary certificates), and historical data that was entered at a national level included purchasing permits (which were issued by the central office to authorise export to other countries). Although there was a potential risk of duplication (national issue of purchasing permits still required exporters to obtain provincial movement permits) it is unlikely as the earliest pro-
Table 4.2: Data fields obtained from the livestock movement permits database for animal movements in the Greater Mekong Subregion.

<table>
<thead>
<tr>
<th>Data field</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>Permit ID</td>
<td>The number of the movement permit, which allowed validation of the data entry.</td>
</tr>
<tr>
<td>Issuing office</td>
<td>The office where the permit was issued.</td>
</tr>
<tr>
<td>Trader</td>
<td>The name of the trader applying for the permit.</td>
</tr>
<tr>
<td>Source country</td>
<td>The country in which the movement originated.</td>
</tr>
<tr>
<td>Source province</td>
<td>The province in which the movement originated.</td>
</tr>
<tr>
<td>Source district</td>
<td>The district in which the movement originated.</td>
</tr>
<tr>
<td>Destination country</td>
<td>As for source country</td>
</tr>
<tr>
<td>Destination province</td>
<td>As for source province</td>
</tr>
<tr>
<td>Destination district</td>
<td>As for source district</td>
</tr>
<tr>
<td>Date of movement</td>
<td>The date that the movement of the animals was planned - not necessarily the date that the animal moved, but the best approximation available.</td>
</tr>
</tbody>
</table>

Vincial movement records captured only date back to 2007, while the purchasing permits mainly included movements prior to 2007. For final analysis, spatial resolution was restricted to province level, analogous to the shire level in WA. Details of individuals were de-identified. These data were encoded to match spatial identifiers and allowed geographic representation and comparison of movements from different sources.

4.2.2 Administrative Data

In WA and the GMS the spatial resolution of movement data were managed at the shire or province level respectively. In order to analyse the movements and display them on maps, it was necessary to have a suitable set of spatial data to be able to identify all necessary locations. This allowed the calculation of centroids which could be used to draw line maps, locate the nearest roads and calculate distances, as well as automatically generate choropleth map outputs.
Western Australia

Western Australia covers an area of 2,532,400 km$^2$ and had 1,997,000 cattle in the 2008/09 year (Australian Bureau of Statistics, 2010). The state is divided into 120 local government areas (LGA) (known as shires or towns, but all referred to in this paper as shires). The human and livestock populations are concentrated in the south west of the state. Outside the Perth metropolitan area cattle production is carried out in all but one of the shires.

The spatial units identified by NLIS PIC codes are the responsibility of the relevant state or territory to identify and administer. In WA they are closely related to current shires in most parts of the state (some boundary changes and amalgamation/separation of these shires have occurred since the property codes were originally allocated, but the properties retain their original code). Because of these changes current administrative boundary data for WA were sourced from the Australian Bureau of Statistics (ABS) (Australian Bureau of Statistics, 2009), and some modifications of the geometry were required to ensure shire PIC codes were matched with an existing areal unit.

Shires were matched to the NLIS data codes for 120 shires, towns and cities outside the Perth metropolitan area. The shire of Ngaanyatjarraku was created in 1993 by dividing the Wiluna Shire. There are no livestock holdings in the Ngaanyatjarraku shire thus all NLIS records referring to Wiluna were left referring to only Wiluna. A new polygon was added to represent the location of the City of Kalgoorlie, which was originally managed as a separate administrative area. The current shire of Broomehill-Tambellup was split East-West to create Broomehill in the north, and Tambellup in the south to approximate original shire codes in the NLIS database.

To account for seasonal management differences the state was split into three regions, summer rainfall, southern rangelands and south west agricultural (Figure 4.2). The boundaries approximate the Northern Rangelands, Southern Rangelands and Agri-
cultural regions used by DAFWA\textsuperscript{4}; however they were constructed using ABS shire boundary data. This allowed analysis of movements by region as well as time.

![Figure 4.2: Western Australia showing management regions used in this study.](image)

**Greater Mekong Subregion**

No consistent, accurate and up-to-date data sets are publicly available that cover the entire GMS\textsuperscript{5}. Most countries have well developed GIS data sets incorporating boundaries of administrative units, as well as up-to-date road information, natural features (including watercourses and population areas) and land use. This data may be available under licence\textsuperscript{6} for use by non-government users. Cameron (1997) noted that Cambodia and Laos have had well developed maps at 1:50 000 resolution in 1997. Other countries have limited up-to-date data available for non-government use (if at all).

Vietnam for example, has relatively well developed spatial data, but charges full cost recovery to other government departments for access to these data, and only makes

\textsuperscript{4}The regions are shown on the DAFWA website at \url{http://www.agric.wa.gov.au/}.

\textsuperscript{5}Or at least, this was the situation when this research started. In fact, efforts such as Global Administrative Areas \url{http://www.gadm.org/} and OpenStreetMap \url{http://www.openstreetmap.org} are improving this situation. There is also the United Nations Second Administrative Level Boundaries project \url{http://www.unsalb.org/} which while promising early has had little progress in the target countries for this research.

\textsuperscript{6}For example, in Australia the ABS provides data under a Creative Commons Attribution licence which is available at \url{http://creativecommons.org/licenses/by/2.5/au/}.
4.2. MATERIALS AND METHODS

a limited subset available to foreigners. As an example a quote of \( \approx \) USD 20 000 was offered for a set of village point location data (Dung Do, personal communication, 2010). Compounding this problem, is a lack of consistency between geographic projection systems (and their application), confusion about datums, nomenclature, standards for naming, and in some cases the adoption (or lack thereof) of the Unicode standard for text encoding, to allow viewing of names in local languages without the need for a specific font set. Over time a number of projects have attempted to rectify this situation, and provide unified consistent data into the public domain, for example Digital Chart of the World.

A data set of provinces was created by using the freely available data set of Digital Chart of the World\(^7\). All data were stored in a relational database PostgreSQL\(^8\) with the spatial extension PostGIS\(^9\). The province data were then updated by comparing the province boundaries with more recent data sets such as the current Cambodian administrative data (supplied by the Cambodian Department of Animal Health and Production), Laos administrative data (supplied by the Laos Department of Livestock and Fisheries) and various copyright restriction free data sets supplied by Non-Government Organisations such as Global Administrative Areas\(^10\) and OpenStreetMap\(^11\). Importantly, this process was required to take into account areal modifications such as the discontinuation of the Xaisomboun Special Administrative Region in Laos. Map editing was done using Quantum GIS\(^12\) and topological correction managed using GRASS\(^13\). Most of these data had poor or limited correction for topology, and some were very high resolution depending on the original data source. This resolution was unsuitable for regional-level analysis and mapping, increasing processing and rendering times, and resulting in large file sizes. To overcome this, the resolution of the underlying data was reduced using the Douglas-Peuker

\(^7\)http://www.maproom.psu.edu/dcw/, no longer available.  
\(^8\)http://www.postgres.org/  
\(^9\)http://postgis.refractions.net/  
\(^12\)http://www.gqis.org/  
\(^13\)http://grass.osgeo.org/
algorithm\textsuperscript{14} implementation in GRASS which preserved the boundary topology. A five-level hierarchical coding system was implemented (country, province, district, commune and village) and stored in the database table alongside the geographic unit name (in English as well as the local language). The hierarchical code was used as the common identifier between data and geographic location. Data matching was done using the application developed for FMD outbreak investigations using Metaphone phonetic matches (previously discussed on page 42).

### 4.2.3 Road Network Data

The distance between two points is often measured as a straight-line between them (the euclidean distance). However, long distance transport of animals is much more likely to follow transport infrastructure (roads or railways) and this is often influenced by natural features and rarely is a straight line. In order to obtain more realistic estimates of the distances animal were travelling, data on road networks were sought for WA and the GMS.

**Western Australia**

A set of detailed WA road data was provided by the Department of Main Roads WA in shapefile format\textsuperscript{15}. The road data were imported into the same spatial database as the NLIS data. To reflect the likely road conditions that might influence the decision of livestock transporters to follow certain routes the road segments were weighted according to the road network type and carriageway construction (dual or single) specified in the data (Table 4.3). These weights were then used to create a graph of probable routes between any two shires, originating or terminating at either an important livestock related feature (saleyard, dipyard) or the centroid of the shire.

The calculation of these routes (and distances) between all 120 shires requires sub-

\textsuperscript{14}http://grass.osgeo.org/grass64/manuals/html64_user/v.generalize.html

\textsuperscript{15}Shapefile format is a de-facto standard for exchange of spatial data.
4.2. MATERIALS AND METHODS

Table 4.3: The weighting applied to different road segments of the Western Australia road network to enhance the probability that the routes chosen by the routing algorithms would match those chosen by livestock transporters.

<table>
<thead>
<tr>
<th>Road Types</th>
<th>Relative Weight</th>
<th>Rationale</th>
</tr>
</thead>
<tbody>
<tr>
<td>Carriageways</td>
<td></td>
<td></td>
</tr>
<tr>
<td>S - Single</td>
<td>1.2</td>
<td>Increasing the effective length of a single carriageway makes it less favourable than a dual carriageway road.</td>
</tr>
<tr>
<td>Road Construction</td>
<td></td>
<td></td>
</tr>
<tr>
<td>M - Main Road (Highway)</td>
<td>1/110</td>
<td>Open road speed for livestock trucks is limited to 100km/h</td>
</tr>
<tr>
<td>L - Local Road</td>
<td>1/60</td>
<td>Local (Urban) Road Speed</td>
</tr>
<tr>
<td>Z - Special Use Road</td>
<td>1/50</td>
<td>Roads with limited access</td>
</tr>
<tr>
<td>PR - Private Road</td>
<td>1/40</td>
<td>Private road average speed around 40km/h</td>
</tr>
</tbody>
</table>

Substantial computing power. To overcome this requirement and speed up analysis all the shire-to-shire combinations (there were 14,280) were pre-calculated. A set of algorithms was developed to identify the most likely location for animal movements in a shire. A set of saleyard locations was obtained from the National Livestock Reporting Service and livestock tick dip yards from DAFWA. In the shires where these facilities existed, the algorithm used their location to identify the nearest road segment in each shire. Where no specific facility existed, the geometric centroid was used where it was contained within the shire, or a point on surface (Open Geospatial Consortium, 2005). Then a shortest route from the origin shire to the destination was calculated in the PostgreSQL database using the Dijkstra algorithm from the pgRouting libraries. These calculations were coded as a set of database procedural language functions and are included in Appendix D.2.

Greater Mekong Subregion

For the GMS, there was very little available in the way of suitable road network data. Considerable effort was spent in researching available data sets and attempting to make corrections for major roads through the region. The best available data set

\[http://www.pgrouting.org\]
was that produced by the OpenStreetMap Project\textsuperscript{17}, however it is most definitely a work in progress, and is inconsistent in its level of development and quality across the region. Perhaps the most noteworthy aspect of the OpenStreetMap data is the rapid progress being made in mapping roads in the GMS. The number of roads in the region has almost tripled in 2011 compared to when first evaluated in 2008. Unfortunately, even with this advance many of the roads were unsuitable for use by routing algorithms, and would require weeks or even months of effort to make suitable for proper road-routing\textsuperscript{18}. This meant it was not possible to evaluate road distances between locations as a part of this component of the research. Further progress was made for main roads in Cambodia in order to improve model fit for the role of price as a predictor and is discussed on page 171.

4.2.4 Statistical Analysis

Analyses of the number of movements by animal (in WA) or animal group (in the GMS), origin, destination, type and region were undertaken. For illustrative purposes the size of the WA cattle herd for each year in the analysis was obtained from DAFWA (Western Australia Department of Agriculture and Food, 2009). Analysis of the temporal trends was conducted using seasonal decomposition by locally weighted smoothing (Cleveland et al., 1990). In WA, the distribution of number of movements by season (for all years) between regions and by region between the total for seasons (of all years) was compared. This was done using the comparison of proportions method as describe by Fleiss (1981). The contribution of season, trend and the residuals was assessed by examining the interquartile range of the smoothed time series. Analyses were conducted using the R Statistical Environment (R Development Core Team, 2011).

\textsuperscript{17}http://www.openstreetmap.org/  
\textsuperscript{18}Routing algorithms typically work by stepping along line segments, and when they reach the end of a line, use the heuristics of the current algorithm to choose the next most appropriate line from those which intersect the node at the end of the current segment. A common problem with many road maps involves lines not joining correctly at nodes, or small breaks in the lines.
4.2.5 Network Analysis

Network analysis terms referred to in this thesis are described in Table 4.4.

A number of the measures used in network analysis are particularly useful for describing the risk of infectious disease spread. In particular, the centrality measures can be used to identify nodes that are of high risk of becoming infected due to the large number of other nodes connecting to them (measured by in-degree), the nodes from which disease may disperse most widely (out-degree) or the nodes that are most likely to become infected regardless of where disease starts or ends (those with high betweenness).

The process of infection is (in network terms) a directed movement—disease can only move from infectious populations to susceptible ones. The Giant Strong Component (GSC) is the largest group of nodes which are connected by arcs (directed edges) in a network and in which any node can be reached from any other node. Infection that starts at any node in the GSC can spread to every other node in the same component, thus providing a lower bound of the maximum size of the outbreak (Danon et al., 2011; Dubé et al., 2009; Robinson and Christley, 2007).

In WA, two networks were considered—a network by category of node (producer, feedlot, abattoir etc) and one by location of the node. As the data available in the GMS were more limited, only the network of location was investigated.

Key elements of each network were described and centrality measures (in-degree, out-degree and betweenness) were calculated in R using the packages sna (Butts, 2008b, 2010), network (Butts, 2008a; Butts et al., 2010) and igraph (Csardi and Nepusz, 2006; Csardi, 2008). Choropleth maps were generated using the sp package (Bivand et al., 2008; Pebesma and Bivand, 2005). These measures were reported as graphs for the node type, and as choropleth maps for the node location with shading used to represent the centrality measures expressed as categorical variables.

To investigate the possible number of shires affected in the event of an outbreak the size of the GSC was evaluated. By specifying the number of days from introduc-
tion until detection it is possible to determine how many shires would potentially be involved in an outbreak by the time of detection and the anticipated cessation of animal movements. For this investigation 3, 7, 14 and 30 day periods were investigated. Although the primary interest is the number of days until detection, the impact of season was also investigated. There was an expectation that seasonal effect may vary from year to year, so data from multiple years was used.

The full set of WA NLIS movement records from July 2005–June 2008 was sampled using a bootstrapping approach. One thousand samples were taken for each season (Autumn, Winter, Spring and Summer) for each year (2005–2008). Each sample started on a random day in the season. This process was repeated for each of the four periods of day until detection (3, 7, 14 and 30). Each sample was used to create a network of movements in that period and key network measures including the GSC was calculated. The range of results of the GSC for each period for each season were displayed as boxplots.
### 4.2. MATERIALS AND METHODS

Table 4.4: A précis of network analysis terms used in this chapter. Definitions sourced from Christley et al. (2010); Danon et al. (2011); Dubé et al. (2009); Martínez-López et al. (2009b); Wasserman and Faust (1994). Figure 4.3 is referred to in the table.

<table>
<thead>
<tr>
<th>Term</th>
<th>Definition</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Network Components</strong></td>
<td></td>
</tr>
<tr>
<td><strong>Graph</strong></td>
<td>A collection of nodes and edges.</td>
</tr>
<tr>
<td><strong>Arc</strong></td>
<td>An edge where the direction of flow is important.</td>
</tr>
<tr>
<td><strong>Edge</strong></td>
<td>The link between any two nodes. Some authors differentiate between an edge (undirected) and an arc (see above). The figure would have 12 edges if undirected, but has 14 arcs.</td>
</tr>
<tr>
<td><strong>Geodesic</strong></td>
<td>The shortest path length between two nodes. There are two possible paths from A to H, but it has a geodesic of one (A → H).</td>
</tr>
<tr>
<td><strong>Node</strong></td>
<td>The unit of interest (sometimes called vertices or actors). May be individuals, groups or locations. In the figure there are 10 nodes labelled A–J.</td>
</tr>
<tr>
<td><strong>Path length</strong></td>
<td>The number of edges that join two nodes. For example, from A → B the path length is 1, but from B → A the path length is 4 (B → C → H → F → A).</td>
</tr>
<tr>
<td><strong>Describing Networks</strong></td>
<td></td>
</tr>
<tr>
<td><strong>Centrality</strong></td>
<td>The role of a node in terms of its relationship with other nodes. See betweenness, in-degree and out-degree.</td>
</tr>
<tr>
<td><strong>Density</strong></td>
<td>The proportion of edges in the network from the total number of possible edges (for example if each nodes was joined to each other nodes). In this example there are 14 edges from a total possible 180, giving a density of 14/180 ≈ 0.07.</td>
</tr>
</tbody>
</table>

*continued on next page*
### Term Definition

<table>
<thead>
<tr>
<th>Term</th>
<th>Definition</th>
</tr>
</thead>
<tbody>
<tr>
<td>Diameter</td>
<td>The longest geodesic in a network. In Figure 4.3 this is from J → D.</td>
</tr>
<tr>
<td>Directed</td>
<td>A graph in which the direction of flow along an edge (or more correctly arc) is important. eg Infection can only move from an infectious animal to a susceptible one.</td>
</tr>
<tr>
<td>Undirected</td>
<td>A graph in which the direction of flow along an edge is not important to the analysis. eg Businesses which have traded animals or products between them are linked, but if the interest is whether or not they have conducted business together, then this would be an undirected network.</td>
</tr>
</tbody>
</table>

### Network Measures of Centrality

<table>
<thead>
<tr>
<th>Term</th>
<th>Definition</th>
</tr>
</thead>
<tbody>
<tr>
<td>Betweenness</td>
<td>The frequency with which a node falls between other pairs of connected nodes.</td>
</tr>
<tr>
<td>Components</td>
<td>Sub-graphs in which all the nodes are connected directly or indirectly to every other node.</td>
</tr>
<tr>
<td>Degree</td>
<td>The number of edges adjacent to a node.</td>
</tr>
<tr>
<td>Giant Strong Component</td>
<td>The largest strong component. This provides a measure of the largest number of nodes that may become infected after the introduction of a disease.</td>
</tr>
<tr>
<td>In-degree</td>
<td>The number of arcs terminating on a node.</td>
</tr>
<tr>
<td>Out-degree</td>
<td>The number of arcs originating from a node.</td>
</tr>
<tr>
<td>Strong Component</td>
<td>A component of a directed graph where all nodes are reachable from any other node. The strong components are shown in Figure 4.3 as the nodes grouped in the dashed circles.</td>
</tr>
</tbody>
</table>

*continued on next page*
4.2. MATERIALS AND METHODS

<table>
<thead>
<tr>
<th>Term</th>
<th>Definition</th>
</tr>
</thead>
<tbody>
<tr>
<td>Weak Component</td>
<td>A component of a graph where all nodes are reachable from any other node, as long as the direction of the edge is unimportant. Note the distinction between Strong and Weak Components — a strong component takes into account the direction of the edge (or arc).</td>
</tr>
</tbody>
</table>

![Network Graph](image)

Figure 4.3: An example of a network graph, as referred to in Table 4.4. The highest out-degree is A (3 edges). The nodes with the highest in-degree are D and F (both 3). The greatest betweenness is node F (17). The two strong components are circled by the dashed lines.

To reduce the time taken to perform this analysis sampling was conducted on an eight core dedicated server running FreeBSD 8, PostgreSQL 9 and running R 2.11 (more details are in Section 4.2.6). This sampling process took 2.6 hours on this server.

The significance of the differences between days to detection, season and year were assessed using a multivariable linear regression using the complete set of bootstrap samples. The GSC (the number of shires potentially infected) was modelled by the
days to detection period, the season and the year, all fitted as categorical variables. Inspection of the plot of GSC vs days to detection by season (Figure 4.12) shows no evidence of interaction as the pattern is consistently repeated in each season so no interaction terms were fitted. Although the effect of year of movement was not significant it was retained in the model as it was a priori thought to be a confounder, and it’s retention improved model fit. The final model is shown in Equation (4.1). As each predictor had four levels, follow-up comparisons of group means were done for each level of days to detection, season and year using t-tests and the standard errors derived from the multivariable model. Normality of residuals and homoscedasticity were checked by inspection of plots of residual and fitted values, histograms of residuals and Q-Q plots.

\[
y_i = \beta_0 + \beta_1 \text{season}_i + \beta_2 \text{days}_i + \beta_3 \text{year}_i + \epsilon_i
\]

\[
\epsilon_i \sim N(0, \sigma_i^2)
\]

\[
i = 1, \ldots, n \text{ samples}
\]

where:

- \( y_i \) is the size of the giant strong component,

- \( \text{season}_i \) is the season in which the samples were taken (autumn, winter, spring and summer),

- \( \text{days}_i \) is the number of days as a categorical variable from introduction until detection of disease (3, 7, 14 and 30),

- \( \text{year}_i \) is the year from which the sample was taken (2005–2008).
4.2. MATERIALS AND METHODS

Web-based query interface

A web application which allowed user specified date ranges and analyses to be conducted was developed\(^{19}\). This provided a range of network summary metrics and a list of centrality measures for the selected parameters as shown in Figure 4.4, and a clickable link to a graphical representation of the network shaded to match the centrality measure chosen for sorting (Figure 4.5). Alternatively if the network was chosen by shire, a map showing the movements between shires could also be produced by ordering the list by the shire, rather than the centrality measure.

---

\(^{19}\)http://nlis.animalhealthresearch.asia/network.php?page=centrality
Figure 4.5: A graphical output of the web based National Livestock Identification System query interface showing the network by node category. The graph shown is coloured according to the out-degree of the node.
4.2. MATERIALS AND METHODS

4.2.6 A Quick Technology Summary

The web application for the NLIS data analysis was developed using a mix of open source technologies. Using the Apache web server\textsuperscript{20} on a FreeBSD\textsuperscript{21} server, a simple php\textsuperscript{22} script created the web page form that allowed the user to select parameters and display the results as shown in Figures 4.4 and 4.5. The NLIS data were stored in a PostgreSQL\textsuperscript{23} relational database and so to reduce processing overhead the queries were handled within the database by developing functions using the PL/R procedural language\textsuperscript{24}. This allows R\textsuperscript{25} scripts to be invoked from within the database, reducing the amount of time required to send data backwards and forwards between applications. The analysis in R used the previously mentioned network, sna and igraph packages, as well as the RColorBrewer package\textsuperscript{26} to select colours that were colour-blind and print friendly, and Cairo for device independent graphical output (in this case png format for web page display. For the spatial data, the rgdal package\textsuperscript{27} was used to import a background map from the PostgreSQL database, and geographically valid centroids (sometimes known as \textit{point in polygon}) for each shire were identified using the PostGIS spatial extension.

\begin{footnotesize}
\begin{itemize}
\item \textsuperscript{20}\url{http://httpd.apache.org/}
\item \textsuperscript{21}\url{http://www.freebsd.org/}
\item \textsuperscript{22}\url{http://www.php.net/}
\item \textsuperscript{23}\url{http://www.postgresql.org/}
\item \textsuperscript{24}\url{http://www.joeconway.com/plr/}
\item \textsuperscript{25}\url{http://www.r-project.org/}
\item \textsuperscript{26}Based on Cynthia Brewer’s ‘ColorBrewer’ system, \url{http://www.colorbrewer2.org/}
\item \textsuperscript{27}A R package which makes the benchmark OGR (\url{http://www.gdal.org/ogr/}) tools for spatial data exchange available.
\end{itemize}
\end{footnotesize}
CHAPTER 4. UNDERSTANDING LIVESTOCK MOVEMENT PATTERNS

4.3 Results

4.3.1 Western Australia

The data for the period July 2005 to June 2008 included 1,737,920 animals completing 3,962,188 movements. This included twelve movements of a total of 107 cattle from eight property source codes and 19 movements of 482 cattle to ten property destination codes which could not be identified due to anomalies in the data. Six of these movements involved the same property. These records were excluded from further analysis as no further information was available to identify to which shire they related. A further two movements to a destination type of ‘exported’ were adjusted to a destination type of ‘port’—in the case where the origin was a Queensland saleyard they were assigned to export through the closest WA port to Queensland—Wyndham. Where the origin was the Swan Valley (the shire where the Midland Saleyards are located) they were assigned to the Fremantle Port—again the nearest live export port. This resulted in a final data set of 3,961,600 movements undertaken by 1,737,833 animals.

The number of movements undertaken by this group of animals ranged from one to 19 during the three year period. 589 animals (0.03%) moved ten or more times. The three animals that moved 19 times travelled together and covered approximately 7,500 km each between January 2007 and May 2008 (Figure 4.6).

Over the period the number of movements recorded on the database have increased, as shown in Table 4.5 (and Figure 4.7(c)). The numbers going to saleyards have not changed substantially during this period, while numbers going to abattoirs and live export have increased.

Over the three years the number of animals moved to the abattoir as their only reported movement with an RFID increased. From the dataset it is not possible to determine whether the animal was moving from its property of birth or had been born elsewhere but moved onto the property of origin for that movement prior to
Table 4.5: Number of movements of cattle (1 000’s) recorded in the NLIS database for WA from July 2005 to June 2008, grouped by a) source type and b) destination type.

