This file is part of the following reference:


Access to this file is available from:


*The author has certified to JCU that they have made a reasonable effort to gain permission and acknowledge the owner of any third party copyright material included in this document. If you believe that this is not the case, please contact ResearchOnline@jcu.edu.au and quote [http://eprints.jcu.edu.au/28985/](http://eprints.jcu.edu.au/28985/)*
The use of Envisat ASAR Global Monitoring Mode data to map rapid broad-scale flood events

Damien O’Grady
School of Earth and Environmental Sciences
James Cook University

Thesis submitted in partial fulfilment of the requirements for the award of
Doctor of Philosophy
November 2012
Statement on the contribution of others

<table>
<thead>
<tr>
<th>Nature of assistance</th>
<th>Contribution</th>
<th>Name</th>
</tr>
</thead>
</table>
| Intellectual support | Editorial assistance | Dr. Marc Leblanc (supervisor)  
                          |                     | Prof. David Gillieson (co-supervisor)                              |
| Financial support    | Grant              | ARC partial funding under  
                          |                     | Discovery grant DP110103364                                       |
| Data collection      | Permission         | Viv Sinnamon  
                          |                     | Kowanyama Aboriginal Land &  
                          |                     | Natural Resources Management Office (KALNRMO)                       |
|                      |                    | Anzac Frank (KALNRMO)  
                          | Rangers            | Phillip Mango (KALNRMO)                                           |
|                      |                    | Raven Greenwool (KALNRMO)                                           |
Acknowledgements

I would like to express my gratitude to my supervisors, Dr. Marc Leblanc and Prof. David Gillieson, for their wisdom, knowledge, support and for their confidence in my work.

Thanks go also the Australian Research Council, who partially funded this research under Discovery grant DP110103364, and to the European Space Agency, for the provision of radar data through project number C1P.5908.

I am grateful also to Viv Sinnamon of the Kowanyama Aboriginal Land and Natural Resources Management Office for his cooperation, and in particular rangers Anzac Frank, Phillip Mango and Raven Greenwool for their help with the data loggers.

Above all, thank you Flora and Ryan, for being just wacko enough to let me do this.
Abstract

This thesis seeks to enhance our ability to map the extent of large floods in near real time using coarse resolution C-band radar remote sensing. The microwave part of the electromagnetic spectrum has a great advantage over visible and infrared light in its ability to penetrate cloud cover, and as radar is an active system, it does not rely on daylight hours for reflected solar radiation. The European Space Agency’s Advanced Synthetic Aperture Radar aboard the Envisat satellite, operating in Global Monitoring Mode (GM), is targeted for particular consideration due to its high temporal frequency, comprehensive coverage and ease of acquisition. Challenges are identified which relate both to the use of radar generally, and also in particular to GM data, in the demarcation of water and land, as well as to the practical business of data processing.

These challenges relate to the way that water is identified, which can be by a low signal where specular surface reflection away from the sensor occurs, or by a high signal where multiple interactions occur between the water surface and emergent structures such as vegetation. Thresholds must make the distinction between the two cases, and as such, some prior knowledge of land cover is needed in the segmentation process. With such coarse data as GM, mixed pixels comprising both high and low water signals are often encountered, which result in a mid-range pixel value that masks the presence of water. The thresholding process is further complicated by the relationship of the signal returned to the sensor with incidence angle, which varies between about 14–44° with GM data. Under some wind conditions, waves of a particular pitch and orientation on the surface of open water cause resonance effects, returning a very high signal - sometimes
even a gain to the sensor. In particular circumstances, where flood waters flow through arid land, the low signal returned from open water due to specular reflection cannot be distinguished from the low signal returned from desert due to attenuation and absorption. In literature surrounding research in this field, results from observations of radar response in wetlands and flooded grasslands are mixed, pointing to the importance for further work in this area. In Australia, the need for a better understanding of the expected backscatter response from inundated areas in tropical savanna, which covers one third of its landmass, is clear.

The computational framework was set up for the efficient download, registration and orthorectification of GM data using scripting and open source software. Full advantage was made of the parallel processing capabilities of James Cook University’s High Performance Computing network, scripts were tailored to GM data’s characteristics and test results proved the method appropriate for the high volume processing required by the large GM dataset. This capability was used to carry out regression on a pixel-wise basis across a year’s worth of GM data, categorised by seasonal rainfall periods, in order to normalise backscatter values with respect to incidence angle. Correlation of the resulting characteristics with surface parameters, such as regolith, vegetation and soil type were observed. The potential confusion between absorption in dry, homogeneous soils, and specular reflection on surface water was predicted. It was observed that the degree of change of backscatter with incidence angle on open water appeared independent of the presence of Bragg Resonance, despite absolute values being at opposite ends of the scale, depending on whether resonance did, or did not, occur.

A major flood event in Pakistan was successfully mapped and made available in near-real time for the disaster relief effort. An image differencing technique allowed the successful separation of low backscatter response from open water with that from the immediately sur-
rounding desert. GM data were found to fill a gap in the period where the flood was obscured to visible and infrared sensors, during the crucial first week of the event. Definition of the extremities of the flood were tackled with a spatial threshold using a region-growing algorithm, and the radiometric backscatter threshold was established using an incremental convergence technique, employing multiple $\kappa$-statistic calculations with contemporaneous MODIS SWIR data. Both the stability of the radar threshold, and the instability of the MODIS SWIR reflectance threshold, were highlighted.

The backscatter responses to two large flood events in the tropical savanna of northern Australia were investigated, showing markedly different results. One flood, in the floodplain of Queensland’s Flinders River, involved total inundation of tussock grasslands over an area of 9000 km$^2$, allowing accurate classification using GM data ($\kappa = 0.7$), with predictable dihedral scattering returns as the flood receded and the emergent tussock grasses caused multiple interactions between the radar signal and the surface water. Inundated areas covered by emergent vegetation in the other flood, in Cape York’s Staaten River floodplain, were almost completely indistinguishable from the surrounding wet vegetation. Data from water height loggers established in the neighbouring Mitchell floodplain over a dry/wet season period provided an insight into the interaction of these particular vegetation conditions under flood. Results concurred with the work of others, that backscatter response is a complex combination of effects depending on relative water height, vegetation spatial density, biomass, and verticality, or enmeshment, of super-surface grasses.

The need for further work is discussed, together with spin-off opportunities, in the context of current and planned alternative C-band satellite data sources. The planned contribution of C-band data, along with contemporary visible/infrared products in the upscaling of current and ongoing JCU research into greenhouse gas emissions in the Mary River in the Northern Territory is outlined, together with the
possible use of C-band radar to gauge fuel moisture content and fire potential, in the light of our findings in the tropical savanna. The potential use of GM data to explore correlation between Gravity Recovery and Climate Experimient (GRACE) data and surface water and soil moisture over time is discussed.
# Contents

Contents

List of Figures vi
List of Tables x
List of Perl Scripts xi
Nomenclature xii

Introduction 1

0.1 Relevance and Importance 1
0.2 Challenges and opportunities 3
0.3 Objectives 5
0.4 Thesis structure 6
  0.4.1 Thesis flow diagram 6
  0.4.2 Chapter flow 8

1 Theory & feasibility - a literature review 11

1.1 Introduction 12
1.2 Theoretical context 14
  1.2.1 Radar remote sensing 14
    1.2.1.1 Modulation 14
    1.2.1.2 Backscattering and Attenuation 16
    1.2.1.3 Noise 18
    1.2.1.4 Radar Wavebands 18
1.2.1.5 Polarisation ............................................. 19
1.2.1.6 Local incidence angle ................................. 19
1.2.1.7 Pulse Frequency ....................................... 20
1.2.1.8 Resolution .............................................. 20
1.2.1.9 Radar Response to Water ............................. 21
1.2.1.10 Hydrological features and parameters .............. 22
1.2.2 Contemporary radar research ............................ 22
   1.2.2.1 Wavelength ........................................... 24
   1.2.2.2 Polarisation .......................................... 25
   1.2.2.3 Incidence angle ...................................... 26
   1.2.2.4 Modulation .......................................... 27
   1.2.2.5 Ancillary data ....................................... 28
1.2.3 Envisat ASAR Global Monitoring mode ................. 29
1.3 Methodological challenges with radar data ............... 32
   1.3.1 Speckle .................................................... 32
   1.3.2 Temporal resolution .................................... 32
   1.3.3 Spatial resolution ...................................... 33
   1.3.4 Ambiguity of response ................................. 33
   1.3.5 Local incidence angle ................................... 36
   1.3.6 Data acquisition ....................................... 38
1.4 Data analysis ................................................ 38
   1.4.1 Preprocessing .......................................... 38
      1.4.1.1 Despeckling ...................................... 38
      1.4.1.2 Georeferencing .................................... 40
      1.4.1.3 Orthorectification ................................ 40
      1.4.1.4 Database registration .............................. 41
      1.4.1.5 Sigma-nought versus beta-nought ............... 41
   1.4.2 Classification ......................................... 45
   1.4.3 Output .................................................... 52
   1.4.4 Analysis of accuracy ................................... 52
1.5 Gaps in existing research ................................ 54
1.6 Conclusion .................................................. 57
   1.6.1 Research questions .................................... 57
2 Practical high volume preprocessing of Envisat ASAR Global Monitoring Mode data over a distributed network

2.1 Introduction ................................................................. 60
2.2 Theory ................................................................. 62
   2.2.1 Theoretical basis .................................................. 62
   2.2.2 Orthorectification .............................................. 62
      2.2.2.1 Establishing the displacement error ................. 62
      2.2.2.2 Redistribution of cell value ....................... 64
      2.2.2.3 Establishing Radar Shadow ......................... 66
   2.2.3 Local incidence angle ........................................ 66
      2.2.3.1 To establish normal vector ....................... 68
      2.2.3.2 Final algorithm ..................................... 71
      2.2.3.3 Methodological errors ............................. 72
2.3 Method ................................................................. 74
   2.3.1 Choice of software and programming languages ........ 74
   2.3.2 Structure ....................................................... 75
   2.3.3 Parameters ..................................................... 78
   2.3.4 Multi-node network distribution .......................... 80
2.4 Results and discussion .............................................. 80
2.5 Conclusion ............................................................ 85

3 Relationship of incidence angle with satellite radar backscatter for different surface conditions

3.1 Introduction ............................................................. 87
3.2 Theoretical Context ................................................... 88
   3.2.1 Local incidence angle ..................................... 88
   3.2.2 Physical models .............................................. 90
3.3 Methodology ........................................................... 90
   3.3.1 Data ............................................................ 90
   3.3.2 Procedure ..................................................... 91
3.4 Results and discussion ................................................ 92
   3.4.1 Relationship of backscatter with incidence angle .... 92
   3.4.2 Variation with land surface properties .................. 94
4 Use of ENVISAT ASAR Global Monitoring Mode to complement optical data in the mapping of rapid broad-scale flooding: case study of the 2010 Indus flood

4.1 Introduction ............................................. 107
  4.1.1 Pakistan floods ...................................... 107

4.2 Study Area ........................................... 109

4.3 Theoretical Basis ...................................... 112
  4.3.1 Expected values ...................................... 114

4.4 Method ............................................... 116
  4.4.1 Data acquisition ..................................... 116
  4.4.2 Coverage ............................................ 116
  4.4.3 Image Preprocessing ................................. 116
  4.4.4 Image Differencing .................................. 117
  4.4.5 Baseline datasets from MODIS ..................... 119
  4.4.6 Thresholding and Classification ..................... 119