(a) Movements by source type

<table>
<thead>
<tr>
<th>Source Type</th>
<th>FY 06</th>
<th>FY 07</th>
<th>FY 08</th>
<th>Total</th>
</tr>
</thead>
<tbody>
<tr>
<td>Abattoir</td>
<td>19</td>
<td>2</td>
<td>1</td>
<td>22</td>
</tr>
<tr>
<td>Agent</td>
<td>0</td>
<td>0</td>
<td>2</td>
<td>2</td>
</tr>
<tr>
<td>Export Buyer</td>
<td>0</td>
<td>0</td>
<td>408</td>
<td>408</td>
</tr>
<tr>
<td>Export Depot</td>
<td>79,621</td>
<td>170,691</td>
<td>190,399</td>
<td>440,711</td>
</tr>
<tr>
<td>Feedlot</td>
<td>5,161</td>
<td>19,295</td>
<td>75,746</td>
<td>100,202</td>
</tr>
<tr>
<td>Interstate PIC</td>
<td>9,442</td>
<td>16,103</td>
<td>37,326</td>
<td>62,871</td>
</tr>
<tr>
<td>Port</td>
<td>241</td>
<td>1,613</td>
<td>2,920</td>
<td>4,774</td>
</tr>
<tr>
<td>Producer</td>
<td>587,488</td>
<td>763,909</td>
<td>1,037,973</td>
<td>2,389,370</td>
</tr>
<tr>
<td>Saleyard</td>
<td>324,414</td>
<td>300,549</td>
<td>335,760</td>
<td>960,723</td>
</tr>
<tr>
<td>Showground</td>
<td>790</td>
<td>697</td>
<td>1,030</td>
<td>2,517</td>
</tr>
<tr>
<td><strong>Total</strong></td>
<td>1,007,176</td>
<td>1,272,859</td>
<td>1,681,565</td>
<td>3,961,600</td>
</tr>
</tbody>
</table>

(b) Movements by destination type

<table>
<thead>
<tr>
<th>Destination Type</th>
<th>FY 06</th>
<th>FY 07</th>
<th>FY 08</th>
<th>Total</th>
</tr>
</thead>
<tbody>
<tr>
<td>Abattoir</td>
<td>29,068</td>
<td>58,434</td>
<td>330,481</td>
<td>417,983</td>
</tr>
<tr>
<td>Agent</td>
<td>152</td>
<td>27</td>
<td>25</td>
<td>204</td>
</tr>
<tr>
<td>Export Buyer</td>
<td>109</td>
<td>3,223</td>
<td>165</td>
<td>3,497</td>
</tr>
<tr>
<td>Export Depot</td>
<td>95,754</td>
<td>167,236</td>
<td>203,811</td>
<td>466,801</td>
</tr>
<tr>
<td>Feedlot</td>
<td>13,999</td>
<td>57,686</td>
<td>80,590</td>
<td>152,275</td>
</tr>
<tr>
<td>Interstate PIC</td>
<td>28,353</td>
<td>13,990</td>
<td>23,948</td>
<td>66,291</td>
</tr>
<tr>
<td>Port</td>
<td>83,617</td>
<td>189,032</td>
<td>195,907</td>
<td>468,556</td>
</tr>
<tr>
<td>Producer</td>
<td>430,497</td>
<td>490,657</td>
<td>512,073</td>
<td>1,433,227</td>
</tr>
<tr>
<td>Saleyard</td>
<td>323,691</td>
<td>290,239</td>
<td>332,197</td>
<td>946,127</td>
</tr>
<tr>
<td>Showground</td>
<td>805</td>
<td>697</td>
<td>1,027</td>
<td>2,529</td>
</tr>
<tr>
<td>Deceased</td>
<td>1,131</td>
<td>1,638</td>
<td>1,341</td>
<td>4,110</td>
</tr>
<tr>
<td><strong>Total</strong></td>
<td>1,007,176</td>
<td>1,272,859</td>
<td>1,681,565</td>
<td>3,961,600</td>
</tr>
</tbody>
</table>

Table 4.6: Number of cattle in WA from July 2005 to June 2008 (source: Western Australia Department of Agriculture and Food, 2009, p34).

<table>
<thead>
<tr>
<th>Cattle Type</th>
<th>FY 06</th>
<th>FY 07</th>
<th>FY 08</th>
<th>Average</th>
</tr>
</thead>
<tbody>
<tr>
<td>Beef</td>
<td>2,127,000</td>
<td>2,275,000</td>
<td>2,223,200</td>
<td>2,208,400</td>
</tr>
<tr>
<td>Dairy</td>
<td>116,000</td>
<td>116,000</td>
<td>104,400</td>
<td>112,133</td>
</tr>
<tr>
<td><strong>Total</strong></td>
<td>2,243,000</td>
<td>2,391,000</td>
<td>2,327,600</td>
<td>2,320,500</td>
</tr>
</tbody>
</table>
Notwithstanding the trend over the three years, there is a significant seasonal component to the variation in the number of movements during the year (Figure 4.7(b)). The movements were plotted by season and region (Figure 4.8). This shows a significant ($p < 0.05$) variation in animal movements between all regions and seasons.

There were 11 types of node in the directed network which are listed in Table 4.7. The largest number of movements was from producer to saleyard, producer to producer and saleyard to producer (Table 4.7). There were also large numbers of animals moving directly from producer to abattoir and export depots.

Over the study period the node type network had a density (the proportion of all possible edges that actually exist) of 0.48 and a diameter (the longest geodesic or
Table 4.7: Matrix of total movements by origin and destination node type. Loops (movements between nodes of the same type) are shown in bold.

<table>
<thead>
<tr>
<th>Origin</th>
<th>Abattoir</th>
<th>Agent</th>
<th>Export Buyer</th>
<th>Export Depot</th>
<th>Feedlot</th>
<th>Interstate PIC</th>
<th>Port</th>
<th>Producer</th>
<th>Saleyard</th>
<th>Showground</th>
<th>Deceased</th>
</tr>
</thead>
<tbody>
<tr>
<td>Abattoir</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Agent</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Export Buyer</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Export Depot</td>
<td>3 844</td>
<td>169</td>
<td>22 554</td>
<td>3 006</td>
<td>1 581</td>
<td>335 580</td>
<td>47 007</td>
<td>26 920</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Feedlot</td>
<td>51 332</td>
<td></td>
<td>3 567</td>
<td>1 152</td>
<td>558</td>
<td>29 082</td>
<td>10 510</td>
<td>3 272</td>
<td>35</td>
<td>694</td>
<td></td>
</tr>
<tr>
<td>Interstate PIC</td>
<td>7 166</td>
<td></td>
<td>163</td>
<td>1</td>
<td>54 075</td>
<td>1 467</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Port</td>
<td>2 446</td>
<td></td>
<td>284</td>
<td>1</td>
<td>2 041</td>
<td>1</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Producer</td>
<td>243 124</td>
<td>5</td>
<td>1</td>
<td>315 335</td>
<td>115 094</td>
<td>46 637</td>
<td>79 090</td>
<td>673 961</td>
<td>910 867</td>
<td>2 494</td>
<td>2 762</td>
</tr>
<tr>
<td>Saleyard</td>
<td>119 099</td>
<td>199</td>
<td>3 327</td>
<td>115 288</td>
<td>32 737</td>
<td>17 500</td>
<td>24 640</td>
<td>643 731</td>
<td>3 599</td>
<td>603</td>
<td></td>
</tr>
<tr>
<td>Showground</td>
<td>584</td>
<td></td>
<td>37</td>
<td>13</td>
<td>15</td>
<td>1 881</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>
minimum path between any two points) of three. The highest betweenness was 20.3 (0.226) for the producer node (Figure 4.9a). This node had out-degree of 10 and in-degree of 8 (Figures 4.9b and 4.9c). No connection existed to producers from export buyer or (unsurprisingly) deceased to any other nodes.

When analysed by location, there were 126 nodes (120 shires and 6 other states or territories). Over the three year period, only two shires (Sandstone and Mount Magnet) and one state (Tasmania) received no cattle. The location network had a density of 0.25 and a diameter of three for the full three years. The Shire of Swan had the highest centrality with a betweenness of 2 561 and in-degree of 116.

These measures of the network will vary when the temporal period is adjusted, as some nodes will no longer be linked by movement. On a weekly basis the network diameter ranged from 5 to 11, with a median 6.

To aid identification of high risk areas for disease spread (high in-degree) and source (high out-degree) as well as locations where prevention activities should be focussed

---

28 The number in brackets in the normalised or relative statistic (Martínez-López et al., 2009b) and allows comparison of the centrality measures between graphs of different sizes.

29 Sometimes also known as degree prestige.
4.3. RESULTS

Figure 4.8: Moves of cattle by season and management zone in Western Australia during the period July 2005–June 2008.

(high betweenness) this was displayed as a map (Figures 4.11, 4.10a and 4.10b respectively). It is important to clarify that these metrics represent the interconnectedness of the shires, but the absolute number of animals actually being moved are not considered in their derivation.

The number of WA shires that had the potential to become infected is shown in Figure 4.12. The high level of connectivity between shires is such that within three days of a shire becoming infected, the outbreak could have spread through movement of infected animals to between 20 and 50 other shires. If detection of disease and implementation of a livestock movement standstill occurred within four days of this, 15 more shires (approximately half the state) may be involved in an outbreak. A delay of a further seven days would increase the number to approximately 70 shires affected. By thirty days this number increases to nearly three quarters of the shires in the state having infected animals. The size varied significantly \( p < 0.05 \) between each category of days until detection. The size varied significantly \( p < 0.05 \) between all seasons except between Winter and Spring. The model checking was done by inspecting plots of residual and fitted values, histograms of residuals and Q-Q plots; no evidence of heteroscedasticity was seen.
Figure 4.9: Centrality measures by node type in Western Australia during the period July 2005–June 2008. The high betweenness of producers in (a) supports the investment in targeting producers as a key role in identifying disease. Producers and export depots are at the highest risk of receiving diseased animals, due to their high in-degree (b), and their high out-degree (c) (along with feedlots and saleyards) would suggest that they are highly likely to be involved in the spread of disease.
Figure 4.10: Choropleth maps showing a) in-degree and b) out-degree by shire in Western Australia during the period July 2005–June 2008. Detail of the south-west of WA is shown in the map on the right for each of a and b. The shires of Swan (SN) and Harvey (HY) both feature prominently as having high in-degree and high out-degree.
Figure 4.11: Choropleth map showing betweenness by shire in the south-west of Western Australia during the period July 2005–June 2008. The whole state map is not shown as no shires outside the south-west had a betweenness of greater than 200. Shires with betweenness of 200 or more were all within the south west agricultural region. The very high betweenness in the shires of Harvey (HY) and Swan (SN) are most likely related to the presence of saleyards.
Figure 4.12: Number of Western Australia shires likely to become infected by season. The solid bar is the median value, the box represents the first and third quartiles.
4.3.2 Greater Mekong Subregion

The details of a total of 15 743 official movement permits were collected during the research. Of these, 3 060 were collected in Cambodia using the Access database system by provincial office staff. Although the database was designed to minimise free text entry (by offering previously entered information in drop down lists for subsequent records) a combination of transliteration issues and multiple users resulted in large numbers of duplicates, particularly with trader names, reasons for movement and routes. These data were imported into the PostgreSQL database that had been developed for online entry of data.

The remaining 11 643 records were entered into the database using the web interface by three national level staff each from Cambodia (5 915 records) and Lao PDR (5 728 records).

Data collection started in January 2008 in Cambodia and July 2008 in Laos. Retrospective records were included in Cambodia, going back as far as 2004. Where it was possible to identify the original source of the data, 3 028 records were derived from sanitary certificates, 2 885 from movement permits and 1 098 from purchasing permits.

A number of provinces submitted very few (if any) movement records. Seven provinces in Cambodia and 4 in Laos submitted less than 100. In Cambodia these provinces were visited and if further records were available they were collected and entered into the system. In a number of these provinces it appears that few movement permits are issued as the province is a net importer of livestock.

A number of records had errors in the movement dates. Six of the records had a start date prior to 2003, seven had a date greater than 2010. The use of permit reference numbers helped greatly here; sequential data from provinces where one part of the date was out of range were relatively straightforward to adjust. In a number of records this was adjusted using other supporting information (such as destination date—where the start and end day and month were the same, but the
4.3. RESULTS

years were 2000 and 2009, the start year was adjusted to 2009). Where it was not possible to make a suitable assumption the record was discarded.

The number of recorded movements that occurred for groups of buffalo, cattle and sheep are shown in Table 4.8. Although the majority of moves were between provinces in the same country, there was a small number of movements into both countries from Myanmar and Thailand, and a small number out to China and Vietnam. The recording on movement permits of destinations in Vietnam is interesting as for most of the period that these data were being collected there was no official mechanism for importing cattle into Vietnam. The number of movements shown in Table 4.8 is likely to under-represent the actual number of movements for a number of reasons, including that data on movements was only collected in Cambodia and Laos (and only two movements of pigs were recorded between these two countries in the three year period), the closed border into Vietnam and the recording of movements between countries other than Cambodia and Laos (which may be due to erroneous data entry).

The numbers of animals moved are shown in Table 4.9. The numbers of animals (For example, < 40 000 per year in Cambodia and ≈ 6 500 per year in Laos) covered by permits is likely to be an underrepresentation of the total number of movements. To illustrate this, during the course of this research some summary data from road checkpoints in Laos during 2009 were assessed. Although there appear to be some errors in the data the number of animals passing through the checkpoints is shown in Table 4.10.

The seasonal decomposition of cattle movements for Cambodia and Laos (Figure 4.13) showed an increased activity each year around July, however this seasonal component (Figure 4.13(b)) only accounts for 10.7% of the total movements. The trend accounts for nearly two thirds of the movements, however the increasing trend may be an artefact induced by the limited success in trying to collect retrospective records prior to 2008.

Seventy of 211 provinces in South East Asia (and 5 of 34 in China) were involved in
Table 4.8: The number of movements within and between countries from January 2007 to December 2009 by species.

(a) Buffalo

<table>
<thead>
<tr>
<th>Origin</th>
<th>Cambodia</th>
<th>Laos</th>
<th>China</th>
<th>Vietnam</th>
</tr>
</thead>
<tbody>
<tr>
<td>Cambodia</td>
<td>557</td>
<td>0</td>
<td>0</td>
<td>1</td>
</tr>
<tr>
<td>Laos</td>
<td>0</td>
<td>2,631</td>
<td>20</td>
<td>1</td>
</tr>
<tr>
<td>Myanmar</td>
<td>0</td>
<td>0</td>
<td>5</td>
<td>0</td>
</tr>
<tr>
<td>Thailand</td>
<td>0</td>
<td>29</td>
<td>30</td>
<td>0</td>
</tr>
</tbody>
</table>

(b) Cattle

<table>
<thead>
<tr>
<th>Origin</th>
<th>Cambodia</th>
<th>Laos</th>
<th>China</th>
<th>Vietnam</th>
</tr>
</thead>
<tbody>
<tr>
<td>Cambodia</td>
<td>4,210</td>
<td>0</td>
<td>0</td>
<td>15</td>
</tr>
<tr>
<td>Laos</td>
<td>0</td>
<td>1,785</td>
<td>10</td>
<td>4</td>
</tr>
<tr>
<td>Myanmar</td>
<td>0</td>
<td>0</td>
<td>1</td>
<td>0</td>
</tr>
<tr>
<td>Thailand</td>
<td>1</td>
<td>112</td>
<td>8</td>
<td>1</td>
</tr>
</tbody>
</table>

(c) Pigs

<table>
<thead>
<tr>
<th>Origin</th>
<th>Cambodia</th>
<th>Laos</th>
<th>China</th>
<th>Vietnam</th>
</tr>
</thead>
<tbody>
<tr>
<td>Cambodia</td>
<td>3,320</td>
<td>2</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>Laos</td>
<td>0</td>
<td>2,633</td>
<td>2</td>
<td>4</td>
</tr>
<tr>
<td>Myanmar</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>Thailand</td>
<td>255</td>
<td>105</td>
<td>0</td>
<td>0</td>
</tr>
</tbody>
</table>
4.3. RESULTS

Table 4.9: The number of animals moved within and between countries from January 2007 to December 2009 by species.

(a) Buffalo

<table>
<thead>
<tr>
<th>Destination</th>
<th>Cambodia</th>
<th>Laos</th>
<th>China</th>
<th>Vietnam</th>
</tr>
</thead>
<tbody>
<tr>
<td>Cambodia</td>
<td>8 013</td>
<td>0</td>
<td>0</td>
<td>18</td>
</tr>
<tr>
<td>Laos</td>
<td>0</td>
<td>21 508</td>
<td>266</td>
<td>14</td>
</tr>
<tr>
<td>Myanmar</td>
<td>0</td>
<td>0</td>
<td>172</td>
<td>0</td>
</tr>
<tr>
<td>Thailand</td>
<td>0</td>
<td>1 027</td>
<td>1 681</td>
<td>0</td>
</tr>
</tbody>
</table>

(b) Cattle

<table>
<thead>
<tr>
<th>Destination</th>
<th>Cambodia</th>
<th>Laos</th>
<th>China</th>
<th>Vietnam</th>
</tr>
</thead>
<tbody>
<tr>
<td>Cambodia</td>
<td>116 885</td>
<td>0</td>
<td>0</td>
<td>1 266</td>
</tr>
<tr>
<td>Laos</td>
<td>0</td>
<td>19 310</td>
<td>188</td>
<td>67</td>
</tr>
<tr>
<td>Myanmar</td>
<td>0</td>
<td>0</td>
<td>43</td>
<td>0</td>
</tr>
<tr>
<td>Thailand</td>
<td>22</td>
<td>3 474</td>
<td>121</td>
<td>300</td>
</tr>
</tbody>
</table>

(c) Pigs

<table>
<thead>
<tr>
<th>Destination</th>
<th>Cambodia</th>
<th>Laos</th>
<th>China</th>
<th>Vietnam</th>
</tr>
</thead>
<tbody>
<tr>
<td>Cambodia</td>
<td>76 527</td>
<td>5 001</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>Laos</td>
<td>0</td>
<td>60 146</td>
<td>23</td>
<td>121</td>
</tr>
<tr>
<td>Myanmar</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>Thailand</td>
<td>34 950</td>
<td>29 751</td>
<td>0</td>
<td>0</td>
</tr>
</tbody>
</table>

Table 4.10: The number of buffalo, cattle, goats and pigs moving through a selection of checkpoints in Laos during 2009.

<table>
<thead>
<tr>
<th>Species</th>
<th>Buffalo</th>
<th>Cattle</th>
<th>Goat</th>
<th>Pig</th>
</tr>
</thead>
<tbody>
<tr>
<td>Buffalo</td>
<td>349</td>
<td>391</td>
<td>20</td>
<td>539</td>
</tr>
<tr>
<td>Cattle</td>
<td>585</td>
<td>507</td>
<td>90</td>
<td>10</td>
</tr>
<tr>
<td>Goat</td>
<td>779</td>
<td>343</td>
<td>0</td>
<td>77</td>
</tr>
<tr>
<td>Pig</td>
<td>239</td>
<td>263</td>
<td>0</td>
<td>24</td>
</tr>
<tr>
<td>5 143</td>
<td>5 698</td>
<td>131</td>
<td>14 616</td>
<td></td>
</tr>
<tr>
<td>Nasaythong</td>
<td>39</td>
<td>55</td>
<td>0</td>
<td>290</td>
</tr>
<tr>
<td>Naxai Neua</td>
<td>9 692</td>
<td>7 832</td>
<td>23</td>
<td>31 746</td>
</tr>
<tr>
<td>Nong Buathong</td>
<td>930</td>
<td>417</td>
<td>590</td>
<td>241</td>
</tr>
<tr>
<td>Phoukoun</td>
<td>1 658</td>
<td>1 172</td>
<td>42</td>
<td>378</td>
</tr>
</tbody>
</table>
CHAPTER 4. UNDERSTANDING LIVESTOCK MOVEMENT PATTERNS

Figure 4.13: Plots of the seasonal decomposition of movements of cattle in Cambodia and Laos from January 2007 to December 2009, showing (a) the raw count of cattle movements, (b) the seasonal component, (c) the trend and (d) the model residuals.

The overall set of data collected between January 2004 and December 2010, including all provinces in Cambodia (24) and Laos (17), and a number in Thailand (12), Vietnam (7), China (5), Myanmar (4) and Malaysia (1). The geographic range of movements identified during this study extended from Chin State in Myanmar and Kelantan in Malaysia through to Liaoning in China. This research was primarily conducted in Cambodia and Laos, so it is ironic that at no time during this period were there any records of cattle movements directly between Cambodia and Laos in either direction, and only very few of pigs moving from western Cambodia to southern Laos.

Perhaps the most important finding is the almost complete change in direction of movement of cattle since 2004 (Figure 4.14), although Phnom Penh remains a major destination for cattle.

When the movement records from Laos are also included, it combines to create a
4.3. RESULTS

Figure 4.14: Maps showing movements of cattle based on Cambodian movement records showing the direction of movement in (a) 2004, where the general movement is from northern parts of Cambodia to the south, and even as far as Malaysia, and (b) 2009, where the pattern has shifted such that most movements are from western Cambodia to the east, or from Cambodia into Vietnam.

Figure 4.16 shows the out-degree(a) and in-degree(b) respectively of the provinces associated with these cattle movements. It is important for the interpretation of these maps that centrality measures in this analysis are functions of the number of provinces, not the total number of movements or animals.
Figure 4.15: Map showing the movements of cattle in 2009 based on the complete set of movement records collected during the research showing the direction of movement. The main feature is the number of provinces from which animals are moving into China. The north-eastern most point is in fact referring to Chinese provinces further away than the Chinese border provinces of Yunnan and Guangxi.
Figure 4.16: Network centrality of provinces in Cambodia and Laos during 2009, showing the (a) out-degree, and (b) in-degree.
4.4 Discussion

This chapter has shown that analysis of livestock movement data can provide valuable information to aid in the understanding of transboundary animal diseases. This information is not just limited to describing general movement patterns, but can be integrated with suitable spatial data to show likely trade routes. Even with incomplete data, network analysis techniques allow the identification of high-risk areas for disease introduction and priority areas for surveillance in the context of reducing disease spread.

4.4.1 Western Australia

Analysis of the full set of cattle movements in Western Australia between July 2005 and June 2008 provides a number of insights into animal movement patterns and the potential for movement associated infectious disease spread in the state. Animal movement across the three regions of the state is not homogenous, with significant differences in numbers occurring between regions and between seasons. Even in the Southern Rangelands region, which had a relatively small amount of year round activity, showed a significant variation between seasons—over the three years of data, nearly 75 000 movements were recorded.

Centrality measures used in network analysis provide guidance to the areas where disease surveillance and intervention activities should take place. By analysing the network by node type it was apparent that producers are a critical node in terms of:

- Prevention activities—their high out-degree means a failure to identify disease in their own herd would lead to rapid dissemination;
- early disease identification—high betweenness allows them to take a key role in recognising disease and reporting it;
- disease control in the case of an outbreak—the high in-degree means that
almost regardless of where in the market chain disease occurred, producers would rapidly become involved.

Saleyards, export depots and feedlots are also important and should remain the focus of disease awareness and preparation activities.

Two shires (Swan and Harvey) are key locations for animal movement centrality most likely due to the location of a major saleyard in the former and a large saleyard and a cattle abattoir in the latter. A number of other shires have high out-degree (or influence) and would be expected to contribute to the rapid dissemination of disease. A smaller group of shires with high in-degree are more likely to be recipients of disease early in an outbreak, and this would justify these shires also being targeted for disease awareness activities. The very high betweenness of the two shires with the two major saleyards may indicate a risk to the industry if there was a disease introduction, as it would be almost unavoidable that both shires would be exposed to disease rapidly, and both saleyards operations suspended. Using network analysis concepts, potential strategies for reducing this risk would include the use of smaller local saleyards (to reduce in-degree and out-degree) or for the existing saleyards to accept ‘batches’ of stock from a limited number of shires in any given week (thus reducing the in-degree). It is unlikely that these measures are likely to gain much support from the livestock industry, and a more practical solution might be enhancing surveillance and producer awareness of disease at the existing saleyard locations.

Although the number of animals moved varies by season, the number of shires involved in livestock transactions does not vary as much. Nearly 30% of shires would have had livestock mixing within a three day period (which may represent movements of carrier animals not expressing clinical disease). If animal movements are allowed to continue unchecked by one month after introduction, three quarters of the states shires would potentially be infected. This is consistent with Gibbens et al. (2001), who identified the delay from introduction to reporting of disease as a key factor contributing to the size of an epidemic of FMD in the United Kingdom.
While homogenous mixing on individual properties may be a reasonable assumption (Danon et al., 2011), random mixing across entire shires is unlikely to occur. In many parts of Western Australia homogenous mixing is unlikely to occur on individual properties. However, this reflects the greatest possible impact and would provide guidance in the determination of disease control and movement restriction areas. Dubé et al. (2008) report that failing to account for the sequence of movement of animals from one farm to the next will increase the size of the giant weak component. A recent innovation has been the use of network metrics that account for the temporal sequence as well as directedness, such as infection chains (Dubé et al., 2008; Nöremark et al., 2011). Inferences about the risk of disease spread are limited by a lack of data on the numbers of animals moving between locations—network analysis is not seeking to address this, but modelling risk of disease spread by the use of probability distributions weighted by average volume could improve prediction outcomes.

Animal movement data may provide an insight into the activity of a livestock business, therefore privacy issues are an important consideration when accessing and analysing these data. Because of this, the full set of animal movement data was used for analysis but with the resolution of the origin and destination of the movement specified as the shire, not the individual property. The advantages of examining the movements by shire only (rather than individual farm location) are reduced computing pressure with only a limited resolution; and the lack of need for a homogenous national cadastral map dataset. The resolution of interstate movements is currently limited to the destination state. Future research could improve this with a whole country model that could be focused on the appropriate shires (or equivalent spatial unit) in each state as required.

There may be a number of factors contributing to the trend of increasing numbers of movements being recorded on the NLIS. The total size of the WA cattle herd did not change enough to explain the increase in animal movements. Because there are

---

30In fact, Danon is discussing data from the British Cattle Tracing System, where property sizes are much smaller than some in Western Australia.
exceptions for recording movements when animals move directly from property of birth to slaughter or live export not all movements are or will be recorded. Some livestock producers were initially resistant to the regulatory requirement to record all movements however the authorities were quick to prosecute those failing to comply. Additionally, although there is a cost involved, the increase in overall movement records may be due to the relative ease of applying an ear tag to a young animal instead of trying to apply them to a fully grown animal at the time of transport. The relatively constant number moving through saleyards over the three years may be a reflection on the attendance of regulatory authorities in the saleyards and strict enforcement of compliance during the introduction of NLIS.

Early criticism of the NLIS often included the high cost of the RFID tags and the equipment for reading them. Since its introduction, the system has been subject to a number of reviews. NLIS has been shown to have extremely high accuracy (no data were lost during an audit) (Eclipse Group, 2004) and to be cost effective (typically costing around AUD 2.50 per animal owned) (Alliance Consulting and Management, 2004). Using these assumptions WA producers invested approximately AUD 4.3 million in animal identification between 2005 and 2008, adding an average of AUD 1.10 per head to the cost of an animal movement. There is inadequate recent published data to allow a comparison of the total cost of livestock identification systems between countries, due in part to difficulties in assigning the costs borne by government and those paid by industry (Erik, 2011). The NLIS is unusual amongst animal tracing systems globally in that the full cost of its operation is funded by livestock producers, not government as happens in many other countries. More recent interest has focused on the benefit-cost of livestock tracing systems, but this is still limited (NAIS Benefit Cost Research Team, 2009).

Most centrality measures are based on evaluation of the geodesic (shortest path) between the start and end of a movement. This is appropriate in the absence of any other information. However this may not reflect the full range of possibilities of animal movement either geographically or purely based on network structure, where
multiple simultaneous paths are possible (Borgatti, 2005). To account for the former, rather than use euclidean distance the road network was incorporated using routing algorithms weighted to reflect factors likely impact on livestock transport. Straight lines by definition are the shortest line between two points, and thus underestimate the number of potential contacts animals moving between two shires might have. The ability to visualise the actual path taken by animals, and relate that to livestock holdings or areas of potential livestock (or human) density offers the opportunity for an enormous boost in assessing the risks associated with livestock movement. The functions developed in this work were designed to work with the best available data, be it a single point, a property polygon or an entire local government area. Further development of these techniques (including the use of probability distributions to introduce stochasticity into the choice of route) and more accurate tracking using individual property data are thus possible where privacy issues can be managed. The development of more sophisticated centrality measures for livestock movement such as disease flow centrality (Natale et al., 2011) and infection chains (Dubé et al., 2008; Nöremark et al., 2011) are subjects of research which will undoubtedly have application for the future analysis of NLIS movements.

As well as being a valuable exercise in its own right, the analysis of the WA NLIS data demonstrated the role network analysis can take in analysing data. This was further explored in the GMS data.

4.4.2 Greater Mekong Subregion

Consistent collection of data was difficult and best managed by using trained dedicated staff. Results such as those in Table 4.10 show the poor coverage of animal movements obtained by less dedicated staff. For example, more goats were personally observed passing through Ban Din Dam on one day than are shown in this table. Near-real time visualisation of movements was possible through a web-based map

\(^{31}\)Disease flow centrality attempts to incorporate temporality by only allowing edges to be included in the measure if at least one node is infected.
interface, however obtaining the data in near-real time was hampered by lack of computer network availability in agriculture offices in provincial parts of Cambodia and Laos. Between 2004 and 2009 a major shift in animal movement patterns in the region was observed. High numbers of livestock movements routinely occur from one side of a country to another, however a major limitation in this study was the limited data on what was occurring in neighbouring countries. A further limitation was that it was not possible from the movement permits to identify any locations where animals may have been rested enroute. Future work should aim to identify a region extending beyond national borders and seek cooperation in assessing movements at the supranational level.

The incomplete nature of the movement network data represents a form of selection bias, which is difficult to assess objectively. The advantage of using network analysis without any expectation of robust quantitative numbers is that it allows a focus on where movements are occurring, rather than what movements are occurring. Nonetheless, the lack of information of any sort between some locations raises questions about movements that may be occurring yet these provinces are not included in the network analysis. Any future work in this regard should thus focus on ensuring that some level of reporting (maybe just one week a month) is occurring from each province, and ideally from neighbouring country provinces. This would improve the validity of the network parameters by allowing relative and consistent comparisons between provinces.

The lack of consistent, well constructed road network data across the GMS limited the ability to apply some of the techniques developed for the WA data which assessed the actual distance travelled. This problem is not new, and a proposed approach to improving the nature of this information is discussed by Nelson et al. (2006), however the issues of ownership and funding are yet to be resolved\textsuperscript{32}. As distance travelled is a major contributor to the cost of livestock (discussed more in Chapter 5), it will be important to have information on road distance and road surface condition for

\textsuperscript{32}In fact, at this stage the greatest advances in resolving this issue appear to be either entirely public domain (OpenStreetMap) or private sector (Google).
modelling the relationship between animal prices and movement.

Network analysis again provides a useful framework for analysing these data. In Figure 4.14 and Figure 4.15 the relative out-degree of each province is shown as it provides a useful indication of the number of other provinces that are being reached from the source of the movement. Recognising that the quality and quantity of movement data varies between the provinces\textsuperscript{33}, rather than focus on the total number of movements, Figure 4.16 provides an indication of the level of risk each province poses for spreading disease to other provinces (the out-degree in (a)) or receiving disease (based on the in-degree, shown in (b)). Knowing the relative importance of the provinces as a source or recipient of disease is an important component of developing a risk model for disease spread in the absence of high quality data on animal movements.

Changing wealth in the region has most likely contributed to changes in livestock dynamics. Although the decision by the Vietnamese government in 2010 to allow imports of cattle through controlled pathways may have contributed to this change, in reality the illegal trade in animals already existed and barely changed. Conversely, a crackdown by Yunnan authorities on illegal imports through Northern Laos in 2010 appears to have changed the dynamics of animal movements along this pathway completely (Bourgeois-Lüthi N, personal communication, 2011). In 2011, the falling value of the Vietnam Dong and rising value of the Thai Baht has reportedly reduced the volume of livestock trade through Cambodia to Ho Chi Minh (Sieng S, personal communication, 2011).

The changes in livestock movements patterns has significant implications for the major disease control programmes in the region, which may need to adjust their control zones to match changes in livestock sources and destinations. Managing this will require ongoing information on movement patterns to be available.

\textsuperscript{33}And the reasons for this are beyond the control (or budget) of an Australian student.
4.4. DISCUSSION

4.4.3 Summary

The analysis of two data sets on livestock movement provides an opportunity to describe the issues associated with incomplete data collection, and how the observed situation (in this case the numbers of movements of groups and the structure of these movements) might vary from the true situation.

In the WA NLIS data, coverage of animal movements was likely to be excellent — with the exception of the animals moving from property of birth direct to slaughter or live export, all animal movements should have been recorded. The NLIS is relatively inexpensive by international standards, but it comes at a price that not all livestock producers where happy to pay. Authorities in WA have rigorously pursued those who failed to comply, causing further dissatisfaction. As such, it is highly likely that the observed situation provides an excellent fit to the true situation, and this allows the opportunity for robust analysis of the movements to guide policy, and rapid tracing forwards and backwards in the event of an outbreak. A limitation on this at the moment is the ability to make the data available for research is very limited due to concerns by industry about privacy. Unless suitable expertise and interest exist within the responsible department these data are likely to be under utilised. In addition, for investigating infectious disease spread it would be useful to be able to identify which animals had travelled as groups. With knowledge of movements at an individual premises resolution, it would be possible to analyse all movements from property A to property B on a given date as a single movement group. A further limitation on the NLIS is that while the database provides a very good representation of cattle movements, it contains no information about the cattle being moved. In some cases it is possible to infer some of this information from the destination (such as animals moving to an abattoir), but details such as age, sex, breed and purpose are not collected at any time.

In contrast to the WA data, in the GMS not only was the total amount of data likely to be a small percentage of the total number of livestock movements, but there are quite significant selection biases in the data that has been returned. Some
provinces returned almost no information on animal movements, others provided every record that they had, however field experience showed that large numbers of animal movements are not captured by the current movement permit system in Cambodia or Laos. This may be a reflection on the limited ability of the animal health departments in GMS countries to identify and pursue offenders, but it is likely to also reflect the restrictions that are placed on traders if they try to work within the regulatory framework, such as mandatory quarantine and the requirement for informal payments.

Although difficult to quantify, the information that is collected requires forms and labour and so comes at a price. Gathering and collating these data costs more again, although concerns about privacy are less and more information (such as details about the animal, purpose of movement and route) may be available and so the opportunity for more detailed analysis is greater\footnote{Lack of a consistent ‘standard’ format for reporting the additional information currently limits the ability to analyse this data.}.

A number of opportunities exist for better integrating legislative requirements and data collection. Currently the data is required to be collected, but no use is made of it. Relatively inexpensive solutions now exist which would allow the data captured on paper recording forms to be gathered into a central database at the time of data collection (some specific examples are discussed in Section 7.2.1). This would allow regular analyses to be used to detect changes in animal movement patterns.

During this research both the prospective animal tracking study and the movement permit analysis suffered in part due to a limited geographic range. The range of many livestock movements in the GMS is larger (and involves more countries) than typical research projects can hope to cover, and this is an important consideration for the design of any further attempts to capture animal movements.

Better animal identification is becoming a priority in the GMS. A key benefit of the NLIS in Australia is the agreement between all states on how animals should be identified and a common principle (enshrined in law in all states) of not tam-
4.4. DISCUSSION

 interoperating with identification devices. Although decisions on how to manage animal identification will be the responsibility of individual nations, the speed with which animals transit borders makes the development of regional standards for animal identification a priority.
Chapter 5

A Description of Livestock Market Prices in Cambodia and Laos, 2007–2010

“...yet, if he sells it at a price which does not allow him the ordinary rate of profit in his neighbourhood, he is evidently a loser by the trade; since, by employing his stock in some other way, he might have made that profit.” (Smith, 1776)
5.1 Introduction

The key incentive to trade any commodity is the opportunity to realise profit. This becomes all the more important when the costs associated with the transaction are higher, such as when expensive long distance movements are undertaken, interest is accruing on the investment, or the commodity may lose value during handling—such as livestock becoming ill and losing weight or even dying.

The previous chapter showed not only the large number of movements of livestock which occur throughout the Greater Mekong Subregion (GMS), but also by use of the network analysis framework, the relationship between provinces in terms of whether they are a source of livestock or a destination. The collection of these data was problematic due to the dependency on livestock moving through legal pathways (and that a record of movement was created to accompany them) and the compliance of provincial staff in submitting these data for analysis. Although the former problem was insurmountable within the scope of this work, the use of other approaches such as sociological investigations of the motivation of traders (as discussed in Section 2.5) may provide better insights. The provision of paid incentives improved the compliance of the provincial staff. Both of these problems justified the interest in identifying a proxy measure of livestock movement and attempting to validate that with the known movements to see if it could be usefully used to approximate patterns of livestock movement. Given the need to realise a profit in order for trade to occur, the hypothesis to be tested is whether the difference in price of livestock between provinces varied in a predictable manner. The prices were the reported market price on a day in each province, and not necessarily associated with animal movements.

The primary objective of this part of the work was thus to determine livestock prices at a province level across Cambodia and Laos. In Chapter 6 the difference in price between provinces derived from this work are modelled against the movements of animals between the province based on the results from Chapter 4 to determine the
5.1.1 Overview of methods

Livestock are not just a homogenous unit, but come in a range of ages, body conditions, two sexes\(^1\), a number of breeds or mixes of such and often have different purposes. As a result, a number of steps were required to collect and analyse price data, which are summarised here.

First, it was necessary to identify which parameters would be important to collect along with price to enable comparison between prices for different commodity classes. This was done using consultation with animal health staff in Cambodia and Laos initially, and then canvassing the opinion of a larger group of experts working in the region.

Once this was done, provincial animal health staff in Cambodia and Laos were trained to collect data and report them to the central office on a regular basis. At a larger scale, some data on prices in China (Yunnan and Guangxi) from 2000–2008 were also obtained\(^2\) and the change in prices over time provides a valuable insight into the increasing demand in China and corresponds with the change in direction of animal movements in the GMS.

The data that were collected were then analysed with the intention of identifying provinces with higher and lower animal value and quantifying the differences between provinces. To account for the different categories on which the prices were based, the analysis used a multilevel linear model with province as a random effect. The use of a constant slope model with varying intercepts made it possible to determine a single value for a standardised livestock unit in each province in Cambodia and Laos.

---

\(^1\)Arguably, sex could be classified into more than two categories but desexed animals were not considered separately in this study.

\(^2\)These data were extracted from Chinese Year Books for this period by Chang Cai in 2010 and generously provided for illustration of this increase.
5.2 Establishing Standard Categories for the Price Determination

5.2.1 Introduction

The valuation of livestock is considered an artform by some, black magic by others. Collection and analysis of livestock prices is made difficult by the complexity associated with the specification of the commodity of interest. The bottom line is that the level of skill required for livestock valuation is uncommon and dependent on experience and a contemporary understanding of the market. However, for this research, to be able to analyse price it was necessary to limit the classes of livestock to those which may be clearly understood and consistently applied by relatively untrained data collectors—but not simplify the process so much as to lose the value by collecting a meaningless group of numbers. In order to do this, opinion was sought on how best to describe the value of livestock.

There is little published literature on the standard approaches to valuation of livestock in the GMS (or anywhere). A comprehensive system such as that previously described for the National Livestock Reporting Service (NLRS) in Australia may be unaffordable and unnecessary in the GMS (Meat & Livestock Australia Limited, 2005a,b). Existing agricultural market price systems in the GMS offer no clear suggestions on their methodology. For example, the Cambodian Agricultural Market Information System\(^3\) (CAMIS) simply reports on prices for four livestock categories, ‘Carcase’, ‘Live Chicken’, ‘Live Duck’, ‘Live Pig’, but does not explain how it comes to these values. A simplified system using a smaller number of parameters is required, but it is important then to identify which parameters should be recorded within the system.

\(^3\)http://www.agriculturalmarketinformation.org.kh/en/
5.2.2 Methodology

Consultation

In order to allow analysis and useful interpretation the data needed to be collected for clearly defined categories or classes of animals. It was important to establish that the categories to be collected were likely to be understood in the context of local animal management, therefore consultations included discussion with local animal health staff in Cambodia and Laos and international researchers working on animal production research in GMS countries.

All the parameters were recorded as categorical data to simplify recording and improve consistency (for example, age was recorded as calf, weaner, adult etc instead of months or years).

Expert Opinion

Early in the price collection exercise attempts to analyse the data revealed some issues with the data categorisation. To gain further insights into the best way to identify parameters of importance in collecting market prices expert opinion was sought. Expert opinion elicitation is a technique that is especially suited to problem definition in the absence of objective data.

An invitation was sent to eighteen people living and/or working in the GMS in livestock production or animal health to complete a questionnaire via a customised online survey system\(^4\). Participants were purposively selected based on having established some credentials for working in the GMS on projects related to livestock and their current involvement in the region. All participants were known to the author either through prior contact or by knowledge of their work in the region. Some of the participants had been involved in the consultations as described above.

To address the limited existing knowledge, the methodology deliberately chose to

\(^4\)http://trade.animalhealthresearch.asia/expert/
5.2. ESTABLISHING STANDARD PRICE CATEGORIES

use open questions to encourage suggestions from participants. The results were anonymous—although it was possible to identify if users had answered a question, the survey system did not allow the linking of answers to participants. As some participants may have had limited knowledge in some of the question topics it was not a requirement to answer all questions.

The results of this questionnaire were collated and common themes examined and described in a discussion document. This document was circulated by email to those respondents who had completed the questionnaire for further feedback.

5.2.3 Results—Consultation

On the basis of the consultation, data on the following variables (for each of cattle, buffalo and pigs) were collected along with the price:

- currency (Cambodian Riel, Laos Kips or United States Dollars),
- unit of sale (per head or per kilogram),
- date (on which the price was valid, at least month and year),
- location (at least province and preferably district),
- sex (male or female),
- age (calf/piglet, weaner, juvenile, adult, mature/old),
- condition (good, medium or poor),
- breed (imported, local breed or cross-breed) for pigs only.

The variables and their definitions for cattle and buffalo are shown in Table 5.1. The categories for pigs are shown in Table 5.2. Without taking into account the date or location of the price data collection the number of combinations of predictors that this allows is illustrated in Figure 5.1.
### Table 5.1: Definitions for price categories for cattle and buffalo.

<table>
<thead>
<tr>
<th>Variable</th>
<th>Type</th>
<th>Values</th>
<th>Notes</th>
</tr>
</thead>
<tbody>
<tr>
<td>Age</td>
<td>Nominal</td>
<td>Calf</td>
<td>A calf was a suckling animal, a weaner in the immediate post-weaning age category, adult was a fully grown animal and mature was an elderly animal (most often old cows or bulls).</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Weaner</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>Juvenile</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>Adult</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>Mature</td>
<td></td>
</tr>
<tr>
<td>Sex</td>
<td>Nominal</td>
<td>Male</td>
<td>Entire/desexed was not recorded.</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Female</td>
<td></td>
</tr>
<tr>
<td>Condition</td>
<td>Ordinal</td>
<td>Poor</td>
<td>Good condition would equate to about a condition score 3 on a 1-5 scale in Australia</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Medium</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>Good</td>
<td></td>
</tr>
<tr>
<td>Location</td>
<td>Nominal</td>
<td>Province Name</td>
<td>24 in Cambodia, 17 in Laos</td>
</tr>
</tbody>
</table>

### Table 5.2: Definitions for price categories for pigs.

<table>
<thead>
<tr>
<th>Variable</th>
<th>Type</th>
<th>Values</th>
<th>Notes</th>
</tr>
</thead>
<tbody>
<tr>
<td>Age</td>
<td>Nominal</td>
<td>Piglet</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>Weaner</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>Grower</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>Adult</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>Mature</td>
<td></td>
</tr>
<tr>
<td>Sex</td>
<td>Nominal</td>
<td>Male</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>Female</td>
<td></td>
</tr>
<tr>
<td>Breed type</td>
<td>Nominal</td>
<td>Local Breed</td>
<td>Exotic breeds refers to landrace/large white type breeds</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Exotic Breed</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>Crossbreed</td>
<td></td>
</tr>
<tr>
<td>Condition</td>
<td>Ordinal</td>
<td>Poor</td>
<td>This was simplified to account for the extra complexity of recording breed.</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Good</td>
<td></td>
</tr>
<tr>
<td>Location</td>
<td>Nominal</td>
<td>Province Name</td>
<td>24 in Cambodia, 17 in Laos</td>
</tr>
</tbody>
</table>
5.2. Establishing Standard Price Categories

Price

Cattle
Calf
Weaner
Juvenile
Adult
Mature
Buffalo
Pigs
Good
Medium
Poor
Male
Female
Male
Female
Male
Female
Good
Medium
Poor
Male
Female
Local Breed
Cross Bred
Exotic Breed

Figure 5.1: A graphical representation of the complexity associated with categorisation of livestock by species, age, condition and sex (and breed for pigs).
5.2.4 Results—Expert Opinion

Eighteen individuals were invited to participate in the expert opinion questionnaire. Of those invited, 13 clicked on the supplied link to view the questions and 10 of these 13 (77%) answered some or all of the questions. The anonymity feature of the survey system meant it was not possible to link responses to individuals. These 10 included an economist, two research project managers, an animal production researcher, an animal disease researcher, a zoonotic disease researcher, a veterinary consultant, an epidemiologist and the head of a government department of animal health. These respondents worked primarily in Laos (3), Cambodia (2), Thailand (1) and Australia (4) and had collective recent experience in Laos (5), Cambodia (4), Thailand (3), Vietnam (2), Myanmar (1) and Indonesia (1). As the survey contained open questions a summary of the results is provided below.

Age

There was little support for the suggestion of recording the age in months/years, although two respondents suggested that for pigs it may be possible. It was reported that although some traders would probably mouth cattle to age them there was no clear understanding on what the implications of this process would be. It was likely that it was purely to ascertain whether the animal was still growing (limited permanent dentition) and hence would have some value-adding potential, or whether it was now at a mature weight.

The weight of the animal was far more important in the decision on the value of the animal, but tended to indicate that the weight was important for determining the final value of the animal based on a per kilogram liveweight price the trader was willing to pay. Consequently, in Laos pigs were normally weighed as part of the sale process; however, in Vietnam the skill of the trader was in their capacity to accurately gauge the amount of 1st grade meat\(^5\) and they could expect to recover

\(^5\)1st grade meat is the best eating meat derived from musculature, as opposed to bruised meat, offal and other carcase products.
A number of respondents questioned whether the ultimate end-use of the animals was likely to be more important in driving the price—if they were destined for immediate slaughter, had alternative uses (draught power or breeding) or offered potential to value add through growth.

**Sex**

Where an animal is being purchased for breeding sex is important—and the value of the animal is likely to be based on capacity to fill this purpose. While a preference for castrated pigs was raised, the supply was considered to be limited due to a lack of people who were able to castrate their animals. Otherwise, where an animal was being purchased for slaughter, gender was seen as less important.

Price data collection commenced prior to the expert opinion, and early analysis of this data suggested that the sex of the animal appeared to have a consistent effect on the price paid. This information was communicated to participants as part of the questionnaire. A number of respondents suggested that this is mainly due to the impact of gender on size and ultimately weight. Therefore it would seem that capturing sex of the animals may be less important than originally thought, if the purpose of use (for example, breeding would often imply female) is known.

**Breed**

Most respondents thought that breed was important for pigs, but less so for cattle. Suggested categories for breed included local breed, exotic breed (for example ‘large white’ type animals), or a crossbred animal (in Cambodia).

Terminology is important here, as in Cambodia the term ‘imported’ was used to describe animals from improved genetic lines (exotic breeds), but in Laos it was suggested that due to differences in feeding, the exotic breeds were also valued

---

6Presumably to reduce boar taint in the meat.
differently based on whether they were raised in Thailand (and imported) or raised in Laos.

In cattle, as for sex, it appeared that the biggest impact of breed was the size of the animal, and possibly the carcass recovery. Respondents with experience in both Vietnam and Cambodia noted that the addition of brahman genetics was valued over the locally adapted cattle breeds.

**Condition**

The key message from responses about condition was that ultimately it was related to how much meat a trader could expect to net from the animal. Responses were evenly split over the number of ‘categories’ of condition that should be used—half thought that three categories would be adequate:

- Poor, skinny, thin, less than ideal
- Good, OK, Medium, ideal
- Fat, more than ideal

The remainder thought that expanding it to five categories was necessary to record the extremes of condition, and suggested ultimately that ‘too thin’ or ‘emaciated’, and a ‘too fat’ category should be included in the mix to allow these animals to be appropriately described.

Respondents were asked for their opinion on introducing a standard system for scoring animals. Amongst other comments the point was made that condition scoring was a useful animal management tool, but for trade ultimately weight was likely to be the most important variable influencing price, and this was normally a subjective assessment made by the buyer (and to some extent the seller). In the trade context, it was pointed out by one respondent that “it seems to work—so why change it?”

---

7 Fat in Cambodia and Laos would be equivalent to around a score 3–3.5 out of 5 by Australian standards of condition scoring.

8 This is, of course, heresy to a research student.
5.2. ESTABLISHING STANDARD PRICE CATEGORIES

Purpose

Throughout the responses the ultimate use or intention for the animal (purpose) was referred to as being highly relevant to the price being paid. The most commonly suggested groups were summarised as:

- **Direct to Slaughter**—animals that were destined for immediate slaughter, so the price is probably directly related to the carcass attributes of the animal and the prevailing market prices for meat.

- **Breeding or Draught power**—animals which have an inherent value based on characteristics such as sex (for breeding) and breed or age for draught power.

- **Growing**—young animals which were ultimately destined for slaughter, but the opportunity exists to make money by feeding them. The value of these animals may vary depending on their current weight, breeding and feed availability.

Secondary feedback

Two participants provided further responses to the summary of the results. As a result of this process, extra data were collected on purpose of the purchase when collecting data from a pre-existing source.

5.2.5 Discussion

The aim of this work was to gain a sense of what information (variables or factors) about livestock would need to be collected alongside price, to make sense of the information—conceivably a standard economic problem, but there was little literature about how best it should be approached. Due to time pressures, by the time an expert elicitation was undertaken data collection had already commenced and the data being collected was very difficult to compare due to the large range of
covariates. The use of an expert elicitation panel was an appropriate way to begin to develop an understanding on issues affecting price of livestock in the GMS, but the use of a purposive selection process was likely to have resulted in some bias about what was important. With more time and resources a random cross-sectional sample of active participants in the market would have provided a much more robust set of data from which to base further data collection activities. As well as limited resources available to conduct such a survey, it was not possible to obtain a comprehensive list of the target population from which to develop a sampling frame to ensure a random sample. One of the broader aims of the project was to develop a better understanding of the trading practices in the GMS to hopefully overcome this limitation in the future.

During the course of the research it emerged that traders tend to arrive at their offer price for purchase by estimating the yield of saleable meat from an animal. They then offer a price based on their likely sale price minus the costs associated with transport (and slaughter if they are the final stage) plus their intended margin (Bourgeois-Lüthi N, personal communication, 2009; Scoizec A, personal communication, 2010). On this basis, the ideal approach to accurate livestock pricing would be to divide the animal up into the component pieces, put a price per kilogram on each part and weigh them. It was clear from survey respondents that this complex process is what is going on in the heads of a trader when they set their prices. However, the system is complex and relies on subjective assessments of the value of the animal by highly skilled practitioners. It was necessary to objectify this assessment process to allow the best chance of inexperienced staff across a number of countries, cultural groups and species being able to use it.

The expert opinion suggested a number of courses of action for ongoing data collection:
5.2. ESTABLISHING STANDARD PRICE CATEGORIES

Limit data collection

Only recording prices for transactions relating to a finite number of animal classes (for example direct to slaughter) would substantially simplify the process. Animals going to slaughter could have a per kilogram price, breeding cows a per head price, maybe considering condition and breed. Simply obtaining regular market prices for meat would be relatively straightforward in Cambodia. In Laos, market price for meat in each province is normally prescribed by authorities and so while it would be almost trivial to collect, it would be highly unlikely to accurately reflect the cost of buying and selling animals.

A drawback to this approach would be the loss of information about the other movement types.

Decision tree to guide simplified data collection

Providing a decision tree for information required based on animal purpose would add some complexity to the user and to the database system, although neither of these is insurmountable. An example tree is shown in Figure 5.2.

The potential extra complexity for users could be mitigated through the use of an interactive interface—for example, choose a purpose, then choose from the next set of options etc until the report is completed. This user side application would make submission easier, and increase the likelihood of submissions being useable and repeated—no SMS messages coming back asking for it to be done again, or paper forms with non-sensical data being returned.

The limits would be the cost of developing an application that could accept the options and encode the results into an SMS; and supplying suitable hardware (for example, a smartphone); and possibly the difficulty in evaluating the groupings. To investigate the cost of developing a suitable application for SMS encoding a visit to the Innovative Support To Emergencies Diseases and Disasters⁹ (InSTEDD)

⁹http://www.instedd.org/
CHAPTER 5. LIVESTOCK MARKET PRICE

Project’s Phnom Penh office (known as iLab) was undertaken. InSTEDD offered the opportunity to create mobile ‘non-smartphone’ applications to assist with encoding data. At the time of the visit (early 2008) the iLab was only just developing, and although they appeared interested in this project, they did not appear to have adequate programming resources available at the time for a new project.