4.5 Results .............................................. 122
  4.5.1 Coverage ............................................ 122
  4.5.2 Image Differencing .................................. 122
  4.5.3 Thresholding and Classification ..................... 128
  4.5.3.1 Bivariate sensitivity analysis to determine threshold 128
  4.5.4 Inundation Dynamics ............................... 135
  4.5.5 Accuracy ............................................ 138

4.6 Discussion ........................................... 140
  4.6.1 Natural disaster response ........................... 140
  4.6.2 Use and limitations of the GM data for flood mapping 142
  4.6.3 A complement to other mapping techniques ........... 144
  4.6.4 Other applications of GM data ....................... 145

4.7 Conclusions .......................................... 146

5 Effects of vegetation on mapping of floods using satellite radar data ........................................... 148
## CONTENTS

5.1 Introduction .................................................. 149
5.2 Study Area and Flood Events ............................... 150
  5.2.1 Flinders .................................................. 151
  5.2.2 Staaten .................................................. 152
5.3 Theoretical Basis ............................................. 156
5.4 Method ....................................................... 159
5.5 Results and Discussion ..................................... 161
  5.5.1 Flinders .................................................. 161
  5.5.2 Staaten .................................................. 169
5.6 Conclusion .................................................... 171

6 Conclusion ....................................................... 172
  6.0.1 Research questions ...................................... 172
6.1 In summary ................................................... 174
6.2 Beyond the thesis ........................................... 178
  6.2.1 C-band synthetic aperture radar ....................... 178
  6.2.2 Mary River, Northern Territory ....................... 179
  6.2.3 Combination with other data types .................... 181
  6.2.4 Characteristics of fire occurrence and spread under a changing Australian environment .................. 182

Appendix A - Scripts ............................................. 186

Appendix B - Data ................................................ 210

References ....................................................... 224
List of Figures

1. Flow diagram outlining structure of thesis .............................................. 7
1.1 Map showing the frequency of ASAR GM mode coverage acquired for Australia .............................................................. 13
1.2 Relief displacement and foreshortening ................................................. 15
1.3 Layover ................................................................................................. 15
1.4 Radar shadow ...................................................................................... 16
1.5 Radar reflection from various surfaces .................................................... 17
1.6 Journal papers published on SAR techniques and applications between 1985 and 2006 .............................................................. 23
1.7 ASAR swath designations .................................................................... 30
1.8 Southern Cape York Peninsula: Mean of 12 GM images in 2009 .......... 34
1.9 Dry and wet backscatter reference curves ........................................... 37
1.10 Two GM data images of the Mitchell catchment using GM data ......... 44
1.11 ASAR Image with variance and entropy .............................................. 47
1.12 Influence of window size on the average overall accuracy of the classification results of texture features for different land cover types 48
1.13 Decision tree design from multi-temporal RADARSAT SAR images 49
1.14 Images obtained using MAR and PNN algorithms ............................ 51
1.15 Flow diagram outlining one possible processing sequence using methods discussed ................................................................. 53

2.1 Geometry used in the calculations for orthorectification .................... 63
2.2 Orthorectification value reassignment .................................................. 64
2.3 Satellite–Target geometry .................................................................. 66
2.4 Signing convention used in calculations .............................................. 67
LIST OF FIGURES

2.5 DEM Grid surrounding target pixel of height $e$ .......................... 68
2.6 Errors ($\Delta \alpha$) in local incidence angle ................................. 72
2.7 Antenna elevation pattern against Elevation angle ....................... 73
2.8 Class outline used ................................................................. 76
2.9 Flowchart showing preprocessing carried out by one node running process.gm.pl ................................................................. 77
2.10 Transect value profiles in Azimuth and Range directions across three bodies of water in Australia ........................................ 83
3.1 Function $f(x) = \cos^a \alpha$ ...................................................... 89
3.2 Range independent backscatter coefficient $\gamma_{30}$ ......................... 93
3.3 Slope $\Delta \gamma / \Delta \alpha$ (dB per degree) from linear regression for each quarter of 2009 ................................................................. 94
3.4 Comparison of $\gamma_{30}$ and $\Delta \gamma / \Delta \alpha$ for the third quarter of 2009 . 95
3.5 Scatter plot of MODIS Band 6 vs. $\gamma_{30}$ .................................. 96
3.6 Scatter plot of MODIS Band 6 vs. $\Delta \gamma / \Delta \alpha$ ......................... 97
3.7 Map showing regolith of Queensland ......................................... 99
3.8 Map showing dominant vegetation species of Queensland ............ 100
3.9 $\gamma$ and $\Delta \gamma / \Delta \alpha$ against dominant vegetation species ....... 101
3.10 $\gamma$ and $\Delta \gamma / \Delta \alpha$ against lithology classes .......................... 102
3.11 GM images of the Aral Sea on 18 June and 21 June 2010 ............ 104
4.1 Pakistan and the Indus-Chenab flood plain ................................. 110
4.2 Daily rainfall (mm) across the Pakistan region ............................ 111
4.3 Key map to describe range of $\Delta \sigma^0$ values derived from image differencing process ............................................................. 114
4.4 Density plot of MODIS Band 6 reflectance values over the flooded region on 27 and 29 August 2010 ........................................ 120
4.5 Count of frequency of cover by GM data over the 98 day study period .......................................................... 123
4.6 Percentage of full flood extent covered on each day of August 2010 by MODIS Terra, MODIS Aqua and GM data .......................... 124
4.7 Probability density functions for $\Delta \sigma^0$ for water and land ........ 124
4.8 The region between Jacobabad and Nawabshah in mid August 2010
### LIST OF FIGURES