Persist with the broad approach

Persisting with training and instruction on the use of the current system has some obvious benefits albeit at the risk of less information. These benefits include the comprehensive nature of the data that are being gathered, and the time and costs associated with developing alternatives. Under the circumstances, as data collection was already underway and concerns about whether data collected in a different
manner could be utilised the status quo was retained.

It is necessary to collect enough information about the commodity to allow a reasonable comparison of the prices of like products and determine how much difference in price is enough to cause a net flow of animals. However, this must be done without compromising the cooperation of contributors by making the requirements overly onerous or time consuming. In econometric terms this concept is known as the efficiency of a parameter.

This research attempted to use a relatively complex system to allow categorisation of data. This was difficult to explain to potential users, and required individual coaching. None-the-less, the Cambodian Agricultural Market Information System (CAMIS) SMS interface discussed in Section 2.6 used a much simpler approach (with just four livestock categories, ‘carcase’\(^{10}\), ‘live pig’, ‘live chicken’ and ‘live duck’) and still experienced an error rate of around 20–25%.

### 5.3 Collection and Analysis of Prices

#### 5.3.1 Introduction

A number of different technologies were used to collect data. A monthly report was requested from staff in each provincial animal health office, which could be provided on paper, directly into an online database or submitted by text message.

#### 5.3.2 Methodology

**Data Capture System Design**

The database structure was developed to receive ‘reports’ (based on a species, location and date), and any report could have one or more ‘prices’ associated with it (a

---

\(^{10}\)Email requests in June and July 2011 for more detail on what ‘carcase’ meant were received (an automated message receipt email was sent back) but no further information was forthcoming.
record of a price with the associated description of the livestock). Although most location data was recorded at the province level, the database could also accept a more detailed location description (such as district, commune and village).

A web based database interface was developed that allowed for the collection of data into a cross tabulated table structure which varied according to the species. The table was dynamically generated according to the required detail on parameters as discussed on page 122 which could be adjusted by the system administrator.

Training

Training courses on collecting market price data were conducted in Cambodia and Laos for one staff member from each provincial office, concurrently with the training on the animal movements database discussed in Chapter 4. Staff were brought to the national capital and provided with a background to the reasons for this study, how to collect and describe the data and how to enter it into the proforma sheets, or the online database interface if they had internet access.

The opportunity was also used during the abovementioned course to demonstrate the use of a mobile phone short message service (SMS) system for submitting market prices, however at the time of the training the system was still under development. Consequently, individual training on using the SMS system was delivered to three provinces in Cambodia.

Collection using Paper reporting forms

Provincial staff in Cambodia and Laos collected market price data by visiting markets or talking to traders on a monthly basis for twelve months. Data were recorded on a standardised reporting form. Where possible this was entered directly into the online database created for this purpose. An example of the online web page for price entry is shown in Figure 5.3. Most reports were entered on a printed paper form which was returned to the national office for subsequent transcription.
5.3. COLLECTION AND ANALYSIS OF PRICES

Figure 5.3: A screenshot of the market price data entry form. The prices are the average market price on the day, not an individual animal price. This example shows an outlier price which was adjusted during analysis (the male good condition calf price).

Each report contained details of the country and province, the date of the report and the location that the prices were obtained. It also contained a number of prices for livestock depending on the various parameters described below.

The intention was that these data would be collected in each province at least once a month for each province over a period of at least 12 months, and ideally two years.

Collection by SMS

Mobile telephone networks are well established throughout the GMS—most larger villages and towns are covered by at least one mobile telephone company—and the uptake of mobile telephones is increasing. At the commencement of this project there was limited data capacity in these mobile networks, especially outside major cities. As a result the development of a data collection system using the Short Message Service available on all modern mobile phones was investigated. The use of SMS offers an opportunity to capture data more rapidly, cheaply and without the need for data to be re-entered. On the other hand, the limited number of characters
per message (160 ascii characters) and high cost of sending messages to numbers outside the country also presented some challenges.

The system was based on three components:

1. The development of an encoding system to allow animals to be described in enough detail to make sense of price information. This is described in greater detail in the SMS Manual which was developed for users, an English version of which is attached in Appendix C.

2. The design, construction and installation of a SMS base-station in each country. This was necessary as sending an SMS internally in Laos and Cambodia is extremely cheap compared to trying to send one to an international phone number.

3. The development of a suitable database and query interface to decode and store the data in a central repository.

The SMS base-stations were created using Teltonika ModemCom-G10 GSM modem\textsuperscript{11} connected to a salvaged laptop (for example, an IBM thinkpad with a broken screen donated by DAFWA) running the FreeBSD operating system and using the smstools3\textsuperscript{12} open source library and a customised perl script to receive the messages, transfer them to the central database (see below) and send an acknowledgement to the sender.

The central database for prices was the same one as for the online price data entry, based on a server in Australia. This allowed analysis of all price data concurrently, regardless of the method of submission. A script was written to parse the data from the SMS message, conduct validity checks and if suitable insert it into the database, or if not suitable advise the submitter why it was not satisfactory.

\textsuperscript{11}\url{http://www.teltonika.it/}
\textsuperscript{12}\url{http://smstools3.kekekasvi.com/}
5.3. COLLECTION AND ANALYSIS OF PRICES

Standardising monetary rates

Livestock trading in the GMS normally takes place in the most appropriate currency for the local market (for example, Lao Kip in Laos) but also includes the possibility for using neighbouring country currencies. Some traders actively hedge on the rate differences by buying and selling in different currencies (for example using Thai Baht in Laos and selling in Vietnamese Dong) to take advantage of the differential (Bourgeois-Lüthi N, personal communication, 2010).

Data were collected in the unit of the transaction or convenience for the reporter. This meant that prices could be reported in Cambodian Riel (KHR), Lao Kip (LAK) or United States Dollars (USD). The reporting system could also accept Thai Baht (THB) or Vietnamese Dong (VND) if required. The report could also specify a price per head or per kilogram. Prior to data analysis it was necessary to standardise these prices.

It was not possible to identify what local trade related rates were as this value fluctuates based on a number of criteria including convenience (rounding to nearest convenient monetary unit) and the strength of any parallel economy. As the system allowed for prices to be recorded in a range of currencies, these were standardised to US Dollars. As a result, the rates were calculated by using a daily average rate for transactions at the interbank rate. This was obtained using currency tables published by OANDA\textsuperscript{13}.

Data analysis

As a result of the market price data collection exercise, a substantial pool of price records were available covering cattle, buffalo and pigs collected in Cambodia and Laos between 2008 and 2010. These data were classified according to the age, sex, condition (and in the case of pigs, breed type) for each province each month of the year. A multilevel linear model using the natural logarithm of price was used. In this

\textsuperscript{13}http://www.oanda.com/historical-rates, accessed 2010-03-02.
model the independent variables included the age, sex and condition of the animal (and breed type for pigs) and month of the year as fixed effects, and the province as a random effect. All of the predictor variables were fitted as categorical variables. This is shown in Equation (5.1), based on that proposed by Laird and Ware (1982). In this model:

\[
Y_{i,j} = \beta_0 + \beta_1 \text{age}_{i,j} + \beta_2 \text{condition}_{i,j} + \beta_3 \text{sex}_{i,j} + \beta_4 \text{month}_{i,j} + \gamma_j + \epsilon_{i,j} \tag{5.1}
\]

\[
\gamma_j \sim N(0, \sigma_{\text{prov}}^2)
\]

\[
\epsilon_i \sim N(0, \sigma_i^2)
\]

\[i = 1, \ldots, n \text{ records, and}
\]

\[j = 1, \ldots, m \text{ provinces}
\]

\[Y_{i,j} \text{ is the natural logarithm of price in the } i^{th} \text{ record from the } j^{th} \text{ province,}
\]

\[\beta_0 \text{ is the intercept,}
\]

\[\beta_1, \ldots, 4 \text{ are the coefficients allowed to vary across age, condition, sex and month,}
\]

\[\gamma_i \text{ is the province in which the } i \text{ report was recorded,}
\]

\[\epsilon \text{ is the error term for each report } i.
\]

Simple statistical tools (summaries and plots) were used to visualise the data in the database and via a database connection from R and outliers were investigated. In a number of cases, where clear data entry errors could be identified adjustments were made to correct for this. For example, a price in a Lao province that was recorded in Cambodian Riel and was twice as high as the next highest Lao price when converted to dollars was considered likely to have been the price in Lao Kip. Similarly, where the price of a calf in good condition was more than 10 times the price of a calf in medium condition as shown in Figure 5.3 it was suspected that the price included an inadvertent extra zero and this was removed.
A query to extract the 11 variables shown in Table 5.3 for cattle from the database was validated in SQL and then used in R (using the RPostgreSQL package) to create a data frame.

Table 5.3: The data fields associated with price data extracted from the database for analysis.

<table>
<thead>
<tr>
<th>field</th>
<th>data type</th>
<th>Notes</th>
</tr>
</thead>
<tbody>
<tr>
<td>country</td>
<td>categorical</td>
<td>The iso3 code for Cambodia or Laos.</td>
</tr>
<tr>
<td>province</td>
<td>categorical</td>
<td>The province where the report data were obtained.</td>
</tr>
<tr>
<td>reportdate</td>
<td>date</td>
<td>The date the report was made.</td>
</tr>
<tr>
<td>year</td>
<td>categorical</td>
<td>The year that the report was made as a categorical variable.</td>
</tr>
<tr>
<td>quarter</td>
<td>categorical</td>
<td>The quarter that the report was made as a categorical variable.</td>
</tr>
<tr>
<td>month</td>
<td>categorical</td>
<td>The month of the year that the report was made as an ordinal variable.</td>
</tr>
<tr>
<td>species</td>
<td>categorical</td>
<td>Cattle or pigs.</td>
</tr>
<tr>
<td>sex</td>
<td>categorical</td>
<td>Male or Female</td>
</tr>
<tr>
<td>age</td>
<td>categorical</td>
<td>A category with 5 levels—these were not treated as ordered as it was not clearly linear—price may reasonably be less for very old animals.</td>
</tr>
<tr>
<td>condition</td>
<td>ordinal</td>
<td>These data were considered to be ordered.</td>
</tr>
<tr>
<td>usd</td>
<td>numeric</td>
<td>The price in United States Dollars after conversion from the original currency as discussed in Section 5.3.2</td>
</tr>
</tbody>
</table>

Due to the large number of categorical variables data exploration was carried out using a mixture of plots (histograms, boxplots and dotcharts) and summary data tables. Investigation of the response variable suggested a natural logarithmic transformation was appropriate to manage the right skewed distribution of the data. There were many more reports from Cambodia than Laos, but there also appeared to be much greater range in the number of prices records included in each monthly report. The presence of outliers and homogeneity of variance was assessed using scatterplots and a composite boxplot of all factors. Additionally, the variance of the log transformed values for each group was assessed. Data were inspected for evidence of heterogeneity of variance by looking at crude estimates of variance before the multivariable model was run, and visual inspection of a scatter plot of fitted values vs standardised residuals after the multi-variable model was run (Perkins, 2003). Collinearity of predictors was assessed using pairwise scatterplots, correla-
As the objective was to identify if animal value varied between provinces, it was necessary to account for the range of values, the range of predictors and the province level variation. By using the province as a random effect in a varying intercept model, it allowed evaluation of the fit of the model and the ability to estimate intercept values for each province. Intercept values estimate the price for animals that were in the reference categories for all of age, condition, sex, and month. For cattle, this was female calves in poor condition in January. For pigs, this was female crossbreed piglets in poor condition in January. Due to >90% of values coming from 2009, year was dropped from the model to reduce complexity. The primary objective was to estimate differences in sale values for each of cattle and pigs between provinces. As a result, because there was high collinearity between province and country, very little in the way of border controls between the two countries and only one movement of pigs was recorded between them, country was not included in the model. As the hypothesis being investigated was determining value at provincial level, province was added as a random effect to the model. All remaining explanatory variables shown in Table 5.3 were considered for inclusion in the model as fixed effects. This model was effectively selected a priori through the use of expert opinion (see Section 5.2.2). A number of additional models were tested using approaches recommended by numerous authors (Dohoo et al., 2009b; Gelman and Hill, 2007c; Pinheiro and Bates, 2000). These include:

- Dropping explanatory variables in a backward stepwise manner (as described by Dohoo et al., 2009c, p 386) and comparing plausible models by using Akaike’s Information Criteria (AIC) to find the models with the lowest AIC;

- evaluating the standard errors associated with the coefficients and assessing the likely biological role of the explanatory variables for those where the coefficients are less than two standard errors from zero;

- comparing models using month and those using quarter;
comparing plots of multivariable models with differing combinations of predictor as a random effect in a varying intercept/varying slope model to identify those where slopes varied significantly.

Plausible two-way interactions were investigated (age × condition, sex × condition and month × condition) to check if any interactions that were significant for inclusion in the final model.

The modelling was conducted using the lme4 package (Bates et al., 2011) in R (R Development Core Team, 2011). Restricted maximum likelihood was used to fit the models. Simulation of the random effects (derived using best linear unbiased prediction) was undertaken to identify the robustness of the estimates. Results are presented as regression coefficients and standard errors. Further model checking included comparing standardised residuals with fitted values. Variance ratio and intra-class correlation were calculated to assess the contribution of the province to the price values from within the province using the methods described by Gelman and Hill (2007b) and Dohoo et al. (2009b) respectively.

The prices in each province for each month were predicted using the random effects from the best fit multilevel linear model. The results figures for cattle and pigs are shown adjacent on page 158 to aid comparison.

Spatial autocorrelation of the prices was investigated by using empirical semivariograms. The method is as described by Berke (2004) and Berke et al. (2007) using the geoR (Ribeiro Jr. and Diggle, 2001) package in R. A empirical semivariogram was computed using Hawkins and Cressie’s robust modulus estimator and using a constant trend spatial model. This was fitted using weighted least squares with a spherical covariance model to estimate the covariance parameters for the prices. Although by design the number of spatial prediction locations was limited (only province centroids were used) anisotropy was also evaluated using directional semivariograms. To aid visual interpretation of the range of prices across the two

\textsuperscript{14}Anisotropy is variation dependant on the direction of measurement, and is described in Pfeiffer et al. (2008a, p33).
countries, universal kriging was used to estimate price values every 10 kilometres for
an adult cow in medium condition (and an adult pig in good condition). This was
then used to generate a predicted price for each month across Cambodia and Laos,
which was displayed as an isopleth map.

5.3.3 Results

Data collected

During the period January 2008 to February 2010, 1 534 reports of provincial market
prices were collected. Slightly more than half of these were collected in Cambodia,
as shown in Table 5.4. On one report province was not indicated (although it
was in Laos, no further details were available), so it was discarded for analysis at
province level. Two reports of meat prices (species unknown) from Cambodia were
also discarded.

Table 5.4: The number of reports of market prices received by species and country from January
2008 to February 2010. (Each report may contain more than one price record)

<table>
<thead>
<tr>
<th>Country</th>
<th>Buffalo</th>
<th>Cattle</th>
<th>Pig</th>
<th>Total</th>
</tr>
</thead>
<tbody>
<tr>
<td>Cambodia</td>
<td>232</td>
<td>292</td>
<td>325</td>
<td>849</td>
</tr>
<tr>
<td>Laos</td>
<td>304</td>
<td>276</td>
<td>105</td>
<td>685</td>
</tr>
<tr>
<td>Total</td>
<td>536</td>
<td>568</td>
<td>430</td>
<td>1 534</td>
</tr>
</tbody>
</table>

Data covered a period of 25 months in Cambodia, and just over 22 months in
Laos. Each report could contain between 1 and 60 individual prices for various
combinations of age, sex, condition and (for pigs) animal breed.

Data submission in Cambodia was affected by misunderstandings over the financial
arrangements for the provision of these data. The numbers of reports submitted by
province are shown in Figure 5.4.

An image plot (Figure 5.5) was developed which could be used to quickly assess
whether a province had made a report, and to some degree the strength (number of
different classes) of the report. This also shows that from mid 2008 until January
5.3. COLLECTION AND ANALYSIS OF PRICES

Figure 5.4: Number of market price reports by province and species. It is important to note that a single report is defined by species, date and province, and may contain prices for between one and sixty different classifications of animal.

2009 almost no reports were being submitted.

Monetary rates

The variation in exchange rates during the study period is shown in Figure 5.6. In summary:

- The Chinese Yuan Renminbi (RMB) and Lao Kip (LAK) strengthened by about 6% and 9% respectively against the USD;

- the Cambodian Riel (KHR) weakened by about 6% against the USD, and the Vietnamese Dong (VND) weakened by more than 17%;

- the Thailand Baht (THB) weakened by about 20% during the course of the study, but recovered to end about 10% weaker.
Results of analysis for cattle

In order to take into account the variability in animal value between provinces, it was intended to use a formula that included age, sex, condition, month as fixed effects, and province as a random effect.

A data frame\(^\text{15}\) of 5,142 observations of 10 predictor variables for the price of cattle in Cambodia and Laos was extracted from the database. A natural logarithmic transformation of the response variable was determined to be suitable, and this

\(^{15}\text{A form of array or matrix used by R to hold related variables in a logical framework.}\)
5.3. **COLLECTION AND ANALYSIS OF PRICES**

![Graphs showing the variation in the value of regional currencies from January 2008 to February 2010.](image)

**Cambodian Riel vs United States Dollar − 2008-01-01 to 2010-03-01**

- KHR per USD
- 4200 to 3900, −5.9%

**Laos Kip vs United States Dollar − 2008-01-01 to 2010-03-01**

- LAK per USD
- 9200 to 8600, 8.7%

**Thailand Baht vs United States Dollar − 2008-01-01 to 2010-03-01**

- THB per USD
- 36 to 30, −9.7%

**Vietnam Dong vs United States Dollar − 2008-01-01 to 2010-03-01**

- VND per USD
- 18500 to 16500, −17.2%

**Chinese Yuan Renminbi vs United States Dollar − 2008-01-01 to 2010-03-01**

- CNY per USD
- 7.3 to 6.9, 5.8%

Figure 5.6: Plots showing the variation in the value of regional currencies in the period from January 2008 to February 2010. The net change displayed on each graph is the percentage change in the mean price from the first month to the last month of the period (shown by the horizontal red lines).

Value was calculated and added to the data frame. A small number of outliers (high price) were identified using graphical techniques, however they were retained for this analysis. The prices for October 2009 from Ratana Kiri had a mean of USD 1 132, which was USD 826 greater than the mean price from Ratana Kiri in other months. At the recommendation of the Department of Animal Health and Production in Cambodia, these prices were reduced by USD 500. After this adjustment the prices ranged from USD 20 to USD 1 000, with a median price of USD 205. The highest priced animals were mature male cattle in good condition (and may have been bulls). Visual appraisal of multivariable varying intercept/varying slope models
using province and another predictor as random effects showed that most provinces had very similar slopes for each predictor. Where the slope was obviously different, the data for the province was examined. In most cases the differing variable slopes were related to provinces/predictor combinations where there were only few records existed.

The absolute value of correlations between predictors were all $\leq 0.2$ suggesting little or no correlation between predictors. Zuur et al. (2010) suggest that Variance Inflation Factors (VIFs) over 3 (and certainly those greater than 10) are indicative of collinearity. In this case all VIFs were all around 1, so there is little collinearity between the selected predictors.

The final model included as predictors age, sex, condition and month as fixed effects, and province as a random effect. This model had the lowest AIC (1 650). The AIC for the null model was 8 226 and that for the full model with all interactions was 3 304. None of the interactions investigated improved the fit of the model.

The model coefficients and standard errors for the fixed effects are shown in Table 5.5. The coefficients for age and sex were significant and biologically plausible. An animal in medium condition was significantly more likely to be worth more than one in poor condition. However, good condition had a negative coefficient estimate (the standard error for this estimate was quite large and the 95% confidence interval included 0). The month effect was not significant for all months, it was retained it in the model to allow predictions for different months, and it significantly improved the model fit over a model with only quarter (AIC=1 714, $p < 0.05$), or no temporal parameter at all (AIC=1 806, $p < 0.05$). The negative coefficients for the months more closely associated with the wet season were investigated by using season (‘Wet’ or ‘Dry’) instead of month. The model with month was a better fit and so season was not included to avoid issues with collinearity. The variance ratio for the cattle model was 1.0 (intraclass correlation coefficient = 0.51).
Table 5.5: Statistical outputs from a multilevel linear model using the natural logarithm of price as an outcome for price variation for cattle in Cambodia and Laos. (Coef. Est. = Coefficient Estimate, Std. Err. = Standard Error, Conf. Int. = Confidence Interval)

<table>
<thead>
<tr>
<th>Variable</th>
<th>Level</th>
<th>Coef. Est.</th>
<th>Std. Err.</th>
<th>P-value</th>
<th>95% Conf. Int.</th>
</tr>
</thead>
<tbody>
<tr>
<td>(Intercept)</td>
<td></td>
<td>4.44</td>
<td>0.08</td>
<td>&lt;0.01</td>
<td>(4.29, 4.60)</td>
</tr>
<tr>
<td>Age</td>
<td>Calf</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>Weaner</td>
<td>0.33</td>
<td>0.01</td>
<td>&lt;0.01</td>
<td>(0.31, 0.36)</td>
</tr>
<tr>
<td></td>
<td>Juvenile</td>
<td>0.73</td>
<td>0.01</td>
<td>&lt;0.01</td>
<td>(0.71, 0.76)</td>
</tr>
<tr>
<td></td>
<td>Adult</td>
<td>1.12</td>
<td>0.01</td>
<td>&lt;0.01</td>
<td>(1.09, 1.15)</td>
</tr>
<tr>
<td></td>
<td>Mature</td>
<td>1.25</td>
<td>0.01</td>
<td>&lt;0.01</td>
<td>(1.23, 1.28)</td>
</tr>
<tr>
<td>Sex</td>
<td>Female</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>Male</td>
<td>0.05</td>
<td>0.01</td>
<td>&lt;0.01</td>
<td>(0.04, 0.07)</td>
</tr>
<tr>
<td>Condition</td>
<td>Poor</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>Medium</td>
<td>0.25</td>
<td>0.01</td>
<td>&lt;0.01</td>
<td>(0.24, 0.27)</td>
</tr>
<tr>
<td></td>
<td>Good</td>
<td>-0.02</td>
<td>0.01</td>
<td>0.03</td>
<td>(-0.03, 0.00)</td>
</tr>
<tr>
<td>Month</td>
<td>Jan</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>Feb</td>
<td>0.10</td>
<td>0.04</td>
<td>0.01</td>
<td>(0.03, 0.17)</td>
</tr>
<tr>
<td></td>
<td>Mar</td>
<td>0.31</td>
<td>0.04</td>
<td>&lt;0.01</td>
<td>(0.23, 0.39)</td>
</tr>
<tr>
<td></td>
<td>Apr</td>
<td>-0.05</td>
<td>0.03</td>
<td>0.10</td>
<td>(-0.12, 0.01)</td>
</tr>
<tr>
<td></td>
<td>May</td>
<td>-0.05</td>
<td>0.03</td>
<td>0.12</td>
<td>(-0.11, 0.01)</td>
</tr>
<tr>
<td></td>
<td>Jun</td>
<td>-0.09</td>
<td>0.03</td>
<td>0.01</td>
<td>(-0.15, -0.03)</td>
</tr>
<tr>
<td></td>
<td>Jul</td>
<td>-0.02</td>
<td>0.03</td>
<td>0.35</td>
<td>(-0.08, 0.05)</td>
</tr>
<tr>
<td></td>
<td>Aug</td>
<td>-0.05</td>
<td>0.03</td>
<td>0.09</td>
<td>(-0.11, 0.01)</td>
</tr>
<tr>
<td></td>
<td>Sep</td>
<td>-0.01</td>
<td>0.03</td>
<td>0.37</td>
<td>(-0.07, 0.05)</td>
</tr>
<tr>
<td></td>
<td>Oct</td>
<td>-0.06</td>
<td>0.03</td>
<td>0.04</td>
<td>(-0.12, -0.01)</td>
</tr>
<tr>
<td></td>
<td>Nov</td>
<td>-0.02</td>
<td>0.03</td>
<td>0.32</td>
<td>(-0.08, 0.04)</td>
</tr>
<tr>
<td></td>
<td>Dec</td>
<td>0.00</td>
<td>0.03</td>
<td>0.40</td>
<td>(-0.06, 0.06)</td>
</tr>
</tbody>
</table>

Random effects

<table>
<thead>
<tr>
<th>Groups</th>
<th>Name</th>
<th>Variance</th>
<th>Std.Dev.</th>
</tr>
</thead>
<tbody>
<tr>
<td>Province</td>
<td>(Intercept)</td>
<td>0.1136</td>
<td>0.337</td>
</tr>
<tr>
<td>Residual</td>
<td></td>
<td>0.0789</td>
<td>0.281</td>
</tr>
</tbody>
</table>
The scale and direction of the random effects is shown in Figure 5.7.

The range in price variability at the province level is likely to be driven by differences in the supply and demand of cattle in the different provinces. For example, in Figure 5.7 the province with the largest positive random effect (Ratanakiri) is fairly mountainous terrain generally unsuited to cattle production. Other provinces with large positive values include Kampong Som (now all but occupied by the municipality of Sihanoukville), Takeo, Kandal and Kampot (provinces from where significant trade into Vietnam occurs). The provinces displaying large negative random effects in the model are associated with cattle production and have relatively higher numbers of
cattle per person, reflecting increased supply and reduced demand.

Spherical semivariograms (Figure 5.8) suggested little or no spatial correlation in any month beyond immediate neighbours, with the range typically <100km. Anisotropy was seen with directional semivariograms throughout the year\(^\text{16}\). Figure 5.9 shows an example for December however it may be an artefact of the sample size in any given direction being small (Berke, O, personal communication, 2011). Isopleth maps showing the variation in price in March (highest prices) and October (lowest prices) are shown in Figure 5.10.

---

\(^{16}\)The shortening of the 90°, 135° and omnidirectional semivariograms reflects the narrow east-west shape of Cambodia and Laos. The very short range of the north-south directional semivariogram suggests insufficient samples for the distance in this direction.
Figure 5.9: A plot showing the anisotropy associated with the semivariogram for price in December for cattle in Cambodia and Laos.
5.3. COLLECTION AND ANALYSIS OF PRICES

Figure 5.10: Isopleth map from kriging the price of an adult female cattle in medium condition in Cambodia and Laos in October.