4.9 Landsat composite colour image of the Indus and its floodplain southwest of Sukkur ................................................................. 127
4.10 Comparison of value profiles of $\Delta \sigma^0$ and MODIS Band 6 .... 129
4.11 $\kappa$ statistic calculated for individual classifications of flooding on August 10, 2010 ................................................................. 130
4.12 $\kappa$ statistic calculated for individual classifications of flooding on August 27, 2010 ................................................................. 131
4.13 $\kappa$ statistic calculated for individual classifications of flooding on August 29, 2010 ................................................................. 132
4.14 Flood extent estimates ................................................................. 133
4.15 $\kappa$ statistic calculated for individual classifications of flooding on August 27, 2010 ................................................................. 134
4.16 Selected instances from the time series showing the build-up of flooding and much of its recession .................................................. 136
4.17 Map showing the extent and duration of inundation surrounding the Indus and Chenab rivers ...................................................... 137
4.18 Distance of the flood head and tail along the Indus channel from the foot of the northern ranges at 71° N, 32° S ................................. 138
4.19 Area of inundation over time, of the Indus Channel and the total flood ................................................................. 139
4.20 Comparison of the $\Delta \sigma^0$-derived flood map on from 29 August 2010 against MODIS flood classification ................................ 140
4.21 Average frequency of terrestrial coverage per week by GM data between September 2009 and May 2011 .................................................. 143

5.1 Catchments under study in Northern Queensland, Australia ... 150
5.2 Flinders basin rainfall and river height at Walker’s Bend in early 2009 ................................................................. 151
5.3 GM Image of Flinders flood compared to MODIS Band 6 .......... 152
5.4 Staaten basin rainfall and river height at Dorunda in early 2007 . 153
5.5 Flooding of the Staaten river on 10 Feb 2007, MODIS vs. GM .... 154
5.6 Backscatter value density plots of the Flinders and Staaten regions, during their respective flood events ............................................. 155
<table>
<thead>
<tr>
<th>Figure</th>
<th>Description</th>
<th>Page</th>
</tr>
</thead>
<tbody>
<tr>
<td>5.7</td>
<td>Part of the Mitchell River during a flood</td>
<td>155</td>
</tr>
<tr>
<td>5.8</td>
<td>Dichanthium Sericeum</td>
<td>156</td>
</tr>
<tr>
<td>5.10</td>
<td>Location of water height loggers in the Mitchell floodplain</td>
<td>160</td>
</tr>
<tr>
<td>5.11</td>
<td>Capacitance loggers set up in the Mitchell floodplain</td>
<td>161</td>
</tr>
<tr>
<td>5.12</td>
<td>Variation of average GM backscatter values through the year</td>
<td>162</td>
</tr>
<tr>
<td>5.13</td>
<td>Progression of projected backscatter ($\gamma_{30}$) values in the Flinders region before and after the flood event</td>
<td>163</td>
</tr>
<tr>
<td>5.14</td>
<td>Two GM images of the Flinders floodplain, acquired 11 hours apart on 9–10 March 2009</td>
<td>164</td>
</tr>
<tr>
<td>5.15</td>
<td>Mean $\gamma$ values for classes determined using an unsupervised classification of five GM images straddling the Flinders flood event</td>
<td>165</td>
</tr>
<tr>
<td>5.16</td>
<td>Accuracy of flood classification using GM data against a MODIS Band 6 threshold classification</td>
<td>167</td>
</tr>
<tr>
<td>5.17</td>
<td>Water level against $\gamma_{30}$ for five logger locations</td>
<td>168</td>
</tr>
<tr>
<td>5.18</td>
<td>Tallest stratum growth-form in flooded areas</td>
<td>170</td>
</tr>
<tr>
<td>6.1</td>
<td>Recent rainfall shown as high backscatter values</td>
<td>185</td>
</tr>
</tbody>
</table>
List of Tables

1.1 Synthetic Aperture Radar Roughness at a Local Incident Angle of 45° .................................................. 17
1.2 Radar band designations .................................................. 19
1.3 ASAR Global Monitoring Mode Image product summary ........... 30
1.3 ASAR Global Monitoring Mode Image product summary (contd.) 31

2.1 Average revisit frequency capability per 35-day orbit cycle ........ 61
2.2 Parameters extracted from GM data file using Pds2Ascii ............ 79
2.3 Spatial displacement in pixels between edges of sampled water bod- ies as identified in outputs ............................................. 84

3.1 Singular Values (SV) indicating categorical separability among land cover classifiers ........................................... 98
3.2 Mean and standard deviation $\gamma_{30}$ and $\Delta \gamma/\Delta \alpha$, for regolith classes 103

4.1 GM Data used in this study ............................................. 117
4.1 GM Data used in this study (Contd.) ................................ 118
4.1 GM Data used in this study (Contd.) ................................ 119
4.2 Error matrix and $\kappa$ statistic for the flood map on 29 August 2010 when compared with MODIS flood classification. ................. 141

5.1 Kappa ($\kappa$) results from the accuracy test performed on the classi- fication of flooding in the Flinders basin on 15 February 2009 ....... 169
## List of Perl Scripts

<table>
<thead>
<tr>
<th></th>
<th>Script</th>
<th>Page</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>GrassAscii.pm</td>
<td>186</td>
</tr>
<tr>
<td>2</td>
<td>ORTHO.pm</td>
<td>187</td>
</tr>
<tr>
<td>3</td>
<td>ALPHA.pm</td>
<td>191</td>
</tr>
<tr>
<td>4</td>
<td>GM.pm</td>
<td>194</td>
</tr>
<tr>
<td>5</td>
<td>process_gm.pl</td>
<td>198</td>
</tr>
<tr>
<td>6</td>
<td>pbs.pl</td>
<td>202</td>
</tr>
<tr>
<td>7</td>
<td>regression.pl</td>
<td>203</td>
</tr>
</tbody>
</table>
Nomenclature