(a) Highest prices were in March

(b) Lowest prices in October
Results of analysis for pigs

As for cattle, a data frame of 7,491 observations of 11 predictor variables for the price of pigs in Cambodia and Laos was extracted from the database. A small number (60) of outliers from one province in one month exhibited extremely low prices. Appraisal of these values suggested that they were most likely to have been prices per kg rather than per head, and so they were discarded for this analysis. After excluding these data rows, natural logarithmic transformation of the response variable (price in USD) was determined to be suitable, and this value was calculated and added to the data frame. The data displayed also reasonable homogeneity in the ratio of variance (< 4) for all predictors except the province and country. Variance between provinces in Laos showed the most heterogeneity, whilst the ratio of variances in Cambodia was not as apparent. In general, the variance by country for Cambodia was only 1.9 times greater than that in Laos. As for cattle, the intention was to model the location as a random effect, to allow the determination of the intercept values at each province after accounting for the other parameters.

Again, as with cattle, the absolute value of correlations between predictors were all \( \leq 0.2 \) suggesting little or no correlation, and all VIF’s were in the range 1.0–1.2, so there appears to be little collinearity between these predictors.

The final model included as predictors age, sex, condition, breed and month as fixed effects, and province as a random effect. None of the interactions investigated improved the fit of the model.

The model coefficients and standard errors for the fixed effects are shown in Table 5.6 (on page 154). From this table the coefficients related to age and breed and condition are consistent with expectations—older (bigger) animals in better condition attract a higher price; similarly, a premium was expected for pigs with exotic genetics, and a discount for locally bred animals. Using ANOVA, the difference between the model containing sex (AIC=5 448) and one without (AIC=5 446) was not significant \( (p = 0.72) \) and so it was been retained. As with cattle, the month effect was
significant ($p < 0.05$) so it was retained in the model although some individual months were not. The variance ratio for the pigs was 0.36 (intraclass correlation = 0.28) suggesting that prices are more variable at the individual level than was the case for cattle.

The scale of the random effects in the pig model is shown in Figure 5.11.

As with the cattle, the range in price variability at the province level is likely to be driven by differences in the supply and demand of pigs in the different provinces. In Figure 5.11 the province with the largest positive random effects are important
Table 5.6: Statistical outputs from a multilevel linear model using the natural logarithm of price as an outcome for price variation for pigs in Cambodia and Laos. (Coef. Est. = Coefficient Estimate, Std. Err. = Standard Error, Conf. Int. = Confidence Interval)

<table>
<thead>
<tr>
<th>Variable</th>
<th>Level</th>
<th>Coef. Est.</th>
<th>Std. Err.</th>
<th>P-value</th>
<th>95% Conf. Int.</th>
</tr>
</thead>
<tbody>
<tr>
<td>(Intercept)</td>
<td></td>
<td>3.43</td>
<td>0.08</td>
<td>&lt;0.01</td>
<td>(3.28, 3.58)</td>
</tr>
<tr>
<td>Age</td>
<td>Piglet</td>
<td>Reference</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>Weaner</td>
<td>0.11</td>
<td>0.02</td>
<td>&lt;0.01</td>
<td>(0.08, 0.14)</td>
</tr>
<tr>
<td></td>
<td>Grower</td>
<td>0.67</td>
<td>0.02</td>
<td>&lt;0.01</td>
<td>(0.64, 0.71)</td>
</tr>
<tr>
<td></td>
<td>Adult</td>
<td>1.25</td>
<td>0.02</td>
<td>&lt;0.01</td>
<td>(1.21, 1.28)</td>
</tr>
<tr>
<td></td>
<td>Mature</td>
<td>1.52</td>
<td>0.02</td>
<td>&lt;0.01</td>
<td>(1.48, 1.55)</td>
</tr>
<tr>
<td>Sex</td>
<td>Female</td>
<td>Reference</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>Male</td>
<td>&lt;0.01</td>
<td>0.01</td>
<td>0.39</td>
<td>(-0.02, 0.02)</td>
</tr>
<tr>
<td>Breed</td>
<td>Crossbred</td>
<td>Reference</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>Exotic breed</td>
<td>0.13</td>
<td>0.02</td>
<td>&lt;0.01</td>
<td>(0.10, 0.16)</td>
</tr>
<tr>
<td></td>
<td>Local</td>
<td>-0.43</td>
<td>0.01</td>
<td>&lt;0.01</td>
<td>(-0.45, -0.40)</td>
</tr>
<tr>
<td>Condition</td>
<td>Poor</td>
<td>Reference</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>Good</td>
<td>0.13</td>
<td>0.01</td>
<td>&lt;0.01</td>
<td>(0.12, 0.15)</td>
</tr>
<tr>
<td>Month</td>
<td>Jan</td>
<td>Reference</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>Feb</td>
<td>0.37</td>
<td>0.05</td>
<td>&lt;0.01</td>
<td>(0.27, 0.47)</td>
</tr>
<tr>
<td></td>
<td>Mar</td>
<td>0.06</td>
<td>0.06</td>
<td>0.27</td>
<td>(-0.07, 0.18)</td>
</tr>
<tr>
<td></td>
<td>Apr</td>
<td>0.08</td>
<td>0.04</td>
<td>0.08</td>
<td>(-0.01, 0.17)</td>
</tr>
<tr>
<td></td>
<td>May</td>
<td>0.14</td>
<td>0.04</td>
<td>&lt;0.01</td>
<td>(0.05, 0.22)</td>
</tr>
<tr>
<td></td>
<td>Jun</td>
<td>0.17</td>
<td>0.04</td>
<td>&lt;0.01</td>
<td>(0.09, 0.26)</td>
</tr>
<tr>
<td></td>
<td>Jul</td>
<td>0.04</td>
<td>0.04</td>
<td>0.26</td>
<td>(-0.04, 0.12)</td>
</tr>
<tr>
<td></td>
<td>Aug</td>
<td>0.08</td>
<td>0.04</td>
<td>0.07</td>
<td>(0.00, 0.16)</td>
</tr>
<tr>
<td></td>
<td>Sep</td>
<td>0.04</td>
<td>0.04</td>
<td>0.25</td>
<td>(-0.04, 0.12)</td>
</tr>
<tr>
<td></td>
<td>Oct</td>
<td>0.10</td>
<td>0.04</td>
<td>0.01</td>
<td>(0.02, 0.18)</td>
</tr>
<tr>
<td></td>
<td>Nov</td>
<td>0.11</td>
<td>0.04</td>
<td>0.01</td>
<td>(0.03, 0.19)</td>
</tr>
<tr>
<td></td>
<td>Dec</td>
<td>0.20</td>
<td>0.04</td>
<td>&lt;0.01</td>
<td>(0.12, 0.28)</td>
</tr>
</tbody>
</table>

Random effects

<table>
<thead>
<tr>
<th>Groups</th>
<th>Name</th>
<th>Variance</th>
<th>Std.Dev.</th>
</tr>
</thead>
<tbody>
<tr>
<td>Province</td>
<td>(Intercept)</td>
<td>0.0866</td>
<td>0.294</td>
</tr>
<tr>
<td></td>
<td>Residual</td>
<td>0.2314</td>
<td>0.481</td>
</tr>
</tbody>
</table>
locations for the international movement of pigs (in the case of Sekong, pigs are being imported from Thailand and travelling into Vietnam). The provinces displaying large negative random effects in the model are associated with livestock production and have relatively higher numbers of pigs per person. Comparison of random effects by province for cattle (Figure 5.7) and pigs (Figure 5.11) reveals that provinces which have a large positive random effect for cattle are less likely to have a large random effect (in the same direction) for pigs, and vice versa.

The semivariogram (Figure 5.12) suggested greater spatial correlation than for cattle, with the range extending to 200 kilometres in some months. As with the cattle, Figure 5.13 shows significant evidence of variability when directional variograms were evaluated suggesting increased spatial autocorrelation along the southwest–northeast line. However, the same sample size issues are likely to be present. Iso-pleth maps showing the variation in price in March (highest prices) and October (lowest prices) is shown in Figure 5.14.

Figure 5.12: Semivariograms of price of pigs for each month in Cambodia and Laos.
Figure 5.13: Anisotropy associated with the spherical variogram for price in December for pigs in Cambodia and Laos.

The predicted range of prices for cattle/buffalo is shown in Figure 5.15a and the predicted range of prices for pigs is shown in Figure 5.15b. These graphs identify outlier provinces over the course of a year, and right skewing on the price data, with almost every province showing one or more extreme high prices. The values in these graphs are used in Chapter 6 for assessing the predictive value of price on movement.
5.3. COLLECTION AND ANALYSIS OF PRICES

(a) Highest prices were in March

(b) Lowest prices in October

Figure 5.14: Isopleth map from kriging the price of an adult female pigs in medium condition in Cambodia and Laos in October.
Figure 5.15: The range of predicted prices for each month in each province over 12 months in Cambodia for (a) an adult cow in medium condition, and (b) an adult local breed sow in poor condition. The solid black vertical line represents the median price across all provinces and months, and the grey vertical lines represent two standard deviations from the mean for each province. The circles are outlier data points - more than 1.5 times the distance from the median to the 1st or 3rd quartile.
5.3. COLLECTION AND ANALYSIS OF PRICES

5.3.4 Discussion

In this study the hypothesis was that the difference in commodity price between two locations is a significant driver of trade. The ability to compare prices between different provinces provides an opportunity evaluate the viability of this trade and possibly predict animal movements. Data were collected with the intention of being able to compare the difference in price between any two provinces with the movements between them, based on the work in Chapter 4. The ultimate hope was that rather than going to the expense of collecting data on every animal movement—and with the reality that many movements weren’t going to be reported anyway—the price data would provide guidance as to which provinces had a high demand for livestock, and which ones were low.

A number of related issues also required investigation. Fluctuations in currency will affect the balance of trade. The increase in value in China (Figure 5.16) and Laos and the decline in the southern countries is likely to increase the affordability of cattle in China. Changes in government policy may also affect animal availability, as discussed in Chapter 4.

Generally, the close match between the paper form and the data entry interface meet recommendations for data entry systems (Cameron, 1999; Cameron et al., 2004). There were a number of issues associated with translation, encoding systems and fonts that caused frustration on the part of the system users until they were resolved. The large number of numeric entries was undesirable, but unavoidable for this work.

A number of attempts were made to introduce an SMS based system for data collection. A combination of unreliable power in Phnom Penh, difficulties with remotely administering the pre-paid phone cards and a lack of control over the internet access from the location of the SMS receiver system meant the system spent too much time offline to become a routine part of the data collection activities.

High error rates in comparable systems have already been discussed, such as the
Figure 5.16: Plots showing the change in price of pig meat, beef and mutton in two southern provinces in China from 2001 to 2008. The mean price in 2001 and the mean price in 2008 are shown in black for each commodity. The percentage increase in price from 2001 to 2008 is shown on the right of the graph. Data were generously provided by Chang Cai based on reported data in Chinese yearbooks.

Cambodia Agricultural Market Information System\textsuperscript{17} (CAMIS). Monthly reports for this system (up until the most recently available one for March 2010) would suggest that about 20–25% of SMS messages received are erroneous. As CAMIS has a significantly simpler interface than the one proposed for this study, it is likely that the error rate would have been similar, and thus many potential users may have been turned off by experiencing difficulties. CAMIS also had the advantage that it provided a feedback system (users could request the current price in a given

\textsuperscript{17}http://www.agriculturalmarketinformation.org.kh/en/
5.3. COLLECTION AND ANALYSIS OF PRICES

province) and information was read out daily on radio. Although this would have been possible with the price data collected during this activity and may have improved participation, it was not a core part of the research and neither the time required or cost were able to be met.

A number of aspects of the training could have been done better. Most Cambodians speak Khmer, and many know some English, French, Thai (in the west) and/or Vietnamese (in the East) as a result of a long colonial heritage. However, for training on complex or technical issues, translation into Khmer is essential, even though it slows down delivery. Cultural attitudes towards training and teaching in Cambodia are different to those of many western countries, with an expectation in Cambodia that the teacher will know and instruct, and the student will listen and learn (Chhay and Pearson, 2006). It is possible that the approach of trying to engage the staff during the training by asking for their suggestions as to what data should be collected and how may have sent a message of incompetence rather then encouraged participation.

To compare prices, it is necessary to have enough information about what the price represents to ensure the comparison of like products. There is nothing to be gained from comparing the price of a piglet in Banteay Meanchey with a price per kg liveweight of yellow cattle in Ho Chi Minh City. The issues associated with selecting these predictor variables have been discussed in Section 5.2.5.

The coverage of the price data was not uniform, and in many provinces only one or two classifications of data per species were received in any given monthly report. In others, no data were received for many months (see Figure 5.5 for an illustration of this), and in others almost every conceivable combination of data were regularly reported.

Due to the limited data available for some categories of predictor, it was not appropriate to remove cases as it would tend to lead to biased estimates (or at least greater standard errors), or an inability to make estimates (Gelman and Hill, 2007c). Both Allison (2002) and Hair et al. (2010) suggest the use of modelling based approaches
to impute missing data in these circumstances, such as Expectation-Maximisation (EM) to obtain Maximum Likelihood estimates (Dempster et al., 1977). However, EM assumes that data are missing at random and that they are normally distributed—both unlikely assumptions for this data (Allison, 2002; de Leeuw and Kreft, 1986).

The alternative approach—a multilevel linear model with random effects—offered some important benefits for analysing the price data. This approach was primarily chosen to allow the regression intercepts to vary by province and thus allow the prediction of the price of a single class of animal (for example, adult cow in good condition) between provinces. However, the variable levels of data collection between provinces would have made it impossible to use classical regression within some provinces due to the incomplete nature of the data for various combinations of predictors. The multilevel linear model overcomes that by continuing to draw from the full set of province price data (Gelman and Hill, 2007c). Robinson (1991) suggests that the mean of the sample (in this case the province) may in fact be estimated with greater precision than the overall population. The use of the intraclass correlation coefficient shows that both the province level and the price report level contribute to the variance—there is some clustering based on province.

While the results suggest that the collected data fits and the small standard errors would suggest that the model is robust, the failure to collect information specifically about purpose is likely to affect the value of the analysis. The lack of spatial autocorrelation is a surprise, especially for the cattle, and would suggest that variation in price is not particularly influenced by prices further away than neighbouring provinces. The complexity of the data collection process however would ideally be reduced. This work shows that when a standard combination of classifications was predicted from the range of available data it was possible to create a useful description of the variation in price between provinces, but it was unlikely that this process could be automated. Ongoing collection of price data for a single ‘standard’ animal classification would be much simpler albeit at the risk of no trade in this type of
animal in some provinces leading to gaps.

An alternative to the distance-based approach for analysing spatial autocorrelation uses neighbourhood relationships between areal entities. These relationships can then be assessed for spatial autocorrelation using tests such as Moran’s $I$. Transporting animals over the distances potentially involved (hundreds of kilometres) almost certainly will require the use of trucks, and the costs they incur. By not including this cost (at least relatively speaking by the use of distance) it is almost certainly risking a misspecified model. Adjacency based approaches thus would require substantial manipulation of the spatial weights matrix. In addition, there would be some situations where ‘adjacent’ provinces would have no realistic probability of trade occurring directly due to the presence of lakes, mountain ranges or political borders.

During the closing stages of the research it was also discovered that individual provincial animal health offices in Cambodia provide a monthly return which includes prices for cattle, pigs, chickens and ducks in the province, and for some provinces this is done at the district level. It was possible to obtain some of this price data, and it appears that there is spatial and temporal variation visible in these values. In light of the low cost of obtaining this data, and the possibility of producing automated reports showing changes in price either over time or through the country, it would make sense for future work to investigate the use of this information.
Chapter 6

Assessing the Value of Price as a Predictor of Livestock Movement

“Prediction is very difficult, especially about the future”

Robert Storm Petersen (1882-1949)
6.1 Introduction

The ability to predict the spread of transboundary animal disease is dependent on a clear understanding of the patterns of livestock movement. Data on livestock movement are not always available, may be expensive to obtain or (as in this research) the data may be inconsistent and likely to significantly underestimate the actual number of movements. Because of this, an alternative way of predicting movement is required. In this chapter, the use of the price difference between provinces as a predictor of movement between these provinces was investigated.

As well as the difference in price, a number of other factors may influence the decision to trade livestock between provinces. The cost of transporting animals between provinces is an important factor when evaluating the opportunity to profit by buying animals and selling them in a different province. Although a real cost of transport may not be possible to obtain, the distance between provinces provides a reasonable guidance for the cost of the movement. Demand for livestock varies throughout the year. For example, typically in Cambodia demand for livestock drops during the rainy season when fish are abundant, and increases again during the dry season. The presence of festivals, the start of harvest and disease outbreaks are also likely to impact on the supply and demand of animals, but are variable across countries and the Greater Mekong Subregion (GMS).

The previous chapter used an explanatory regression model to explain the variability in the price of different classes of livestock and then used a defined set of explanatory variables to predict price in each province. The objective of this part of the research was to establish if price data modelled during this project could be used to predict the movement of livestock between provinces. To achieve this the occurrence of animal movement between two provinces was modelled in a logistic regression model against the predicted price difference and distance between the provinces.

In addition, because of the close relationship between cost of transport and distance travelled, then relationship between euclidean (straight-line) and road network dis-
tance was examined to ensure the best measure of cost for using in this model.

6.2 Materials and Methods

6.2.1 Choice of Model Approach

To test the value of price difference as a predictor of livestock movement a logistic regression model was appropriate to better understand the role of price. In this model the outcome variable $\text{move}_i$ was either movement occurring between province a and b, or no movement occurring (note that these are directed movements, $\text{move}_{a\rightarrow b} \neq \text{move}_{b\rightarrow a}$, so each province pair has two movements associated with it). The predictor variables are the difference in price for a standard animal (cattle/buffalo or pig), $\text{d.price}_i$, the estimated cost of the movement $i$ approximated by the euclidean or road distance between the provinces in kilometres, $\text{distance}_i$, and the time of year the movement occurred, $\text{time}_i$.

A model was developed to explain the movement event based on Equation (6.1).

$$
\logit(\text{move}_i) = \beta_0 + \beta_1 \text{d.price}_i + \beta_2 \text{time}_i + \beta_3 \text{distance}_i + \epsilon
$$

$$
\epsilon \sim N(0, \sigma_i^2)
$$

where:

$\text{move}_i$ is a movement occurring from province a to b,

$\beta_0$ is the intercept,

$\beta_1, \ldots, \beta_3$ are the coefficients

$\text{d.price}$ is the difference in predicted price between province a and b,

time is the time of the year the movement occurred,
distance is the distance between province a and b,

\( \epsilon \) is the error term.

This model was subsequently evaluated for fit along with a number of other combinations including a full model with interaction terms. Logistic regression is a subset of generalised linear modelling using a binomial family with a logit link function (Agresti, 2002b; Dohoo et al., 2009a; Gelman and Hill, 2007a). In light of the large number of province pairs where no movement was reported during a month, it may be appropriate to replace the canonical logit link function with the extreme value model using the complementary log-log (cloglog) link function (Agresti, 2002a; Dobson and Barnett, 2008; Zuur et al., 2009). In this work, the use of both functions was tried and the decision on which one to use was made in light of information on the model fit and interpretability.

Only limited data were available on prices and movements from Laos so this analysis was confined to Cambodia. A larger amount of higher quality road network information was available for Cambodia and as the process of developing topologically correct road networks was extremely time consuming, it was not undertaken for Laos.

### 6.2.2 Derivation of the Outcome Variable—Movement

Under the assumptions of making a profit discussed in Chapter 5, it is prudent to test the hypothesis that livestock will move from province a to province b only if there is a reasonable probability that the trader will make a profit.

Summary data on the number of cattle/buffalo and pig movements in Cambodia in 2009 were extracted from the database discussed in Chapter 4 into R. These data were categorised by month, origin and destination. Movements between locations within the same province were excluded from the analysis as the price difference and distance within provinces would always be zero. Months where at least one movement occurred from province a to province b were coded as 1, remaining pair
combinations were coded as 0. Age, sex, condition and breed information were not available for the movement data.

The final outcome variable was a binary variable coding for movement from province a to province b in a given month (0 = no movement, 1 = movement recorded).

6.2.3 Derivation of the Predictor Variables

The predictor variables considered for inclusion in the model were the time of the year, the difference in price between province pairs and the distance between provinces.

Price Difference

To calculate the difference in price between each pair of provinces the price in each province was predicted by month as discussed in Chapter 5. For cattle, the predicted prices for an adult cow in medium condition in each calendar month were used. This combination was chosen as data were available for this combination of categories from the greatest number of provinces over the largest range of months. In pigs the predicted prices for local breed adult sow in poor condition were used on the same premise as for the cattle. Isopleth maps showing the relative prices are shown in Figure 5.10 for cattle/buffalo and Figure 5.14 for pigs.

The predicted price difference was defined as the price in province b (the destination) minus the price in province a (the origin). The values for cattle/buffalo were fitted as a continuous predictor after division by 100 (referred to as price.100 in the cattle/buffalo model), allowing for ready interpretation of the coefficient values and description of the implications of a change in price per USD 100. For pigs the price differences were divided by 10 (referred to as price.10 in the pig model). In a number of cases the price difference between the provinces was negative\(^1\) and thus

---

\(^1\)There were movements in both directions between these provinces—therefore at least one would be moving against the price gradient.
unlikely to relate to market driven trade, so these prices were removed from the analysis. Further aspects of this decision are considered in the discussion (Section 6.4.2).

**Time of Year**

Time of year was modelled as a binary variable representing season (dry vs rainy). Season was chosen as many aspects of the culture are influenced by season, including eating preferences (less beef and pork is eaten during the rainy season when fish is abundant) and ease of travel (flooding affects transport during the rainy season).

**Distance**

Both euclidean distance and road route distance were calculated for inclusion in the initial model. Distance between provinces were measured using either the location of the largest urban area (often the provincial capital) or the geometric centroid for each province.

Euclidean distance (in kilometres) between provinces was calculated using the WGS84 spheroid\(^2\). Road distance was calculated by creating a topologically corrected map by digitising and correcting road data from two sources, Digital Chart of the World\(^3\) and OpenStreetMap\(^4\). Three levels of road quality were in the final map, ranging from highway to dirt road. The cost associated with using a main road was measured in kilometres. As road quality may influence the decision to use a particular road, the cost applied to using other than the main roads was set at 150% of the per kilometre rate for the main road, except for the smaller ‘back-roads’ where the cost was set at 200% of that of the main road. These costs were based roughly on personal experience gained during the research of travelling on similar quality

\(^2\)The WGS84 spheroid is defined as having a transverse radius of 6 378 137 metres, and a flattening ratio of 298.2572235. These values obtained from [http://spatialreference.org/ref/epsg/4326/html/](http://spatialreference.org/ref/epsg/4326/html/).

\(^3\)http://www.maproom.psu.edu/dcw/, no longer available.

\(^4\)http://www.openstreetmap.org/
roads and slower travel speeds and likely increased wear and tear on vehicles. On this basis the Dijkstra routing algorithm\(^5\) used to determine the distance between two provinces did not always return the shortest route, but the most cost-effective route\(^6\).

Using this approach a ‘most probable’ road line between two provinces was created for each province pair using the procedural language code shown in appendix D, and the length of these segments, rounded to the nearest kilometre, was then used. The road lines are shown in Figure 6.2. During the model development process distance was fitted both as a continuous and a categorical predictor. A number of permutations of road lengths were used, including (making the coefficient simpler to interpret), distance in hundreds of kilometres centred on the mean, and a categorical description of the road distances. For cattle, the categorical values were \(<200\text{km}, <400\text{km}, <600\text{km} \text{ and } \geq 600\text{km}\), based on the quartiles of the movements. For pigs, due to lesser numbers of movements over the longer distances, the categorical values were \(<200\text{km}, <400\text{km} \text{ and } \geq 400\text{km}\).

The variation between the euclidean distance and the road network distance between province pairs was investigated by examination of scatter plots and linear regression. To avoid violation of the independence assumption the testing was conducted using the distances in only one direction by taking the lower triangle of the matrix. The \(R^2\) was calculated to determine the proportion of the variance in distance along a road that was explained by euclidean distance. Province pairs where the standardised residuals were greater than 4 when checking the linear regression model fit were also plotted to allow their identification and further investigation. Both the euclidean and road distances were tried in the logistic model to determine which had the best fit.

\(^{5}\)http://en.wikipedia.org/wiki/Dijkstra’s_algorithm\(^{6}\)This is the same process that most GPS car navigation computers are using when they choose to travel further but faster on a motorway instead of crawling through suburban streets.
6.2. MATERIALS AND METHODS

6.2.4 Model Building and Evaluation

Building on the initial model described in Equation (6.1), a full logistic regression model including the interaction terms \( \text{d.price} \times \text{distance} \) and \( \text{d.price} \times \text{time} \) were developed. Separate models were built for cattle and for pigs. A number of combinations of predictors were attempted including distance scaled by 100, distance centred on 0 and distance as a categorical variable. Backwards and forwards stepwise algorithm (as described by Dohoo et al., 2009c, p 386) was then used to evaluate different possible model combinations taking into account biological plausibility. For these models Akaike's Information Criteria (AIC) was calculated for each model using the ANOVA function in R to identify the model with best fit, judged by the lowest AIC and significant p value. The same approach used for the logistic regression model was then applied but using the cloglog link function, and resulted in the same predictors being chosen, but as it provided no extra value and was more difficult to interpret it was discarded.

A number of authors concur that checking logistic regression models for fit is difficult (Agresti, 1990; Dohoo et al., 2009a; Gelman and Hill, 2007a; Zuur et al., 2009). Hosmer and Lemeshow (2000) recommend a number of different approaches due to the limits of tests (such as their own Hosmer-Lemeshow Test) that return a single value which may disguise variability in the data. These models were evaluated using a range of techniques. Linearity in the logit for continuous predictor variables (price difference, and for distance when fitted as a continuous variable) was assessed by plotting coefficients against medians of categories of the predictor variable for the final multivariable model (Perkins, 2011). A combination of goodness of fit tests and visualisation of standardised residuals and leverage as described in Dohoo et al. (2009a) were undertaken, as well as the goodness of fit test described by Cessie and Houwelingen (1991). Binned residual plots were produced and checked to ensure that the majority of residuals were within two standard deviations of the expected value for the bin (Gelman and Hill, 2007a). An option not used in this study is evaluation of Receiver Operating Characteristic (ROC) curves to better assess the
ability of the model to discriminate between movements that occurred, and those that did not (Dohoo et al., 2009a; Hosmer and Lemeshow, 2000).