Greek Symbols

\( \alpha \)  
Local incidence angle

\( \beta^0 \)  
Radar brightness coefficient

\( \phi \)  
Off-Nadir angle

\( \gamma \)  
Range-independent (ground-range projected) backscatter coefficient

\( \kappa \)  
Kappa Statistic

\( \Lambda \)  
Look angle

\( \lambda \)  
Wavelength

\( \Omega \)  
Angle subtending the target and the satellite, at the centre of the ellipsoid

\( \omega \)  
Angle subtending the \textit{apparent} target and the satellite, at the centre of the ellipsoid

\( \sigma \)  
Radar cross-section (m\(^2\))

\( \sigma^0, \sigma_n^0 \)  
Normalised radar backscatter coefficient

\( \sigma^0_\alpha \)  
Radar backscatter coefficient at incidence angle \( \alpha \)

\( \theta \)  
Incidence angle (context-specific)

Other Symbols

CH\(_4\)  
Methane
NOMENCLATURE

\( \text{CO}_2 \)  Carbon Dioxide  
\( \text{NO}_2 \)  Nitrogen Dioxide  
\( \text{det} \)  Determinant of a matrix  
\( G^2_{ele} \)  Two-way antenna elevation pattern gain  
\( \text{HH} \)  Co-polarised Horizontally  
\( \text{HV} \)  Cross-polarised, H on transmission, V on reception  
\( K \)  Absolute calibration constant  
\( S \)  Slant range  
\( S' \)  Nominal zero-Doppler slant range  
\( \hat{v} \)  Unit vector (vector of length 1)  
\( \mathbf{v}_1 \times \mathbf{v}_2 \)  Cross product of two vectors  
\( \text{VH} \)  Cross-polarised, V on transmission, H on reception  
\( \text{VV} \)  Co-polarised Vertically  

**Acronyms**

AEP  Antenna Elevation Pattern  
ANN  Artificial Neural Network  
ASAR  Advanced Synthetic Aperture Radar  
SAR  Synthetic Aperture Radar  
dB  Decibels  
DEM  Digital Elevation Model  
DN  Digital Number - raster image pixel value  
ENL  Equivalent Number of Looks
<table>
<thead>
<tr>
<th>Acronym</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>ESA</td>
<td>European Space Agency</td>
</tr>
<tr>
<td>GDAL</td>
<td>Geospatial Data Abstraction Library</td>
</tr>
<tr>
<td>GIS</td>
<td>Geographic Information Systems</td>
</tr>
<tr>
<td>GLCM</td>
<td>Grey Level Co-occurrence Matrix</td>
</tr>
<tr>
<td>GM</td>
<td>Envisat ASAR Global Monitoring Mode</td>
</tr>
<tr>
<td>GRD</td>
<td>Gross Rock Descriptor</td>
</tr>
<tr>
<td>HPC</td>
<td>High Performance Computer</td>
</tr>
<tr>
<td>JCU</td>
<td>James Cook University</td>
</tr>
<tr>
<td>LS</td>
<td>Lacustrine Sediment</td>
</tr>
<tr>
<td>PNN</td>
<td>Probabilistic Neural Network</td>
</tr>
<tr>
<td>REGEX</td>
<td>Regular Expressions</td>
</tr>
<tr>
<td>SRT</td>
<td>Slant Range Time</td>
</tr>
<tr>
<td>SWIR</td>
<td>Short-wave Infrared</td>
</tr>
<tr>
<td>TS</td>
<td>Tropical Savanna</td>
</tr>
<tr>
<td>VIR</td>
<td>Visual / Infra-red</td>
</tr>
</tbody>
</table>
Introduction

The primary objective of the research described in this thesis is to enhance our ability to map the extents of large floods.

In recent decades, our dependence on empirical data for such a task has given way to rapidly developing satellite imaging technology, which has enabled us to successfully map flood inundation to a fine degree of accuracy, and has contributed greatly to the study of flood dynamics and to our efforts in disaster relief and prevention.

0.1 Relevance and Importance

Flooding amounts to nearly 50% of natural disasters, and accounts for over 70% of all lives affected by natural disasters (Kugler et al., 2007). Justification for research into the understanding of floods covers a broad area of science. In order to understand the hydrology of large river systems, it is vital that the dynamics of inundation patterns are better understood (Frappart et al., 2005). Wetlands receive increasing attention as indicators of changing biodiversity. Powell et al. (2008) tell us that “Recent research into water requirements for wetland systems shows that duration, frequency, depth, timing and extent of flooding are the most important influences on ecological communities”, and that “Modelling these systems is hampered by a lack of data and inappropriate model structures”. Signatories to the Ramsar Convention are encouraged to set up national wetland inventories in order to advise government on policy and management strategies, yet a review has discovered gaps in coverage of such inventories and problems with the associated information management systems; Rosenqvist et al. (2007) point out that The Asian Wetland Inventory and Convention Resolution VIII.6
Introduction

specifically highlight remote sensing as a key tool in rectifying this situation.

Environmental degradation is an increasing concern in respect of river catchments. Some changes in land use and management practices cause increases in surface run-off which, in turn, can have a major effect on the volume and rate of discharge of water. This can cause erosion and large-scale changes to fluvial behaviour patterns as well as increasing the potential for increased phosphates to be carried downstream (Bonn & Dixon, 2005). Henderson & Lewis (2008) point out that mangroves are considered “the world’s most productive ecosystem based on net productivity”, and that over 54% of the world’s mangroves have disappeared over the last few decades. The characteristics of the annual inundation pulse in the Amazon are, according to Rosenqvist et al. (2002), the “dominant environmental factor affecting aquatic biota on the floodplain”.

Rosenqvist et al. (2002) bring up another issue which is widely discussed, and that is the quantities of methane which are produced in flooded areas. This amounts to some forty times the quantity produced by wet soils (Arnell, 2002). Wetlands also play a part in the accrual of heavy metals in surface waters (Baghdadi et al., 2001). Methane is produced during the anaerobic decomposition of biota, and thus is particularly prevalent during the inundation of terrestrial vegetation (Noernberg et al., 1999), and therefore an understanding of flood patterns against vegetation maps will become an increasingly important tool to assist in the modelling of greenhouse gas emissions (Rosenqvist et al., 2002).

Under certain environmental conditions, perhaps the best satellite instruments to use for the separation of water from land are “passive” instruments (relying on reflected solar energy or on thermal emissions, for example) that operate at wavelengths between the visible and infrared ranges of the electromagnetic spectrum, due to the fact that radiation through much of the infrared spectrum is absorbed by water.

There are, however, inherent limitations to the use of passive sensors. The reliance on the relative position of the sun halves the data acquisition period (on average, depending on latitude and season), rendering us blind to night-time flood activity. Perhaps more significantly, radiation at these wavelengths is absorbed, reflected or attenuated by atmospheric water vapour to such an extent that the surface flooding is hidden from the satellite sensors by cloud cover, commonly
present at most flood events.

The research described in this thesis, therefore, singles out the use of active sensors operating in the microwave range ($\lambda \approx 1\text{mm} - 1\text{m}$). Such instruments analyse reflected radiation originating from their own source, and therefore are not reliant on the sun. In addition, atmospheric effects on microwave radiation at certain wavebands (such as C-band, studied here), are negligible by comparison to those which effect the light used by passive sensors.

0.2 Challenges and opportunities

The use of radar data, specifically, to map flooding, is well researched and reported. Although the advantages over passive sensors, as described above, have been confirmed and put to good use, much of the published material on the subject comprises conspicuously large sections documenting the limitations of radar data for the purpose, and for good reason. These limitations stem from major differences between microwave and optical data, and the methods used to acquire images:

1. Optical data is two-dimensional, the physical position of receipt of data on the “retina” of the instrument is analogous to the relative position on the earth’s surface from which the energy was reflected.

   Radar data is one-dimensional, the position of the reflector having to be calculated using a variety of geometric and temporal parameters. Terrain affects radar imagery differently to optical imagery, and requires different processing. Terrain also causes radiometric, as well as geometric distortion of values, and in some cases causes total masking, which must be taken into account.

2. Microwave sensors operate in narrow wavebands, and the radiation is therefore subject to summation interference effects to which all coherent radiation is susceptible, causing a “salt and pepper” effect on the resultant image, known as speckle.
3. Regular textural patterns at certain orientations, such as small wind-induced waves on the surface of water, cause additive effects due to resonance, which result in an amplified signal, the value of which lies at the opposite end of the scale to what we would expect from water.

4. Received values have a relationship with the angle of the transmitted radiation incident to the surface of the reflector. This angle must therefore be known, and the relationship understood. The relationship itself depends on the nature of the reflector (in terms of scale, structure and dielectrics), which compounds the complexity of requirements at the processing stage.

5. Where structures such as vegetation or buildings emerge through flood water, multiple reflections can cause a deceptively high signal response, which falls into the range of values expected from wet soil, making it difficult to accurately define the boundaries of a flood.

6. The radar backscatter response to open water is usually characterised by a low value, due to specular reflection away from the sensor by the surface of the water. Errors in classification are contributed to by other smooth surfaces causing a similar reflection, or by extremely dry, homogeneous surfaces such as desert sand, which present a low radar value due to attenuation and absorption of the incident radiation.

As stated, the above are factors which limit the use of radar data to map flooding, and most have been well documented. What is lacking, however, is a measure of the precise extent to which these factors may be mitigated, and beyond which the use of radar data to map flooding is precluded.

Mapping relatively small floods (such as in urban areas) to a fine detail requires relatively high spatial-resolution data. Data at such a resolution is usually confined to a temporal frequency of one per orbit cycle (usually just over a calendar month). Thus for a fast flood event, there is an element of luck in employing such data, and even when timing is favourable, unless the flood is extraordinarily slow in progression, its dynamics may only be interpreted from one or two instances in time.
Mapping floods that occur on a large scale, spanning tens or hundreds of kilometres, however, may make use of coarse resolution data such as that made available by the European Space Agency (ESA) from their Advanced Synthetic Aperture Radar (ASAR) instrument aboard the Envisat satellite, when operating in Global Monitoring Mode (GM). Much higher temporal frequency is achieved, as this instrument records data across a swath over 400 km in width, which overlaps considerably (depending on the latitude) with each consecutive orbit cycle.

Analysing the response of GM data to large flood events gives us the opportunity to fill the gap in research - to define the constraints governing the use of such data for the important task of understanding large and rapid flood events. At the same time, the massive dataset of GM output affords us the ability to understand the relationship of incidence angle with different surface reflectors, and allows us to investigate how this, and the other effects which compromise the accuracy of flood maps, may be mitigated.

0.3 Objectives

The ultimate objective is, therefore, to record an optimum method to monitor and map large flood events quickly, to a reasonable degree of accuracy, and to understand the environmental and geophysical conditions to which such a method is limited. This objective is approached by seeking to address the following questions:

- How can the large dataset be managed, the incidence angle and other geospatial data be extracted and the imagery preprocessed in a fast, efficient way, and in such a manner as to maintain complete control and transparency for all calculations performed in the process?