6.3 Results

6.3.1 Data Overview

A total of 3 317 movements of cattle and buffalo were recorded wholly within Cambodia in 2009. Cattle/buffalo were recorded moving from 21 of 24 provinces in Cambodia and moving to 22 of 24 provinces. No movements were recorded from Kep\textsuperscript{7}, Svay Rieng or Kandal. No movements were recorded to Pailin\textsuperscript{8} or Mondul Kiri.

For pigs, 2 199 movements were recorded, originating from 12 and terminating in 12 provinces as shown in Table 6.1.

Table 6.1: Provinces from which movements of pigs in Cambodia in 2009 originated and ended. Note that there is no direct relationship implied by the rows, animals from any province may have travelled to any other province in the list (they are alphabetically ordered).

<table>
<thead>
<tr>
<th>Origin</th>
<th>Destination</th>
</tr>
</thead>
<tbody>
<tr>
<td>Banteay Meanchey</td>
<td>Battambang</td>
</tr>
<tr>
<td>Battambang</td>
<td>Kampong Cham</td>
</tr>
<tr>
<td>Kampong Cham</td>
<td>Kampong Chhnang</td>
</tr>
<tr>
<td>Kampong Speu</td>
<td>Kampong Som</td>
</tr>
<tr>
<td>Kampong Thom</td>
<td>Kampong Thom</td>
</tr>
<tr>
<td>Kampot</td>
<td>Kandal</td>
</tr>
<tr>
<td>Kandal</td>
<td>Phnom Penh</td>
</tr>
<tr>
<td>Kracheh</td>
<td>Preah Vihear</td>
</tr>
<tr>
<td>Otdar Meanchey</td>
<td>Pursat</td>
</tr>
<tr>
<td>Prey Veng</td>
<td>Ratanak Kiri</td>
</tr>
<tr>
<td>Svay Rieng</td>
<td>Siem Reap</td>
</tr>
<tr>
<td>Takeo</td>
<td>Stung Treng</td>
</tr>
</tbody>
</table>

\textsuperscript{7}Kep is one of the small provinces (Khet) formed by a Royal Decree on 22nd December 2008 (Sokheng, 2008). It was previously a municipal area (Krong) and it is likely that movements from here are recorded under Kampt.

\textsuperscript{8}Pailin is also one of the small provinces formed from a municipal area, in this case within Battambang province.
6.3. RESULTS

6.3.2 Comparison of Euclidean and Road Distances

A scatter plot showing association between the euclidean distance and the road network distance in Cambodia is shown in Figure 6.1. Movements could occur in either direction, and due to the length of the road segments being used and the calculation of the road closest to the centre of the province, in many cases the distance between two provinces varied depending on the direction of travel. The proportion of variance in distance along the road network that was explained by the euclidean distance was high ($R^2 = 0.93$). The intercept was 1.07 and the coefficient was 1.62 (95% confidence interval 1.58–1.65).

![Figure 6.1: Scatter graph showing association between the euclidean distance and the road network distance between province centres in Cambodia. Outlier pairs are shown with red circles, and are identified with the name of the province pairs. These provinces are labelled in Figure 6.2.](image)

Examination of the plot of fitted values versus the square root of standardised residuals showed three prominent outliers (shown with red circles in Figure 6.1). These were investigated and found to represent roads between Koh Kong and Pailin, Koh
Kong and Banteay Meanchey and Koh Kong and Battambang. The road network used to calculate the distances between provinces is shown in Figure 6.2. The provinces associated with the outlier points are labelled to help illustrate why they have an unusually high road distance compared to the euclidean distance—much of the area between Koh Kong province and Battambang province is filled by the largely inaccessible Cardamom Mountains.

Figure 6.2: A map of Cambodian provinces with the ‘least cost’ path road network between provinces used in the calculation of distances shown in yellow. The labelled provinces are those provinces where the road network is longer than would be estimated by the straight-line distance as shown in Figure 6.1.
6.3.3 Price Differences

The prices in each province for each month were predicted using the best fit multi-level model determined in Chapter 5. The range of prices for cattle/buffalo is shown in Figure 5.15a and the range or prices for pigs is shown in Figure 5.15b.

6.3.4 Model Results

All provinces from which movements had originated or terminated were used in the analysis, even when no movements were recorded directly between two provinces. Although models with road or euclidean distance were significant, using road distance provided a better model than euclidean distance, and so euclidean distance was dropped.

There was little collinearity between the predictors in the final model (road distance, price and season) and the variance inflation factors were all less than 1.5 for both cattle and pigs.

Season did not appear to be a significant contributor to the model for pigs and so it was dropped.

Partial residuals plots demonstrated linearity for the price and continuous distance variables but provided limited insight into the categorical distance predictors. Examination of standardised residuals and leverage, as in Dohoo et al. (2009a), and the use of binned residual plots (Gelman and Hill, 2007a) did not reveal any extreme values or unusual patterns. Goodness of fit tests were not significant at the $p < 0.05$ level.

The final model for cattle is shown in Equation (6.2) and for pigs in Equation (6.3).
\[
\text{logit}[P(Y = 1)] = \alpha + \beta_1\text{price.100} + \beta_2\text{c.dist.100} \\
+ \beta_3\text{price.100} \times \text{c.dist.100} + \beta_4\text{season} + \epsilon
\] (6.2)

\[
\epsilon \sim N(0, \sigma_i^2)
\]

\[
i = 1, \ldots, n \text{ movements}
\]

where:

\(P(Y = 1)\) is the probability of a movement between two provinces during a season,

\(\text{price.100}\) is the difference in price between two provinces in 100’s of USD,

\(\text{c.dist.100}\) is the road network distance between provinces in 100’s of kilometres centred on the mean distance,

\(\text{season}\) is the season in which the movement is occurring (Dry or Rainy).

\[
\text{logit}[P(Y = 1)] = \alpha + \beta_1\text{price.100} + \beta_2\text{c.dist.100} \\
+ \beta_3\text{price.100} \times \text{c.dist.100} + \epsilon
\] (6.3)

\[
\epsilon \sim N(0, \sigma_i^2)
\]

\[
i = 1, \ldots, n \text{ movements}
\]

where:

\(P(Y = 1)\) is the probability of a movement between two provinces,

\(\text{price.10}\) is the difference in price between two provinces in 10’s of USD,

\(\text{dist.100}\) is the road network distance between provinces in 100’s of kilometres.

Statistical output from the logistic regression models for cattle/buffalo and pigs are shown in Tables 6.2 and 6.3.

The role of increasing price, season and distance are shown in Figure 6.3a for cattle.
Table 6.2: Statistical outputs from a logistic regression model using the occurrence of movement between two provinces as an outcome for cattle in Cambodia. (Coef. Est. = Coefficient Estimate, Std. Err. = Standard Error, Conf. Int. = Confidence Interval)

<table>
<thead>
<tr>
<th>Variable</th>
<th>Level</th>
<th>Coef. Est.</th>
<th>Std. Err.</th>
<th>P-value</th>
<th>95% Conf. Int.</th>
</tr>
</thead>
<tbody>
<tr>
<td>(Intercept)</td>
<td></td>
<td>-3.70</td>
<td>0.29</td>
<td>&lt;0.01</td>
<td>(-4.29, -3.17)</td>
</tr>
<tr>
<td>Price</td>
<td>per $100</td>
<td>0.33</td>
<td>0.11</td>
<td>&lt;0.01</td>
<td>(0.12, 0.54)</td>
</tr>
<tr>
<td>Distance</td>
<td>per 100km</td>
<td>-0.81</td>
<td>0.13</td>
<td>&lt;0.01</td>
<td>(-1.07, -0.56)</td>
</tr>
<tr>
<td>Interaction</td>
<td>price.100:dist.100</td>
<td>0.19</td>
<td>0.5</td>
<td>&lt;0.01</td>
<td>(0.08, 0.29)</td>
</tr>
<tr>
<td>Season</td>
<td>Dry</td>
<td>Reference</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>Rainy</td>
<td>-0.67</td>
<td>0.21</td>
<td>&lt;0.01</td>
<td>(-1.09, -0.26)</td>
</tr>
</tbody>
</table>

and Figure 6.3b for pigs. From these graphs it can be seen that the probability of movements occurring rises as the price difference increases.

The distance of the move appears to have a slightly less predictable effect. For cattle, up to about USD 425, the price difference seems to have the greatest effect on the probability of a movement occurring over shorter distances, and less of an impact on longer distance movements. Beyond this price, the probability of longer distance movements increases more rapidly than the shorter distances. This may simply reflect the limits associated with prediction near the ends of the range of data, or it may be an effect of some data on the international cattle trade being mixed with the domestic market.

For pigs, there is a pronounced impact of price difference for moves over a shorter distance, with a price difference of about USD 45 over 100 kms having a probability of about 0.5 of a movement occurring. In contrast to the cattle model, the price difference has very little impact on the probability of a move occurring for move-
Figure 6.3: Plots showing the relationship between the probability of a movement of (a) cattle and (b) pigs occurring between two provinces in Cambodia and the price difference, road distance between provinces and season of travel (for cattle/buffalo).
ments over longer distances. The slightly higher probability associated with price differences up to USD 5 may be an artefact of working near the end of the range of data.

The impact of season on movement of cattle is significant and is reflected in the higher probability of a movement during the dry season when price and distance are held the same.

6.4 Discussion

6.4.1 Comparison of Euclidean and Road Network Distances

This research provides an example of the use of the road network distance to account for variability between points. The road distance was on average approximately 60% greater than the euclidean distance between the same province pair. In the process of fitting the model it was shown that using road distance provided an improvement in the fit of the model compared to the use of euclidean distance. This is an important outcome as although the potential benefit of using road distance rather than straight line distances is recognised (see Section 2.5.1), the implementation of algorithms to identify most likely road routes has not previously been reported in veterinary epidemiology literature.

Although in the context of spread of disease between provinces (rather than individual premises) there is some loss of precision, an important finding is that in Cambodia (a mainly flat country) there is a fundamentally linear relationship between euclidean distance and road network distance. Both of these are plausible due to the intrusion of mountain ranges, a lake and the historical presence of Khmer Rouge rebels limiting road building activities creating the need for quite significant diversions when travelling by road. The impact of mountains or lakes is consistent with the expectation described in Bessell et al. (2008) and Savill et al. (2006). Although neither of these papers mention the impact of guerilla activities, this is likely to be
an important issue in a large number of countries where civil strife is interfering with disease control programmes.

### 6.4.2 Prediction of Livestock Movement Using Price

Price is able to be used as a predictor of movements. Given the importance of realising a profit as an underlying mechanism driving trade, at first the small increase in the odds of a movement per unit of price difference (USD 100 for cattle/buffalo and USD 10 for pigs) would appear to raise questions about the analytical approach and model construction. There are however, a number of reasons why in this situation the predictive ability is so low. To a large extent they are postulated to be related to the difference between:

- The class of animals the price data were collected for, and
- the choice of class of animal for prediction of prices, and
- the class of animals that were actually being moved.

One of the most difficult aspects of the price data collection (as discussed in Section 5.2) was the ability to generate a robust description of the animal to which a given price applied. To overcome this a number of categorical variables (arguably not enough) were captured, along with each price. This was shown to allow a reasonably good fit of the data to most of these variable and allow the prediction of price for predetermined classes of cattle and pigs. However this dataset did not contain prices for every combination of the categorical variables from every location. A prediction approach was therefore used to generate price data in the face of the constraints with data collection and the predicted prices was used in the logistic model instead of the observed prices. The decision to choose an adult cow in medium condition and an adult local breed sow in poor condition as the reference animals was based on observation of a commonly sold animal in markets. It would be possible to repeat the prediction investigations with alternative classes of livestock.
The mechanism for data collection relating to animal movements was inherently less complete. The process of utilising cheap and readily available data from livestock movement permits issued at the province level meant the description of the animal was often limited to simply species, with no indication of sex, breed, age or condition (a form of measurement error). Although some data about the purpose for which the animal was intended may have been available, it was not in a categorical format. Issues associated with translation made it difficult to collect and potentially code this information where non-English speaking persons were primarily responsible for the data entry.

Although this work has focussed on movements between provinces, a large amount of livestock move within provinces on a regular basis, for sale, draft power, breeding and general transport. This would have significant implications for the spread of disease, as once a foci of infection was introduced, it would rapidly spread to cover the province, and from there into neighbouring provinces. The assumption of homogenous mixing discussion in Section 4.4.1 is more likely in the GMS than in Western Australia, and likely to be the most pragmatic option for modelling the risk of disease spread.

It is unlikely that the missing price data is ignorable, but most provinces submitted some data during the course of the research which would reduce the degree of misclassification bias. Use of summary data values derived from multiple imputation is likely to reduce but not eliminate the scale of bias, but the direction of any pre-existing biases will still not be known (Allison, 2002; Dohoo et al., 2009c). This potential variability would be expected to have an impact on the role of price in assessing the probability of a move, but it is not possible to tell whether it might increase or decrease the probability.

The lack of spatial auto-correlation associated with price was discussed in Section 5.3.4. This is in contrast to the general impression of movements travelling over several provinces at once shown in Figure 4.14. It is quite possible that the movement of animals over longer distances are based on supply to different markets to those
represented by the provincial price data. An example of this might be the supply of livestock from Thailand into the Vietnamese market. The traders involved in this may well have obtained movement permits to transport the animals from Western Cambodia to the south east in preparation for export to Vietnam. Because the purchasing and selling of these livestock is occurring in markets not represented by the price data collected in Cambodia any further inferences amount to little more than speculation.
“...no worries, I’m sure I can get this done in two years...”

(Ben Madin, 2007)
7.1 Summary of Findings

7.1.1 Objectives of the Research

The productivity of livestock owned by smallholders in Cambodia and Laos suffers repeatedly from the introduction of transboundary animal disease. This research was motivated by the opportunity to better predict the risk of disease spread in the Greater Mekong Subregion (GMS) by investigating a number of putative factors that may lead to the introduction of an infected animal into a naïve population. In order to achieve this, a number of studies were undertaken.

First, could disease occurrence be reliably detected rapidly enough to be used as an input into an early warning system. Evaluation of information on outbreaks of disease, in this case foot and mouth disease (FMD), at a regional level was undertaken to determine if any relationship existed between outbreak locations over time and whether it would be possible to identify disease outbreaks early enough that they could be used in the prediction of disease spread.

Second, could animal movements be recorded and analysed to identify the movement patterns and use them to identify at-risk areas for the introduction of disease as a result of an outbreak in another part of the region. Information on the movement of livestock through Cambodia and Laos was obtained both retrospectively and prospectively using a number of different approaches to investigate patterns of movement.

Finally, could the use of an inexpensive proxy measure such as animal price obviate the need for an ongoing investment in the relatively more expensive process of collecting animal movement data in near real-time. Data on price were obtained at a provincial level in Cambodia and Laos. These price data were compared with the livestock movement data to identify whether or not animal movement could be predicted.

Using parameters derived from the above work, a model was developed to evaluate
the association of livestock movement with price difference between provinces. In
the event that the answers to each of these questions were positive, knowing price
at a provincial level could allow the prediction of animal movements throughout the
region. Combined with timely knowledge of disease occurrences, this could form
the basis of risk-reduction activities such as targeted movement controls, enhanced
publicity about disease risks and biosecurity, and if required pre-emptive vaccina-


development.

7.1.2 Disease Reporting

For ten years the South-East Asia and China foot and mouth disease (SEACFMD)
programme has been collecting outbreak data for foot and mouth disease from coun-
tries in South-East Asia. There have been significant improvements in the quality
of data being captured during this time; however, this analysis demonstrated that
the sensitivity and specificity of the current disease reporting system offer room for
improvement.

Using various spatial techniques it was possible to identify provinces where disease
outbreaks were occurring more regularly than expected. Due to questions about the
nature of the disease being reported—including the potential for both missing data
and false positive reports—it was important not to over represent the importance
of the identified provinces in the natural history of the disease in the region.

The establishment of realistic population parameters has allowed the derivation
of standardised morbidity ratios to improve the interpretation of outbreak reports.
Traditionally, this has identified disease hotspots where control programme activities
should be based. A potentially more useful outcome of this analysis is the use of an
empirical Bayes estimator (Clayton and Kaldor, 1987) to identify provinces which
are likely to be underreporting the incidence of disease. This offers support to
an alternative paradigm for disease control programmes—where efforts should be
directed at the provinces that are highly likely to be underreporting the disease

incidence, rather than those which already appear competent.

An online interface was developed to aid in the early detection of increases in disease reporting. A number of different cluster/anomaly detection algorithms are available through this system. These algorithms were not formally evaluated during this research, and it will be important for them to be properly parameterised and assessed for their sensitivity and specificity.

The ability to use molecular data for epidemiological purposes has improved greatly in the last few years. Even as most people struggle to deal with FMD being a general term for three viruses in the region that behave independently, the use of molecular techniques is showing that understanding the topotypes involved in disease outbreaks is vital. Although the current reliance on the VP2 segment of the genome somewhat limits the resolution of these techniques, the falling cost of genetic sequencing means that the use of full genome sequences will improve this situation.

\subsection{Movement}

\textbf{National Livestock Identification System}

A comprehensive assessment of the information on individual cattle movements in Western Australia (WA) stored in the National Livestock Identification System (NLIS) has provided important insights into the number of movements, distance travelled and seasonal nature of these movements.

A number of studies have identified that using euclidean distance\(^1\) is a poor alternative to knowing the actual distance travelled by road, although to date there is little published material using road distance. By being able to apply routing algorithms to road network data sets it is now possible to incorporate information on the distance (and by inference, cost) associated with the movement of livestock. This work showed that in mostly flat, open countryside road distance is \(\approx 60\%
\[^1\text{The euclidean distance is the shortest distance between two points.}\]
greater than the straight line distance, which may have the potential to significantly affect the interpretation of studies using the euclidean distance. The correlation is likely to decrease where more mountainous terrain and other topographic features exist.

The potential use of NLIS data has previously been discussed. The application of network analysis techniques has identified of premise types which are important in terms of disease surveillance and control. It has also identified shires in WA that will be more or less important for disease spread, identification and control. The methods used in this approach were applied at a shire level due to concerns over privacy, but they could be applied at the property level and provide a greater depth of understanding on the spread of disease in the state.

An analysis of predicted spread of disease was also undertaken by bootstrap sampling the existing data to investigate the potential spread of disease and implications of delay in implementing movement controls. Although this analysis was somewhat limited by the assumption of total mixing within shires, it provided a valid approach for modelling the implications of delays and should be directly transferable to property level modelling, where the assumption of complete mixing is more likely to be satisfied.

Greater Mekong Subregion

In the GMS, in contrast to WA, information on the movement of individual animals is scarce. However for longer distance movements, Cambodia and Laos both require the issue of official movement permits. These permits are issued on a consignment basis, and consignment may include a mix of livestock species. There is also a relatively ambivalent attitude towards the need to obtain these permits, in contrast to the strictly enforced requirement to report movements in WA.

The important movements from a transboundary disease control perspective are, of course, transboundary movements. This research was targeted at Cambodia and
Laos, and during the course of data collection no direct movements were recorded between these two countries. This is probably related to the fact that the adjoining districts are not heavily populated, have limited roads through mountainous terrain and the large number of international movements occurring in each country are a result of larger scale movements between Thailand and Vietnam. This makes it difficult to extrapolate the results received within each country to the broader region.

Nonetheless, it was possible to use the same network analysis techniques that were used in WA, and at a relatively similar spatial resolution, although the temporal resolution was not as high. Again, these techniques identify critical provinces for the surveillance and control of disease, and provide guidance as to the direction of flow of animals.

Prospective Animal Movements

The small-scale trial of animal tracking failed to provide any objective insight into the movement of animals. After 24 months, very few of the 1 500 tagged animals had actually left the village in which they were originally tagged, which in itself is useful information on the population dynamics in small villages.

Small scale tagging studies such as this one suffer from a number of limitations:

- Trial sized sample sizes are not large enough to incorporate a significant number of movements;
- even when larger sample sizes are used, the number of potential destinations for any animal is large, and in the current environment the provision of an appropriate number of detection locations would be not be feasible, thus reducing the sensitivity of the system;
- the low adoption of animal tagging in the region leads to a perception that tags are unusual and leads to the removal of tags by owners and traders who do not want the animal to stand out.
The very small number of animals that had been sold is an interesting finding, and investigation of village level population dynamics should be considered for future research.

While an exciting concept, the ability to use a ‘GSM rumen bolus’ (discussed in Appendix B) to track animals that are not moving through legal pathways failed due to technical barriers with transmitting signal through the animal. Even if it had been technically feasible to track cattle and, by inference, humans involved in illegal activities this may have led to prosecution and sanction of the people involved in the trade. This raised ethical concerns about the implications it may have had for the traders, which may have limited the ability to obtain human research ethics approvals to demonstrate its usefulness.

7.1.4 Prices

The capture of livestock price information requires the development of an approach to categorising animals that may be applied by non-specialists yet is adequately able to describe the animal for which the price applies. It was very difficult to achieve consensus on how this should be done. Reference to both the Australian National Livestock Reporting Service and other regional agricultural market information systems suggested that complex systems are expensive and simple systems provide little value. In the absence of objective animal weights, categorical indicators of animal characteristics were used to classify animals when prices were collected. The decision to record species, age, sex, condition and breed created a complex array of data requiring complex analytical techniques. The results of the expert opinion suggested the reason for sale (or subsequent purpose for the animal) and/or animal weight would offer considerable value to the data set, albeit at the risk of increased complexity.

The utility of price as a predictor of movement on a regular basis is limited by the cost of collecting such data. Even with considerable investment of energy and
7.1. SUMMARY OF FINDINGS

funds, during this research it was not possible to consistently collect high quality price data from each province. This level of expenditure is not sustainable under current circumstances, hence a number of alternative sources of data were identified, including research funded market information systems (for example the Cambodian Agricultural Market Information System), government reporting systems (the incorporation of monthly reports from provinces into a national monthly report) and industry driven systems (such as the Australian National Livestock Reporting System). Only the latter system collected data at an adequate resolution to improve the quality of price data over that which was collected as a part of this project. The cost of this would be unsustainable in Cambodia or Laos under current conditions. The methods associated with the collection of data were also problematic, with paper forms requiring data entry for analysis being associated with transcription errors, web-based submission systems requiring internet access, local database systems requiring collection, computer virus cleaning and merging and the use of a more advanced alternative such as text messaging being highly susceptible to power failures, lack of internet bandwidth and limited familiarity.

Even with these challenges, it was still possible to collect a substantial amount of price data covering much of Cambodia and Laos over a 12 month period. It is not possible to conclude that the missing data was missing completely at random (Gelman and Hill, 2007c), however it is possible to use the available data in a multilevel model to evaluate the fit of that data for further prediction of values at points where no data exists. In this case, the use of multilevel modelling techniques provides a robust basis for predicting missing data and thus being able to create maps of the variation in price of different classes of livestock.

An unexpected outcome of this was the apparent lack of spatial autocorrelation beyond the range of neighbouring provinces. This suggests that the prices being paid in any given location are not particularly affected by those prices being paid only a few hundred kilometres away. Whether this represents a failure of market systems or limitations associated with the imputation process is beyond the constraints of
this study. The spatial units of interest (the province) are few and thus relatively large across the scale of the region—had it been possible to sample prices at a higher resolution (for example the district level) it is possible that a greater degree of spatial autocorrelation may have been observed. In a few cases where adequate raw data were available to conduct these analyses without requiring prediction, the autocorrelation was also very limited. Using an adjacency based approach instead of a distance-weighted approach was not used in this work as the presumption was that animals being moved over scales of hundreds of kilometres would be transported by road, and there are a number of examples in the region where neighbouring provinces have no direct transport links (other than walking) between them.

7.1.5 Price as a Predictor of Movement

The use of a regression model contributes to the understanding of the role of price, distance and season in predicting movements. Price can be used to predict the odds of a movement occurring between two provinces but it may only have a limited significance in light of the wide confidence intervals associated with the results. The difference in price between provinces was a weak predictor of the probability of livestock movement. It was most important for movements of less than 200 km, but for movements across the entire country, price difference appears to have almost no impact on the probability of a movement occurring.

During the period of this research there was a steady trade in cattle from Thailand to Vietnam through Cambodia and Laos. This country to country trade is a structured operation managed by a number of larger organisations and appears to be driven by demand from Vietnam with little or no involvement in province to province trade. A plausible explanation is that the local price data collected during this research reflects the supply and demand within a few hundred kilometres, but does not accurately reflect the true prices driving the international trade.
7.2 Implications and Future Directions

This research provides important insights into key aspects of transboundary animal disease management, particularly where information is lacking. A number of opportunities exist to overcome these limitations.

7.2.1 Tools

Advances in Technology

This work commenced prior to the widespread availability of the internet and smartphone technologies in Cambodia and Laos. As a result, the concept of timely capture of data was measured in weeks and months, not days or minutes. These technologies are now becoming commonplace even in increasingly remote areas. The opportunity to use mobile telephone text messaging to overcome some of this disadvantage was implemented as a component of this research. At the time, the underlying technology was still not robust enough, and the coding information for text messaging using cards with encoding keys was cumbersome (and at times frustrating) for users unfamiliar with text messaging. At the same time as this was being tried, a number of organisations in Cambodia and Laos were trying to do similar things. The Innovative Support To Emergencies, Disease and Disasters (InSTEDD) project was one of these, working on developing mobile phone applications\(^2\). During 2010 they announced\(^3\) that they were trialling a ‘coding wheel’. This simple device provides an intuitive method for users to select the reporting options they want and be able to read off the required code from the other side of the wheel, which could then be submitted by text message or an automated phone dialling system.

One of the most limiting aspects of the analysis of FMD in the region was the poor resolution associated with the spatial and temporal location of the reports that were

\(^2\)These were applications for non-smartphone mobile phones, not ‘apps’ known in the iOS/Google phone systems.

received\textsuperscript{4}. The area affected by FMD might best be specified at the village or district level, and analysed using appropriate techniques. However, due to redistribution of administrative boundaries it might not be possible to identify these locations for subsequent analyses. Having an accurately located outbreak point independent of these boundaries is vital to overcome this problem. In the last 15 years, one of the important advances in spatial analysis has been the ability to access the Global Positioning System (GPS) from ever-cheaper handheld devices (Clarke et al., 1996). Increasing availability of these devices to staff involved in field investigations will undoubtedly help to improve the quality of disease reporting.