- What is the relationship between incidence angle and radar backscatter for different surface properties? Can knowledge of this relationship be used to enhance our ability to delineate flooded regions? If necessary, how can incidence-angle effects be most accurately mitigated?
Introduction

- How can ambiguities in radar backscatter due to absorption, under particular environmental conditions, be reduced? How can the most accurate thresholds be established in floods whose scale, timing and geographical constraints forestall the acquisition of empirical data? Answering such a question may be critical for disaster response efforts.

- How do existing vegetation and rapidly-growing aquatic vegetation affect radar response in flooded regions? To what extent can these effects be mitigated or taken advantage of? To what degree do these effects preclude the use of GM and other radar data under certain environmental conditions?

0.4 Thesis structure

This thesis records research and analysis that was carried out in order to answer these questions. Following the introductory sections which cover theory and related research, the questions are tackled systematically in independent chapters. The major successive chapters are structured as comprehensive stand-alone papers, for submission to research journals. Two of the articles are under review, and one has been published. The format and content of the articles are altered here only to provide continuity and to avoid undue repetition of common but necessary theoretical explanations which will appear independently in the individual submissions.

0.4.1 Thesis flow diagram

Figure 1 shows a schematic outline of the structure of this thesis. The diagram shows the overall flow of the work, the broad reasons for paths taken, the direct and indirect outcomes of the research and brief summary of work that is proposed to follow on from it. Colours are coded roughly as follows:

<table>
<thead>
<tr>
<th>Colour</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>Black</td>
<td>Sections, chapters of the thesis</td>
</tr>
<tr>
<td>Red</td>
<td>Gaps, questions, obstacles</td>
</tr>
<tr>
<td>Blue</td>
<td>Solutions</td>
</tr>
<tr>
<td>Green</td>
<td>Outcomes, direct or indirect</td>
</tr>
</tbody>
</table>
Figure 1: Flow diagram outlining structure of thesis
Headings in the left margin refer to relevance within the thesis. In the diagram, what is shown in the Introduction and Literature Review sections are summaries of the major challenges and the logic that leads on to the decided focus of research, whereas these actual chapters in the thesis cover much more detail which need not be shown schematically here. The Gaps section of the diagram results from the concluding part of the literature review, and represents the focus of the work to follow. From this point of the diagram onwards, items are arranged in a matrix, with horizontal groupings separating articles from direct and indirect outcomes, and with vertical columns grouping problems, solutions and outcomes under the article, and therefore chapter, in which they arise. The findings listed in these groups all form a part in the Conclusion section of the thesis, which is not explicitly labelled in the diagram.

A Chapter context diagram is shown at the beginning of each chapter to explain the relevance of the chapter in the context of the whole thesis, and to prevent the reader from becoming lost. These are based on the Thesis Flow Diagram above (Figure 1). The connectivity diagrams are milestones, or place markers, and therefore do not appear in the List of Figures.

0.4.2 Chapter flow

The chapters appear as follows:

- The front matter of the thesis includes the contents, lists of figures and tables and a list of the nomenclature (symbols and acronyms) used throughout.

- Following the introductory chapter, a literature review presents an explanation of the theory behind the use of radar remote sensing in the mapping of floods, in the context of research literature to date. Here we discuss the parameters that play a part in the processing of radar data for this purpose, and the challenges that they present.

- The method of managing the data, extracting the parameters and performing the geo-referencing and orthorectification, using object-oriented scripting and parallel processing over a distributed network, is detailed. This
Introduction

is a necessary precursor to all the work that follows. Methods to address the questions discussed in this introduction, and arising from the literature review, form the basis of the following chapters, each of which represents a journal submission:

– The effects of incidence angle on radar backscatter, given different surface characteristics, such as regolith, vegetation and soil types, is investigated. In particular, the separability of water in comparison to those parameters is measured. The relationship of incidence angle with backscatter is established by carrying out regression for each pixel over time, making full use of the high temporal frequency of GM data available. The work results in a means to normalise radiometrically across a wide swath with respect to incidence angle, which is essential to be able to apply a threshold to the pixel values to determine the presence of flooding. Two problems are highlighted as a result of this chapter. One is a phenomenon known as Bragg Resonance, for which a potential solution is discussed. Another is a problem with ambiguity in backscatter values which could represent water, due to absorption in dry sand. A solution to this problem is tackled in the following chapter.

– Is this problem with ambiguity due to absorption likely to present a real problem? How might a solution be found? A case study is performed, mapping the rapidly developing extents of a major flood event in Pakistan in 2010. This study was initiated as the result of a request by a UNESCO team, who were about to fly to the region as part of the relief effort, lacking maps detailing the extent of the flooding. Methods to mitigate incidence angle effects and sand absorption ambiguities are successfully carried out, and a method to non-empirically optimise thresholding is detailed.

– Having successfully established the suitability of complementing remote sensing data from the optical / infra-red spectra with GM data to map floods in an arid / semi-arid environment with a low-height vegetation matrix, the applicability of this method to a northern Aus-
Introduction

Australian environment is investigated. This addresses the fourth major challenge highlighted in the gaps section of the literature review, related to dihedral scattering. An investigation into the limitations presented by vegetation on the use of radar remote sensing to delineate water is carried out, by studying two major floods in the north of Queensland, which took place in catchments with different vegetation matrices. This chapter highlights the need for more detailed research, which is outlined in the subsequent chapter. Backscatter observations are compared to fieldwork data, with interesting results.

- The following chapter outlines the proposed next step, to take the benefits arising from the research carried out for this thesis and pave the way for its application to the broader environment. The suggested work builds on research that has been commenced and is ongoing by a team from JCU’s Hydrology department, who are taking measurements in the field in a bid to estimate fluctuations in greenhouse gas emissions in a flood plain in the Northern Territory. The research described in this proposal would serve to provide the existing team with the remote sensing function required to upscale their own results to the wider floodplain, whilst at the same time using its spatially detailed data to work towards a better inversion model for the interpretation of radar data to map flooding. The potential use of C-band radar to investigate fire fuel conditions, given findings from the previous chapter, is outlined.