More efficient and effective data collection is likely to be possible in the future due to the wider reach of mobile telephone networks and their associated data capacity. Improvements in portable technology such as netbooks, smart phones, and other electronic recording devices mean that they are becoming affordable even for smaller projects. Within the last few years, there has been a revolution in the ability to access the internet through the consumerisation of smartphones. The combination of portability, internet access and GPS location into one device has proved fertile ground for developing location based information and reporting systems, such as EpiCollect (Aanensen et al., 2009).

Increased Adoption of Sociological Methodologies

In parallel with this research, other work using sociological techniques has been ongoing in the region (Bourgeois Lüthi, 2011; Cocks et al., 2009; Van Kerkhove et al., 2009). This work has focussed on survey and interview techniques with livestock traders and has demonstrated an ability to rapidly collect information on the nature of the trade and the drivers of the trade for different individuals. Such work is expensive to undertake, relies very heavily on trained personnel with appropriate language and cultural skills and has some limitations in terms of ethical responsibilities to individuals involved in the research. There now exist a number of

\textsuperscript{4}In some cases the information that was provided was likely to be accurate (for example the name of the village), but difficult to accurately locate on a map.
suitably experienced practitioners who can apply this approach, and it seems likely that it can continue to provide vital information for disease control planning.

7.2.2 Disease Reporting

The opportunity to improve the sensitivity of the reporting system for FMD is enhanced by the presence of a systematic approach to identifying the provinces which are likely to be underreporting disease outbreaks. The analyses which identify these provinces are complex, but are adequately robust that they can be included in an automated analysis system. Better insights into the system sensitivity could also be achieved by the application of new but proven methodologies (Martin et al., 2007; Martin, 2008) to utilise negative disease reports, the facility for which does not really exist in the region currently.

Improvement of specificity will also be vital and encouraging the submission of samples to reference laboratories should remain a priority to ensure the best possible information on the disease is obtained; the use of pen-side tests should also be encouraged for basic differentiation between FMD and other diseases.

Utilisation of genetic tools for routine disease investigations has enormous implications for improving the specificity and thus the understanding of the epidemiology of disease in the GMS. The use of full genome sequencing has already begun in Malaysia (King, 2010) and will allow improved differentiation between the virus topotypes causing regular outbreaks, and those topotypes involved in the more explosive outbreaks discussed in Section 2.2.4.

The introduction of the new technologies discussed in Section 7.2.1 offers unprecedented opportunities to improve the spatial and temporal resolution of the disease reporting. Simple widespread adoption of GPS, coupled with the increasing availability of Geographical Information Systems (GIS) in the region (as described by Tum et al., 2004)—and perhaps more importantly, staff capable in their use—should make it feasible to develop far more accurate analyses than are currently possible. A
combination of reduced disease reporting times associated with faster options for the transmission of data than reliance on taxis and buses; a reduced need for pre-analysis data manipulation (for example Section 3.2.2); coupled with viable and automatic cluster and anomaly detection systems (such as those described in Section 3.4) will support the timely implementation of highly efficient, targeted interventions.

7.2.3 Movement

From the perspective of animal disease control, the most important movements are the transboundary movements, not the intra-provincial ones. To overcome the limitations experienced with both the movement permit analysis and the animal tagging study, it is vital to ensure that future work is targeted at the appropriate scale. This has a number of potentially serious cost and political implications, however some opportunities already exist.

An alternative to tracing a cohort prospectively is to work backwards—trace the animals from their destination back to their origin. As the animals are probably not tagged, a paper trail is unlikely and the current owner may be reluctant to provide this information, one approach relies on using geological or biological markers carried on or in the animal. An example of using such markers is ‘soil fingerprinting’—comparing soil samples taken from an animal with the known distribution of similar soils using x-ray fluorescence\(^5\). An alternative approach is using molecular techniques similar to those used for viruses, described under disease reporting (Section 3.2.3). In this situation, genetic variance between a living organism is compared—some examples might include actual animal genetics, a (possibly non-pathogenic) virus carried by the animal, or even plant material collected from faecal samples.

A better alternative in the long term is to support the regional adoption of an animal identification system that would replace the dependency on indirect measures of animal movement with capture of animal movement data. Although support-

\(^5\)A project trialling this work has been funded by ACIAR and will be starting in Cambodia during 2012.
7.2. IMPLICATIONS AND FUTURE DIRECTIONS

In concept (OIE South East Asia Regional Coordination Unit, 2011; Stratton, 2010), under current conditions it is unlikely that the costs of implementing such a system would be supported by producers on the grounds of livestock disease management\(^6\). It is more likely that trade-related drivers will provide the impetus for such initiatives. Consumer demand for traceability due to concerns about food safety and national authorities requiring robust certification to support initiatives such as requiring pre-entry vaccination, rather than on-arrival quarantine will provide an immediate economic incentive to participate\(^7\). This is not ‘out there’—outcomes from this research have already contributed to, and will continue to guide an Asian Development Bank/Food and Agriculture Organisation project on enhancing market access through traceability systems. The benefit of this, from a disease prediction perspective, would be the availability of information in ‘near’ real time on animal movements (as in Australia) without having to rely on indirect predictors such as price.

7.2.4 Prices

The complexity associated with classifying animals makes it likely that better matching of the prices at origin and destination with the class of animals being moved will contribute to the improved ability of price to predict the movement of animals.

The greatest improvement to recording the value of animals would be to include the weight of the animal. Although during the course of this research the use of scales was observed for small ruminants and young pigs, the use of scales is not a routine component of the sale of larger stock. In the absence of suitable scales, training in estimation of liveweight and the use of girth tapes to assist in this process would be

\(^6\)Disney et al. (2001) are clear that the cost of whole of life traceability is high, and may be unnecessary for diseases such as FMD due to the short incubation and disease course. In fact, even the short duration of immunity from vaccination for diseases like FMD means that an animal based semi-temporary system (such as wrap-around tailtags or hoof/hair brands) may be perfectly adequate.

\(^7\)This would also support opportunities such as the development of a ‘fast-track’ for animals to travel directly to slaughter after entering the country, thus reducing the potential for local animals to become infected while allowing authorities to have confidence that animals could be traced.
Reliable methods for encoding this data for submission on a regular basis were discussed, and code sheets produced. It was noted that although it was possible to train individuals to do this, it was clear that the system was difficult to navigate. A promising recent development to assist this has been the reporting wheel discussed in Section 7.2.1.

7.3 Overall Conclusions

The risk of transboundary animal disease spread in the Greater Mekong Subregion could be predicted by price data but at the moment this is not feasible. The data collected during this research can predict movements, albeit poorly. There are a number of reasons for the limited predictive ability.

Currently, the sensitivity, specificity and timeliness of reporting into the regional animal health system are inadequate to provide an accurate picture of the true disease status. The use of FMD as a proxy for transboundary animal disease was a suitable choice for this research as it is a highly contagious disease of livestock with a significant economic impact, but little zoonotic risk and significant interest at the national and regional level. Nonetheless, without an accurate picture of this status at the origin of movements, being able to predict movements would not extend to being able to predict disease outbreaks at the destination. Reporting of disease requires increased awareness amongst producers of the need to report and awareness amongst animal health staff on the need for high quality investigations and expedient reporting. Without improvements in these areas, it is unlikely that the SEACFMD campaign can make significant progress in the control of FMD in the region. This situation is unlikely to be restricted to FMD, and it is perhaps of concern that even diseases with zoonotic potential such as Highly Pathogenic Avian Influenza and Rabies may be underreported.

Current data on animal movements in the GMS were inadequately categorised to
allow it to be matched to the price data. The data were incomplete—even if had been possible to collect all the data captured by movement permits, these are known to underrepresent the total number of livestock movements that occur. Most of the data are unavailable for practical use as the records are stored in paper books at the provincial level, making it difficult and expensive to collate, and impossible to do in a timely manner. By focussing on the relationship between origin and destination of movements, rather than the individual movements, it was still possible to gain useful information about high priority provinces for disease surveillance. More affordable technology and better availability of data networks offer a number of opportunities for improving the quality of data collected. Animal health continues to compete for limited resources in Cambodia and Laos and is unlikely to be a priority in Cambodia and Laos in the near future.

On the prediction side, poor quality price data (due to an incomplete classification of the animal to which the prices relate) and a lack of data from a combination of provinces and time periods conspired to reduce the accuracy and precision of the data on prices. International trade also complicates the picture, as the large number of movements associated with this trade (typically from Thailand to Vietnam during this study) are unlikely to be driven by the local market prices in provinces within Cambodia and Laos. None the less it was possible, even with all these constraints, to demonstrate a relationship between movement and price.

An important outcome of this research was the development of a method to improve the estimates of the cost of travel between provinces. Instead of using straight line distances (which are often completely improbable), distances were calculated based on the actual length of the road network most likely to be travelled between provincial centres.

The role of profit as a driver of trade is a key element of economic theory. This thesis has shown that the opportunity to make money can underpin the ability to predict animal movement. Combining this ability with improved data on disease will allow the prediction of the spread of transboundary animal disease.
Appendix A

Prospective Animal Tracking Study
A.1 Introduction

This study was a pilot attempt to prospectively track a cohort of animals to determine if they could be detected along the trade pathway subsequent to their departure from a village. A further outcome from this trial would be improved understanding of when and where the animals were moved to. Options for animal tracking were discussed in Appendix B. Due to the high infrastructure cost of machine readable devices, the choice was made to use plastic ear tags and advertise their presence to the potential detection points (the checkpoints and abattoirs). The work was done in collaboration with ACIAR Project AH/2006/159: Best practice health and husbandry of cattle and buffalo in Lao PDR. This project was conducting interviews and recording information about the effect of animal health interventions in villages and were thus well placed to insert the tags and record any subsequent tag loss (either from or with the animals) from their study site. It also tied in to other components of the Understanding Livestock Movements project, namely the checkpoint monitoring and abattoir monitoring in Laos.

A.2 Method

A sample of adult animals (≈250) in each of six villages in Northern Lao PDR were ear tagged with Plastic ‘Number 2’ ear tags (Allflex Pty Ltd). Each village had tags with a different colour (yellow, blue, pink, white, orange or green) and each tag had ULM\(^1\) written on it, and a unique number identifying the animal (Figure 1). The location of the villages is shown in Figure 2.

These tags were inserted during December 2008 by the District Agriculture Office staff as shown in Table A.1. The villages were visited regularly as part of a best practice research project, and any change in the animal population of the villages was recorded on a monthly basis. Any animals born or brought into the village during

---

\(^1\)An acronym for Understanding Livestock Movements, the name by which this activity was referred to in Laos.
Table A.1: Details of ear tagging of cattle in six villages in three provinces in Northern Laos

<table>
<thead>
<tr>
<th>Location</th>
<th>Village</th>
<th>Colour</th>
<th>Number</th>
</tr>
</thead>
<tbody>
<tr>
<td>Pakou, Luang Prabang</td>
<td>Ban Hardpang</td>
<td>Blue</td>
<td>258</td>
</tr>
<tr>
<td></td>
<td>Ban Hueypen</td>
<td>Orange</td>
<td>255</td>
</tr>
<tr>
<td>Pek, Xiengkhouang</td>
<td>Ban Nadee</td>
<td>Pink</td>
<td>250</td>
</tr>
<tr>
<td></td>
<td>Ban Nong</td>
<td>Yellow</td>
<td>250</td>
</tr>
<tr>
<td>Viengthong, Huaphan</td>
<td>Ban Naviang</td>
<td>White</td>
<td>249</td>
</tr>
<tr>
<td></td>
<td>Ban Nakud</td>
<td>Green</td>
<td>252</td>
</tr>
</tbody>
</table>

the study period were also tagged. The checkpoints along the main route from Vientiane to Xieng Khouang via Phoukhoun (near Luang Prabang) were provided with posters showing the ear tags and asked to be alert to the presence of these tags, such that if any animals were noticed in consignments they should be noted. Additionally, visits were made to abattoirs in Vientiane and Phonsavan (the capital of Xieng Khouang) and the project was discussed with owners and staff. A bounty was offered for any project tags that were recovered.

A.3 Result

After 24 months, none of the checkpoints along the main trade route between Vientiane, Luang Prabang and Xieng Khouang had reported seeing any of the tags. One tag was recovered from a slaughterhouse, but it was not one of the project tags used in this project.

A.4 Discussion

This activity was intended to trial an animal identification device for a larger animal tracking study. In this case, due to pragmatic considerations including the wide range of potential destinations and the cost of recording data at these destinations, a robust, simple and human readable device was chosen (in this case a plastic ear tag) with a range of colours and hindu-arabic numbers. A further hope of this pilot
study was to identify if any of the local village animals that were sold moved into the export chain into Vietnam or back towards Vientiane.

Although a number of tagged animals were sold from the villages during the period of this trial, very little information is available within the village on the destination of the animals. Even with a range of potential locations where the animals might be identified after movement from the villages, from none of these were animal tags reported. A number of reasons exist for this:

- animals passing through checkpoints were not unloaded, and observation of checkpoint activities would suggest that it is unlikely that more than a cursory glance is given to the contents of the truck on most occasions,
- some traders prefer to travel at night as their trucks are open on top, and although this reduces the heat-stress on the animals, it makes it nearly impossible to see anything in the truck,
- some checkpoints are not permanent or operating 24 hours,
- traders may remove obvious ear tags,
- there are a large number of abattoirs and slaughter points in Laos, and it was not feasible to have a presence at more than a few,
- abattoir staff may not take notice or remember the request to notify their findings (although the reporting of one tag not belonging to the project would not support this).

As a result of this activity it is concluded that such a capture-recapture trial design (a small scale trial constrained to a limited geographic area) is unlikely to be more successful without a larger investment in improving capture activities at a larger number of potential destinations.
Appendix B

Livestock Tracking and Tracing
B.1 Livestock Tracking and Tracing—Current Status

Gathering suitable information on livestock movements is a difficult task in the current regulatory environment of the Greater Mekong Subregion (GMS). Data were collected from a number of sources to allow analysis of the movements of livestock in Cambodia and Laos. These data were generally being collected for other purposes and was gathered retrospectively for analysis. While this provides valuable information about the nature of animals moving, it was quite likely that some animals would have moved that were not captured in these studies.

The method employed during this research relied on collating data from official forms associated with livestock movement into a database. This was a practical and efficient solution in the absence of any existing methods for producing data summaries.

In order to provide some validation of this information it was intended to conduct a small-scale prospective trial of animal tracking. As livestock in the GMS can move almost anywhere within the region, a method suitable for a large geographic area was required.

B.1.1 Tracking Options

Two methodologies currently exist for the prospective evaluation of animal movements—real-time tracking and capture-recapture techniques.

When the location of the animal at regular or specified times is desired, real-time tracking of animals using satellite or radio based locators is commonly used. This is suitable for the tracking of free-ranging animals, most commonly wildlife (Breed et al., 2010), feral animals (Johnson, 2000) or other non-livestock species (Wiehe et al., 2008).

Where animal presence at a known point is enough information, capture-recapture approaches using animal identification are often used. These have been used for
wildlife and are more common for livestock species (ie where the known point may be a farm, market or abattoir).

Whether or not the movement of the animals is under the control of livestock owners (and thus we are identifying owner behaviour, not animal behaviour) or the animals are free living may influence the choice of methodology. Tracking livestock over large distances may use either of these methods. Both are applicable for cattle and require the application of a tag or transmitter. When the movement of animals is conducted illegally, the application of a visible tag or tracking device may be ineffectual or the devices may be removed.

B.1.2 Tagging/Banding

This requires a visual tag with the capacity to incorporate enough information to allow identification of the likely origin. A simple brand, colour or shape may be sufficient for limited point-to-point tracking. If more details were required, a higher level of data complexity (ie alphanumeric codes) would be needed.

An apparent advantage of tags is that they are relatively cheap to produce, simple to apply and once applied tend to have a long life. However, they require a ‘capture’ network to be established to report on this system. This may have a significant cost, and may require incentives for reporting. This, in turn, may be open to abuse.

Tagging in this context also includes the use of Radio Frequency Identification Devices (RFID) such as ear tags and rumen boluses, as used in the National Livestock Identification System (NLIS) in Australia.

B.1.3 Satellite

The use of satellite based position location using triangulation\(^1\) for tracking of wild animals has become more common in recent times. There are numerous issues to

\(^1\)http://www.argos-system.org/
consider in relation to their use in livestock:

- Although much higher in cost, this provides enhanced spatial resolution, and a limited need for observers or reporting networks.

- The cost makes it prohibitive to monitor a large number of animals.

- Transmitters are limited in their size only by the capacity of the animal to live with the transmitter attached (which would not be likely to be a problem in cattle); and the need to provide power over longer periods of time.

- Such tags are commonly used for wildlife which are free to choose where and when to move. Livestock rarely move long distances of their own volition, and a transmitter/aerial would in most cases be anticipated to act as a deterrent to animal purchase.

### B.1.4 Biometric Analysis

A recent development in this area has been the use of biometric approaches (muzzle or retina scans) to identify animals along the production chain (Gonzales Barron et al., 2009). Although muzzle impressions have been used since the 1920s for high value stock\(^2\), this is a relatively new area for commercial use and the equipment has not been developed to manage it commercially.

### B.1.5 Genetic Fingerprinting

Genetic fingerprinting offers some advantages over biometric analysis for broader scale application, notably that regional variation may be identifiable in livestock at their origin. This can be linked to samples taken from animals at destination abattoirs in order to identify their origin.

Similar techniques are routinely used in tracing of bacteria and viruses and have been used in wildlife. They have also been used in livestock for determination of

\(^2\)Pre-dating fingerprinting in humans.
paternity during research (Holroyd et al., 2005).

B.1.6 Better Resolution—Global Positioning System

Instead of relying on satellite triangulation (as with the Argos system), animals equipped with a global positioning system (GPS) data-logger can provide very accurate location information constantly. The location data can be very cheaply logged to a data card, which requires the recapture of the animal and is suitable for local trials. Alternatively, it may be transmitted via radio or satellite to allow it to be traced in real time (and allow for the recovery of the equipment) (Clark et al., 2006; Trotter et al., 2010). The need to retrieve data cards (or have them pass within or past a specific base station) limits their usefulness for tracking animals over a large range, and engaging with satellites has a high cost associated with power requirements, and hence the size of the final device. In light of the difficulty of recovery of the equipment, the card based data-logger idea is unrealistic and will not be considered any further here.

A newer development is satellite based GPS linked animal tracking systems. This requires only a device that is the size of a normal ear tag. These have been described in very limited detail\(^3\). Attempts to learn more about these systems by contacting the companies involved typically have received responses along the lines of “We can’t discuss the technical aspects as they are proprietary.” and “The system is under construction.” (Stassen, G, personal communication, January 27 2009). Although both companies have registered with the Australian Communications and Media Authority, there is no evidence of any further development\(^4\).


B.2 A Proposed Novel Solution—A GSM Rumen Bolus

B.2.1 GSM Rumen Bolus Concept

A micro-generator (as found in a kinetic watch or torch) is connected to a power storage device and a GSM transmitter (Figure B.1). Normal ruminal activity moves the bolus and the micro-generator charges up the power storage. When the amount of power stored reaches a threshold value an electronic circuit is activated, turning on a GSM transmitter. The transmitter only needs to engage with the GSM network and register the device. The location can then be established by identifying which towers the device communicated with during the registration process. Information on the time, the location of the tower and the signal strength is then available through the network management system, and this information allows the animal to be approximately located within the working range of the tower—normally about 25km for GSM.

Although this area ($\approx 2\text{ 000 km}^2$) would not allow an individual animal to be ‘found’, it would provide sufficient resolution to allow useful information at a country scale to be recorded (for example, Figure B.2). After slaughter signal transmission would cease due to the lack of ruminal activity.
B.2.2 Rumen Boluses

Rumen Boluses are well established for long term pharmaceutical/mineral dosing, and are used with embedded RFID in the Australian NLIS (Fallon, 2001). For identification purposes the RFID boluses require a dedicated ‘reader’ device, which is not only expensive, but a relatively bulky object, thus making it unsuitable for tracking animal movements over long distances without the high cost of infrastructure. They also require excellent restraint of livestock for safe and successful insertion.

The use of a rumen bolus overcomes a number of issues associated with existing options, including:

- **Visible Deterrent**—The device is no longer visible to animal handlers, so should have no impact on the purchase price of the animal or it’s value in trade. This serves the dual purpose of enhancing co-operation amongst animal
owners, removing the risk of economic disadvantage that may be associated with visible devices and the risk that the device might be removed.

- **Price**—Some verbal quotes from electronics manufacturers suggest that devices could be produced relatively cheaply, with an end-point goal of less than $10, or even ideally closer to $6 per bolus. This compares favourably with the cost of plastic ear tags (around $0.70) and satellite tracking equipment (around $2 500).

- **Retention** Rumen boluses in use for NLIS have a retention rate exceeding 99.8% when inserted by an experienced operator (Ghirardi et al., 2006a). The most likely failure time is immediately after insertion, and so careful monitoring of animals is likely to allow identification of animals where insertion has been unsuccessful.

Data Capture can be fully automated through the GSM Mobile phone system—animals within the range of a mobile telephone tower during the time the system operates will be registered to the network at the tower location and this can be queried in real time to obtain information on the last data registration.

**B.2.3 Core Design Requirements**

**Range**

Range must be adequate to ensure reasonable coverage across multiple GSM towers. Transmission range for GSM is normally in the order of 25-30 km. This is restricted by protocol requirements (the requirement for signals to travel back and forward within a certain amount of time) more than transmission power.

**Power**

The power required to work across this range will be dependent upon whether it is necessary to go past the normal registration protocols to register the device to the
Normal registration ‘handshaking’ protocols only transfer a limited amount of data, although at a reasonably high level signal strength. Were it to be necessary to actually transmit a message, then it is likely that substantially more power may be required, and for a longer duration.

**Frequency**

Transmission frequency will also be important in determining the viability of the device.

In order to capture movement, ideally including at least some location recordings while the actual movement is being undertaken, then transmissions would be occurring around every six hours.

In the context of the spatial resolution offered by the transmission towers it is unlikely that more frequent reports will offer substantially improved resolution, and in fact too much reporting would be likely to place unnecessary strain on the system.

Conversely, reporting less than once a day would risk animals being moved to slaughter without any indication that they had ever left the property of origin.

**Message**

It is likely the minimum requirement would be undertaking a normal registration ‘handshaking’ protocol for GSM. The assumption so far is that the minimum requirement is establishment of power level for transmission and sharing of International Mobile Equipment Identity (IMEI) details to register the GSM device. If it were not possible to establish device location just by IMEI registration, the sending of a message could aid this. It is unlikely that this would allow geolocation of the animal if it were not possible to access tower details.
B.2.4 Bolus Physical Limitations

Overview

Rumen boluses are now established as a valuable alternative for permanent identification of ruminant livestock, using passive transponders. Initial research in this area was based on powered transmitters (and is recorded in Hanton J.P. (1981) Rumen-implantable electronic identification of livestock. Proc. U.S. Anim. Hlth Assoc., 85, 342-350, but it was not possible to obtain this reference).

Fallon (2001) lists some previous applications of rumen boluses:

- Delivery of trace elements including Selenium, Copper and Cobalt, using soluble glass or cement boluses.
- Dispensing growth promoters through a spring driven matrix of growth promoter that dissolved from one end of the bolus.
- Treatment of conditions including bloat (releasing detergents) and radiation poisoning (using hexacyanoferrates).

Rumen boluses have been in use for many years and the specific limitations on their design are generally well understood. These limitations are primarily related to the retention of the bolus in the rumen.

Dimensions

The size of the bolus must be small enough to allow its placement, but large enough to ensure that it is retained within the reticulum. Designing and matching the bolus to commercially available products would allow the use of a commercially available applicators for inserting the bolus into cattle. Commercially available products typically have a diameter of 2 cm and a length of 6 cm. It is feasible to increase this length by about 50% if required to fit extra components. Researchers have determined that volume is important for retention, but the specific gravity of the bolus is also important (Ghirardi et al., 2006b).
Specific Gravity

The specific gravity of the bolus is critical to retention, to ensure it falls neatly into the reticulum and is retained there throughout the life of the animal. It is also possible for the bolus to be retained in the general rumen, and Ghirardi et al. (2006a) report a bolus being recovered from the abomasum without apparent damage or blockage occurring.

A set of equations was determined from large scale trials, conducted in cattle in Spain, that determine the relationship between volume and specific gravity to maximise bolus retention in the cattle (Ghirardi et al., 2006a).

\[
W = 27.14 + 1.513V \tag{B.1}
\]

\[
SG = \frac{27.14}{V} + 1.513 \tag{B.2}
\]

Thus, assuming a length of 6cm and diameter of 2cm, a volume of \(6\pi\) would require a specific gravity of 2.95.

This must be achieved using materials which are not likely to impede radio transmissions, ie ceramic boluses or concrete filling of the bolus if using a plastic shell.

A glass bolus with a steel end has also been developed by Nedap Agri\(^5\). Depending on the material available this is likely to significantly reduce the available space for internal components.

Slightly lengthening the bolus may be an option, and would have a dual benefit of requiring a lower SG in the final product.

\(^{5}\text{http://www.nedap-agri.com/}\)
B.3. SIGNAL STRENGTH TESTING

Summary of Design Limitations

Bolus components are limited by a requirement to fit entirely within a limited space. It is not safe to assume that the components can completely occupy this space unless they are heavy as well. Unless the specific gravity of the device is within a specific range, it is likely that the bolus will not be retained in the reticulum and may pass into the small intestine or be regurgitated.

B.3 Signal Strength Testing

One of the important questions to answer was could the GSM transmitter put out enough power to travel through in excess of 30 cms of isotonic cow materials. Cows have an approximate osmolality across the plasma, interstitial and intracellular compartments of 280mOsm/litre (based on human values in Guyton, 1991). To develop some insight into this, a number of tests were run to evaluate GSM signal penetration through an ionic water solution. In this case, the decision was made to use sea water, which has is approximately 34 parts per thousand salinity equating to an osmolality of $\approx 1\,000$ mOsm/litre, about $3 \times$ that of normal tissues.

B.3.1 Method

Prior to application for animal ethics approval a preliminary trial was run using locally available salt water\(^6\). A test-rig was built using a GSM modem\(^7\) with an antenna on a 1.5m cable which could be enclosed in a water-tight enclosure\(^8\) sealed with silicone sealant\(^9\). The GSM modem was connected to a netbook computer\(^10\) running FreeBSD and a perl script was written to poll the modem every 10 seconds for signal strength\(^11\), and this was displayed on the screen. This test rig was mounted

\(^6\)Indian Ocean
\(^7\)Teltonika ModemCOM-G10, http://www.teltonika.it/
\(^8\)Graduated 1 litre leakproof flask, Decor
\(^9\)Roof and gutter sealant, Selleys
\(^10\)ASUS eeePC 701
\(^11\)Using an AT+CSQ request
A testing regime was identified that included testing for signal strength in eight situations. All sites were located on a beach in WA to minimise the risk of interference from other structures. The nearest GSM capable mobile phone tower was approximately 5km from the sites\textsuperscript{12}.

Testing regime:

1. sited on top of an esky at the beach in the vertical position;
2. sited on top of a rock at the beach in the vertical position;
3. sited on top of a light gauge steel tube\textsuperscript{13} on top of an esky at the beach in the vertical position;
4. sited in a plastic container on top of an esky at the beach in the horizontal position;
5. sited in a plastic container on top of an esky at the beach in the vertical position;
6. sited in a plastic container in the water at the surface in the horizontal position;
7. sited in a plastic container in the water at a depth of 10cm in the horizontal position;
8. sited in a plastic container in the water at a depth of 20cm in the horizontal position.

For each item above, 10 consecutive received signal strength indication (RSSI) values were recorded by a helpful research assistant (Figure B.3). As the mapping of RSSI to signal strength in decibels (dBm) is not linear the results are presented in RSSI. The mapping for the Teltonika modem is shown in Table B.1 for reference (Teltonika, 2006). These were subsequently analysed for summary statistics and presented as a boxplot. An intended replicate of these results was not performed due to water

\textsuperscript{12}Telstra Corporation, Australia
\textsuperscript{13}A dog food tin
ingress into the antenna enclosure

Figure B.3: Testing for signal strength when antenna is sited on top of a light gauge metal surface.

Table B.1: Mapping of received signal strength indication (RSSI) to signal strength (in decibels) for the Teltonika ModemCom-G10.

<table>
<thead>
<tr>
<th>RSSI</th>
<th>Signal Strength (dBm)</th>
<th>Display bars</th>
</tr>
</thead>
<tbody>
<tr>
<td>&lt; 4</td>
<td>&lt; -107 or unknown</td>
<td>0</td>
</tr>
<tr>
<td>&lt; 10</td>
<td>&lt; -93</td>
<td>1</td>
</tr>
<tr>
<td>&lt; 16</td>
<td>&lt; -71</td>
<td>2</td>
</tr>
<tr>
<td>&lt; 22</td>
<td>&lt; -69</td>
<td>3</td>
</tr>
<tr>
<td>&lt; 28</td>
<td>&lt; -57</td>
<td>4</td>
</tr>
<tr>
<td>⩾ 28</td>
<td>⩾ -57</td>
<td>5</td>
</tr>
</tbody>
</table>

A second test was conducted by enclosing a mobile phone\textsuperscript{14} (with vibrate mode activated) in a snaplock bag. A research assistant sent five consecutive text messages to the waterproofed phone while it was held 10 cm above the water, 10 cm below the water and 20 cm below the water. The receipt of these messages was noted by the primary researcher (Figure B.4) and confirmed by trainee research assistants. This sequence was repeated twice.

An experiment was also planned using cadavers associated with a local pastoral operation. Unfortunately, during the period of this research they did not identify any dead animals which could be used in this experimentation. The ingress of

\textsuperscript{14}Motorola Razr V
Figure B.4: Testing for receipt of text messages at specified depths below sea level.

Sea-water into the snaplock bag during the text message testing meant that the mobile phone which was intended to be placed in the body cavity of the cadaver was damaged and subsequently failed.

B.3.2 Results

The results of the signal strength experiment are shown in Figure B.5. This demonstrates that above water, varying the position of the antenna has little effect on the signal strength, but at a depth of 10cm there is nearly no signal, and at 20cm depth no signal was received.

The results of the text message experiment are presented in Table B.2. This suggests that even though some signal is known to penetrate to 10cm of depth, it is not adequate to allow the transmission of even a small packet such as a text message.

Table B.2: The number of text messages (SMS) received by a mobile phone at a given altitude/depth above the ocean surface at Cable Beach, Western Australia.

<table>
<thead>
<tr>
<th>Phone height</th>
<th>SMS sent</th>
<th>SMS received</th>
</tr>
</thead>
<tbody>
<tr>
<td>+ 10cm</td>
<td>5</td>
<td>5</td>
</tr>
<tr>
<td>- 10cm</td>
<td>5</td>
<td>0</td>
</tr>
<tr>
<td>- 20cm</td>
<td>5</td>
<td>0</td>
</tr>
</tbody>
</table>
B.3. SIGNAL STRENGTH TESTING

225

Figure B.5: GSM signal strength recorded in various antenna positions and locations. The signal strength is in arbitrary units (the vendor return signal strength indication) which can be mapped to the

B.3.3 Conclusions

Transmission of GSM radio signals was adversely affected by immersion of the antenna in an ionic solution. Furthermore, even at a depth where network signal strength could be measured, it appeared that it was too weak to be able to complete a low-bandwidth network procedure such as delivery of a 160byte text message. Although sea water has a concentration more than three times that of normal mammalian tissues, it was a convenient and readily available non-animal approach to testing the potential penetration of radio signals.

The primary goal of developing a GSM rumen bolus was to insert a device into a cow that would allow it to be tracked unknown to any subsequent purchaser. Concurrent with this research, the primary researcher was also involved in the formulation of an application to the Human Research Ethics Committee at Murdoch University\textsuperscript{15}. Based on this experience it was considered that inserting hidden devices into animals with the intention of tracking and being able to locate the animals that are being

\textsuperscript{15}Murdoch University HREC 2008/158: Identifying the cultural, social and economic drivers of animal movements in Cambodia and Laos
moved as a result of illegal activity by subsequent owners who may then face sanction by local authorities was unlikely to be deemed ethical research.

Due to the limited evidence of successful signal propagation and low probability of obtaining further research approvals development of the GSM rumen bolus was not pursued.
Appendix C

Reporting Market Prices by Text Message

Manual (English) and Code Sheets
C.1 Sending Text Messages

C.1.1 Project Overview

Diseases like Foot-and-Mouth Disease and Classical Swine Fever spread rapidly across long distances when animals are moved as a part of trade. Trade is important for increasing the profitability of livestock producers, and providing food in markets where supply does not meet demand. However, the spread of disease can cause significant hardship and losses to producers if their stock are affected. This project is working to better understand the patterns of livestock movement, and what causes animals to move around the region. We also hope to be able to use this information to rapidly assess the impact of a disease outbreak in terms of spread to another region.

Figure C.1: Map of Project Region
C.1.2 The Importance of Market Price Data

Animals clearly move in response to the price offered, and the greatest opportunity for profit for a trader is to move animals from locations where they can be purchased for less and sold for more. Because of this it will almost certainly be the most important component of understanding why animals move. Commonly when a disease outbreak occurs, the owners of stock in the vicinity of the outbreak will be offered lower prices for their animals – to offset the risk they face in keeping the animals and the animal becoming infected. However, at the time the animals are purchased, they may already be infected, but not showing clinical signs (the latent period).

C.1.3 The Problem of Data Collection

Collection of data across a large area can be difficult to coordinate. In order to make the best use of data, we need to receive it as soon as it is apparent. This is very important for disease outbreaks, so we can respond rapidly. It is also important for other factors that might influence a disease outbreak, especially animal movement. Normally the best way to collect data is through an established network, and feed it into an analysis that is prepared prior to the data being collected, making the results of the evaluation available rapidly for decision makers to assess. The best available network for the sort of data we require is probably the internet, or world-wide-web, but it is not available everywhere, and may not be available in rural areas at any stage. In fact, this project has a website\footnote{The website is http://trade.animalhealthresearch.asia/} which people who are submitting data can log into to see what the results the system is finding. However, the mobile telephone network in South East Asia now covers a very large percentage of the population, and many people are competent already in the use of mobile telephones, so we are looking to utilise it.
C.1.4 Using the Mobile Telephone Short Messaging Service (SMS)

Obviously, the easiest way to send data by telephone is to ring someone up and tell them. However, when we are only after small amounts of data, this is quite inefficient, as not only do we pay for the call cost, but also we require someone to answer the phone, take down the information, and then enter it into the computer for analysis. The last two steps also run a risk of the data being incorrectly recorded, leading to errors in the data input, and most likely the analysis.

But we can overcome this by instead using the Short Messaging Service (SMS) to send a text message with the information instead. This data can then be input directly into the computer, and the analysis run without requiring input from anyone other than the person submitting the data.

C.1.5 Limitations (and Solutions) to SMS

There are some limitations to a system like this.

1. Text Messages can only be 160 characters (letters, numbers, symbols etc) long.\(^2\)

2. Computers can be programmed to recognise words, but typing errors and multiple languages make this difficult, leading to erroneous results. Additionally, many handsets do not offer support for all the languages commonly used in the region (ie Lao, Khmer)

This is approximately as long as this sentence, which although it seems like a lot can be quite annoying when you suddenly run out of space to finish your message.

To overcome these limitations, we want to try using a message sending system where instead of sending whole words, we use single letters from the English alphabet to

\(^2\)In fact, they can be longer, but in fact the telephone splits the message into two or more parts and sends them both without you knowing – but you pay for each of them!
code for the message we want to send. In fact, this means we can send more than one price in a single message.

C.1.6 What Information Do We Need?

The information we need to make use of the data is actually very simple. For the price of any given commodity (ie a ‘type’ or ‘class’ of livestock) we need to know:

1. the details of the stock or commodity. a. Species – Cattle, Buffalo, Swine b. Sex – Male or Female c. Age – Young animal, Adult d. Condition of the animals – on a scale of 1 (poor) to 5 (very good) e. Breed (ie Exotic or Local for pigs)

2. the price a. the currency (ie USD, Kip, Real), b. the actual number value, and c. the units (per Head, per Kilogram)

3. the date the price was valid

4. The location the price was valid.

And that’s it... just four bits of information. So how do we do that quickly and easily, and fit it into the one SMS.

C.1.7 Two Rules

First we make two rules.

- One SMS can make one report for a date and location—but it can report many prices. So if you stop at a market or are talking to some farmers or traders, if they give you the price for 3 or 4 different types of animals, you can send them in the one SMS. If at the end of the day you have three prices for three different districts, that requires three messages.

- You put the information in the order below, and separate bits with spaces.
C.1.8 Message Format

The format of the message is simple, and looks a bit like this:

[Type] [location] [date] [Price data] [Price data]

and example message might look like this (don’t worry about the codes yet):

price KHM.2.5 25/11/2008 CMA3!U280H SMP3E!U30H

for a total length of only 54 characters. In fact, we could fit many more prices in the one message.

Type — price

This will always be the same – the message type will be price. This serves a number of purposes, such as ensuring that the message received is about price, and leaving the way open for the system to be used for other types of data collection (ie disease outbreaks)

Location — KHM.2.5

Although this doesn’t look very nice, it is actually very simple. The first three letters (KHM in this example) are the code for the country. The second number (2) is the number representing the province (in this case Battambang), and the third number (5) is the district (Aek Phnum) in the province. If you look in the back of this manual, there is a code sheet with these on. You will only need as much as covers the districts you are likely to be working in.

In fact, the letters for the country are not necessary if you are registered in the country you are submitting the data for. Just put the province number and the district number.
Date — 25/01/2008

If you are sending the date, please use the format day/month/year as above. Use only the slashes to separate the date, and don’t put any spaces in.

** If the price data is for today, you don’t need to put a date in.

Price Data — CMA3!U280H

Although this looks like a mess, it is made up of two main parts, the description of the stock and the price details, separated by a !, so let's look at them separately.

The description is pretty straightforward. We code animals according to the code sheet at the back, and it depends on their species. For cattle, we need (in order) Species, Sex, Age and Condition Score. Using the code sheet in the back, we can decipher CMA3:

C = Cattle
M = Male
A = Adult (fully grown animal)
3 = Condition score 3 (fair condition)

For the price, we always use the format Currency, Value, Unit; so we can also see how U280H is derived:

U = US dollars
280 = the price in US Dollars
H = Head -- the unit of price

So to code one ourselves, let’s try to code a prices for male piglets, largewhite in fair condition, selling for 30 US Dollars per head:

First, Species is Swine (S), Sex is Male (M), Age is piglet (P), Condition is fair (3) and breed is Exotic (E). SMP3E – easy.

Next, Currency is US Dollars (U), the value is 30 (30) and the unit is per head (H). U30H – also easy.
C.2 LOCATION CODE SHEETS

Put them together with a !, and we are done.

\text{SMP3E!U30H}

C.2 Location Code Sheets

To be supplied at an individual level

C.3 Commodity Code Sheets

\text{C} \quad \text{Cattle}

\begin{tabular}{ccc}
\text{Sexes} & \text{Ages} & \text{Conditions} \\
\text{Code} & \text{Sexes} & \text{Code} & \text{Ages} & \text{Code} & \text{Conditions} \\
M & Male & C & Calf & 5 & Good \\
F & Female & W & Weaner & 3 & Medium \\
C & Castrate & A & Adult & 1 & Poor \\
? & Unknown & M & Mature & & \\
X & Mixed & J & Juvenile & & \\
\end{tabular}

Figure C.2: SMS codes for cattle prices
### Appendix C. Reporting Market Prices by Text Message

#### Species

<table>
<thead>
<tr>
<th>Code</th>
<th>Species</th>
</tr>
</thead>
<tbody>
<tr>
<td>P</td>
<td>Pig</td>
</tr>
</tbody>
</table>

#### SMS Code Sheets

You can change the language by clicking on the appropriate flag in the top corner.

<table>
<thead>
<tr>
<th>Code</th>
<th>Sexes</th>
<th>Code</th>
<th>Ages</th>
<th>Code</th>
<th>Conditions</th>
<th>Code</th>
<th>Breeds</th>
</tr>
</thead>
<tbody>
<tr>
<td>M</td>
<td>Male</td>
<td>M</td>
<td>Mature</td>
<td>5</td>
<td>Good</td>
<td>X</td>
<td>Cross Breed</td>
</tr>
<tr>
<td>F</td>
<td>Female</td>
<td>A</td>
<td>Adult</td>
<td>3</td>
<td>Medium</td>
<td>L</td>
<td>Local</td>
</tr>
<tr>
<td>C</td>
<td>Castrate</td>
<td>P</td>
<td>Piglet</td>
<td>1</td>
<td>Poor</td>
<td>E</td>
<td>Exotic</td>
</tr>
<tr>
<td>?</td>
<td>Unknown</td>
<td>W</td>
<td>Weaner</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>X</td>
<td>Mixed</td>
<td>G</td>
<td>Grower</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

Figure C.3: SMS codes for pig prices
Appendix D

Code examples discussed in the thesis
D.1 Code for Standardised Morbidity Ratios

The below R code is an example of the code used to generate Standardised Morbidity Ratios from the years 2000–2009 in this work, as well as the Empirical Bayes Estimates using the lognormalEB function from the R package DCluster (Gómez-Rubio et al., 2005). The gms.1 data frame contains the information by province on the number of outbreaks for each year, and the number of villages.

```r
for(i in 2000:2009) {
    rx<- paste("reports",i,sep='\'')
    reports <- gms.1[[rx]]
    reports[is.na(reports)]<- 0
    r<- sum(reports)/sum(gms.1$villages, na.rm=TRUE)
    gms.1$totalexp <- gms.1$villages * r
    gms.1$Observed <- reports
    gms.1$Rate <- gms.1$Observed / gms.1$villages
    gms.1$Expected <- gms.1$totalexp
    gms.1$SMR <- reports / gms.1$totalexp
    gms.1$Expected[is.na(gms.1$Expected)] <- 0.00000001
    ebln <- lognormalEB(gms.1$Observed, gms.1$Expected)
    gms.1$EBLN <- exp(ebln$smthrr)
    brks <- c(0,1, 2, 5, 8, 20)
    grps<-as.ordered(cut(gms.1$SMR, brks, include.lowest=TRUE))
    SMRs <- cbind(SMRs,SMR=gms.1$SMR)
    EBLNs <- cbind(EBLNs,EBLN=gms.1$EBLN)
}
```

D.2 Procedural language functions for routing

The use of software for routing is increasingly common\(^1\). Before such software can work correctly, a number of aspects of the underlying map data need to be properly developed. The main one is the identification of all possible joins between line segments so that they can be described in a way that can be used by software. This

\(^1\)For example, using in car navigation devices or services such as Google Maps (http://maps.google.com/).
is known as network topology.

In this example, the routing is done using Dijkstra’s algorithm\(^2\), a popular and straightforward solution.

The first issue when attempting to undertake logical routing of animal movements between shires or provinces was to find a part of the shire to start the movement from or send it to. Based on a risk type perspective, it made sense to look for a saleyard or similar venue. These functions all use the open source pgRouting library\(^3\).

```
CREATE OR REPLACE FUNCTION gis.saleyard_exists(
    shire varchar(4),
    OUT saleyard geometry)
AS $BODY$
DECLARE
    BEGIN
        SELECT s.the_geom INTO saleyard
        FROM saleyards s
        JOIN lga t
        ON st_within(s.the_geom, t.the_geom)
        WHERE t.shortcode LIKE shire
        LIMIT 1;
    END;
$BODY$ LANGUAGE plpgsql;
```

If this function didn’t return any geometry two approaches were used. In Australia the centroid of the shire was used to approximate the spatially weighted mean of livestock properties. In Cambodia the location of the largest urban centre within the province was used. The next issue was to find the nearest road to the ‘centre’ of the shire.

This function requires a point geometry (ie from the `gis.saleyard_exists()` function and a nominated return value, and will return the unique record identification (gid), source or target value for the closest line segment work on the un-ambiguity principle, you can choose to spell out the return value you want, or just the first letter. Some of the conditions are:

- both geometries in the same projection system.

\(^2\)http://en.wikipedia.org/wiki/Dijkstra’s_algorithm
\(^3\)Available from http://www.pgrouting.org/
• geometry in the the_geom column

• geometry table already contains a unique id, source and target field

Obviously dynamic nomination of the table would be nice, but the function was written on a 14 hour plane trip, and I was vaguely under the impression that it stops caching and slows it down, and I was looking through 150 000 rows for the Western Australia road data.

It would be correct to point out that working in degrees does not give accurate distances, however in this function the accuracy of the distance is not vital, but the net order is the most important issue. If the absolute distance was important it would be necessary to reproject the data and use a projected measurement system (ie metres).

```
CREATE OR REPLACE FUNCTION gis.find_nearest_road(
    point geometry, sf varchar(6), OUT value int) AS
$BODY$
DECLARE
    max_search_radius real := 5.0; -- assuming working in degrees
    search_radius real := 0.01;
    rec mrwa_net%ROWTYPE;
BEGIN
    LOOP
        SELECT * INTO rec
        FROM mrwa_net m -- and you might want to change this
            WHERE st_expand(point, search_radius) && m.the_geom
            ORDER BY st_distance(point, m.the_geom)
            LIMIT 1;
        IF FOUND THEN
            IF substring(sf from 1 for 1) iLIKE 's' --start or Source
                THEN
                value := rec.source;
            ELSIF substring(sf from 1 for 1) iLIKE 't' --target
                OR substring(sf from 1 for 1) iLIKE 'f' --finish
                THEN
                value := rec.target;
            ELSIF substring(sf from 1 for 1) iLIKE 'g' --gid
                OR substring(sf from 1 for 1) iLIKE 'i' --id
                THEN
                value := rec.gid;
            END IF;
        END IF;
    END LOOP;
    value;
$BODY$
```
THEN
    value := rec.gid;
END IF;
EXIT;
END IF;
search_radius := search_radius * 2.0;
EXIT WHEN search_radius > max_search_radius;
END LOOP;
END;
END;
$BODY$ LANGUAGE plpgsql;

Example of using the function:

SELECT
    find_nearest_road(the_geom, 'source') as source,
    find_nearest_road(the_geom, 'target') as target
FROM saleyards
WHERE gid = 59;

CREATE OR REPLACE FUNCTION gis.join_shires(
sourcetta varchar(4), targettta varchar(4), algorithm TEXT,
OUT route geometry) AS $BODY$
DECLARE
    startpoint geometry;
    endpoint geometry;
    rsource int;
    rtarget int;
BEGIN
    -- look for saleyards
    SELECT saleyard FROM saleyard_exists(sourcetta)
    INTO STRICT startpoint;
    SELECT saleyard FROM saleyard_exists(targettta)
    INTO STRICT endpoint;
    IF startpoint IS NULL
    THEN
        SELECT c_geom INTO startpoint
        FROM lga
        WHERE shortcode LIKE sourcetta;
    END IF;
    IF endpoint IS NULL
    THEN
        SELECT c_geom INTO endpoint
    END IF;
    END;
$BODY$ LANGUAGE plpgsql;
D.2. PROCEDURAL LANGUAGE FUNCTIONS FOR ROUTING

FROM lga
WHERE shortcode LIKE targettta;
END IF;
SELECT value FROM find_nearest_road(startpoint, 'source')
INTO rsource;
SELECT value FROM find_nearest_road(endpoint, 'target')
INTO rtarget;
SELECT make_route(rsource, rtarget, 'dijkstra') INTO route;
END;

CREATE OR REPLACE FUNCTION gis.make_route(
    source int, target int, algorithm TEXT, OUT route geometry
) AS $BODY$
DECLARE
    BEGIN
    IF substring(algorithm from 1 for 1) ILIKE 'd' --dijkstra
    THEN
        SELECT st_collect((rt.the_geom)) INTO route
        FROM mrwa_net,
        (SELECT gid, (st_dump(the_geom)).geom as the_geom
        FROM dijkstra_sp_delta(
            'mrwa_net',
            source,
            target,
            3000) as di
        ) as rt
        WHERE mrwa_net.gid=rt.gid;
    END IF;
    END;

D.2.1 Road distance calculations in South East Asia

In order to calculate the road distances between provinces in South East Asia (SEA) it was necessary to:

1. identify suitable road data
2. manually edit it to ensure suitable line segments at the right locations

3. generate topology

4. locate a suitable province ‘centre’

5. create a line geometry between each province and measure the line

D.3 Livestock movement directions for online mapping

Being able to display the thrust of online movements in an ‘on-demand’ system using real-time data meant being able to determine a suitable start and end location for the movement, join those points to create a geometry suitable for online display (in this case a directed line to allow placement of an arrowhead) and adjust the length of the line to clearly show the route, but not end up with such a density of lines at the origin or destination that the information was obscured.

This code was used for dynamic map generation in a MapServer based system querying the database through a secure read-only connection.

```sql
SELECT s.id as mid, s.commodity, s.shortname, s.amount, s.purpose,
       s.origprov, s.destprov, ST_length(s.line) as distance,
       CASE
           WHEN ST_length(s.line) > 12
           THEN ST_Line_Substring(s.line, 0.05, 1 - (ST_length(s.line)-10) / ST_length(s.line))
           ELSE ST_Line_Substring(s.line, 0.05, 0.85)
       END as route, to_char(s.origindate,'DD/MM/YYYY') as date
FROM (SELECT m.id, c.commodity, sp.shortname, c.amount,
       initcap(r.name) AS purpose,
       initcap(o.name) as origprov,
       initcap(d.name) as destprov,
       ST_MakeLine(ST_Centroid(o.geom),
       ST_Centroid(d.geom)) as line,

```
D.3. LIVESTOCK MOVEMENT DIRECTIONS FOR ONLINE MAPPING

```
origindate
FROM gms.movements m
LEFT OUTER JOIN locations o
    ON m.originprovince=o.hiercode AND o.del=0 AND o.level=2
LEFT OUTER JOIN locations d
    ON m.destinationprovince=d.hiercode AND d.del=0 AND d.level=2
JOIN commodities c
    ON m.id = c.movementid AND c.del = 0 AND c.commodity = 1
LEFT OUTER JOIN species sp
    ON c.commodity = sp.id AND sp.del = 0
LEFT OUTER JOIN reasons r
    ON c.purpose::int = r.id AND r.del = 0
WHERE m.del = 0
    AND originprovince is not null
    AND originprovince not like '0'
    AND destinationprovince is not null
    AND destinationprovince not like '0'
    AND destinationprovince not like originprovince
) AS s
WHERE s.origindate > '%STARTDATE%'
    AND s.origindate < '%ENDDATE%'
ORDER BY s.origindate
```
Bibliography


Alliance Consulting and Management (2004). Cost analysis of NLIS compliance


Department of Agriculture and Food Western Australia (2005). NLIS minimum requirements for Western Australia.

Department of Agriculture and Food Western Australia (2006). Western Australia NLIS Cattle Directory.


Gilbert, M., Xiao, X., Pfeiffer, D., Epprechht, M., Boles, S., Czarnecki, C., Chait-


Hickler, B. (2007). Bridging the Gap Between HPAI “Awareness” and Practice in
Cambodia: Recommendations from an Anthropological Participatory Assessment. Technical report, FAO Regional Office for Asia and the Pacific, Bangkok.


Keeling, M., Woolhouse, M., Shaw, D., Matthews, L., Chase-Topping, M., Haydon,


Kitching, R. and Hughes, G. (2002). Clinical variation in foot and mouth disease:


OIE South East Asia Regional Coordination Unit (2011). *SEACFMD 2020: A roadmap to prevent, control and eradicate foot and mouth disease (by 2020) in South-East Asia and China*. OIE Sub-Regional Representation for South East Asia, Bangkok.


Pfeiffer, D. U., Robinson, T. P., Stevenson, M. A., Stevens, K. B., Rogers, D. J., and


Rweyemamu, M., Roeder, P., Mackay, D., Sumption, K., Brownlie, J., Leforban, Y.,


Scoizec, A. (2009). Study on cross-border movement and market chains of large ruminants and pigs in the Greater Mekong Subregion (GMS) - Lao PDR sur-
vey report. Technical report, Australian Centre for International Agricultural Research.


United States Board on Geographic Names (2010). Geonet names server. https://www1.nga.mil/ProductsServices/GeographicNames/Pages/.