- A conclusion summarises the results of the thesis, discusses shortfalls and summarises the avenues of future research that arise from it.

- The appendix contains the program code developed to pre-process the data, as discussed in Chapter 2, and that used in the regression analysis discussed in Chapter 3. It also lists the data used in some of the chapters.
Chapter 1

The Theory & feasibility of mapping surface water with satellite radar data - a literature review

Chapter context
Abstract

A review of the mapping of water extents, using radar data, is made, briefly outlining the theory behind the field, as well as the many challenges faced, including speckle, spatial resolution, variance with incidence angle, temporal frequency and ambiguity in interpretation. With regards to opportunities presented by the large time series of data available, four broad areas are identified that contain gaps in existing literature: The practical methodology behind the registration and orthorectification of a very large data set; the relationship of backscatter response with incidence angle specific to pixel-scale regions; the separation of water with dry soil, which share low backscatter values due to absorption; the separation of high backscatter response from dihedral scattering with that from wet soil.

1.1 Introduction

This review aims to study the use of radar remote sensing to map the extents of water inundation over time. Of particular interest is the feasibility of using Envisat Advanced Synthetic Aperture Radar (ASAR) Global Monitoring Mode (GM) data as a source for this task, complimented with data from other ASAR modes and from other active and passive sensors. The detailed specifics of GM data will be discussed later in section 1.2.3. Briefly, ASAR provides radar backscatter measurements in the C-band ($\lambda = 5.6\text{cm}$) in a 35 day repeat cycle orbit, with data being provided in a variety of modes with different specifications (ESA, 2007a). For certain modes, temporal frequency is greatly increased using ASAR’s ability to steer its beam to different elevation angles. A consequence of this is that spatial resolution is compromised, and a greater complexity associated with data coming from a variety of incidence angles is introduced. GM is one such mode, having the lowest spatial resolution and highest temporal frequency of all of the ASAR modes. A graphic representation of the temporal frequency of GM data available covering the Australian continent is shown in figure 1.1.

GM data has a swath width of 405km, a pixel size of 500m and a spatial resolution of 1km. Such a spatial resolution would seem to constrain the value of GM data to broad-scale views of radar-distinguishable features that cover areas
greater than a few square kilometers. The European Space Agency (ESA) originally envisaged GM data’s greatest value being in the monitoring of Antarctic ice sheets, with ASAR’s Image Mode Precision (IMP) mode targeted for flood monitoring, having 30m spatial resolution and a 5 day repeat coverage available on request (ESA, 2007a). However, in the past some frustration has been expressed over the tendency for the data request and acquisition process (generally) to cause peak flood events to be missed (Oberstadler et al., 1997; Sanyal & Lu, 2004). Additionally, it will be shown that the potential for detailed and extensive time series analysis afforded by the high temporal frequency and ready availability of GM data allows us, to some extent, to overcome issues of low spatial resolution, providing us with the means for the effective distinction and mapping of inundated areas not envisaged previously, enabling also the mapping of flood durations and patterns. Broadly, this review seeks to explore the extents (and therefore the limitations) of using GM data for this purpose.
1. Literature review

1.2 Theoretical context

1.2.1 Radar remote sensing

This section is intended to serve as a theoretical introduction to radar remote sensing for the mapping of water extents.

1.2.1.1 Modulation

The first step to analysing any image data is to understand what the values associated with each pixel represent. The two parameters considered in respect of coherent electromagnetic radiation are amplitude and phase.

The phase difference between two radar signals is used in interferometry to provide a very accurate elevation value in, for example, SRTM DEM data (Lillesand & Kiefer, 2004). The value of each pixel in a DEM image represents this elevation. Phase information is included with some data products (Single Look Complex, for example) usually in the form of two files comprising the real and complex conjugates.

The value dealt with in this analysis is the Digital Number (DN), which is the pixel value of the GM data and is related to radar brightness $\beta^0$ by the equation

$$DN^2 = K \cdot \beta^0 = K \cdot \frac{\sigma^0}{\sin(\alpha)}$$  \hspace{1cm} (1.1)

where $\sigma^0$ is the radar backscatter coefficient, $\alpha$ the local incidence angle and $K$ the absolute calibration constant (ESA, 2004). The significance of $\sigma^0$ will be dealt with in detail in section 1.4.1.5.

Whilst the radar brightness $\beta^0$ is a measure of amplitude and is analogous to the brightness value from an optical sensor, there is a fundamental difference, and this is to do with the way that a return signal’s spatial attributes are determined. Most electromagnetic sensors, such as a camera, are directional, in that the part of the sensor in which the signal is received depends on the spatial position of the area from which the signal was reflected. In this respect, the received signals form a two-dimensional analogy of the area under observation. With a radar sensor carried on an aeroplane or a satellite, only the spatial dimension along the
azimuth, or flight path, is controlled. The returning signal is one-dimensional, and the spatial dimension perpendicular to the azimuth (the range dimension), is determined by calculating the time it takes for the signal to get back to the sensor. This leads to a problem which is fundamental to understanding radar images: a single signal can return to the sensor at the same time from two different positions along the swath. This is demonstrated in Figure 1.2, taken from Delio Tortosa’s website *Geoforum* (Tortosa, 2008). In this case, C and D are in separate locations along the swath, but are combined in the return signal (represented by C’D’), as they are the same distance from the sensor. This phenomenon, tending to shorten the apparent displacement between two points, is known as *foreshortening*.

![Figure 1.2: Relief displacement and foreshortening (Tortosa, 2008)](image)

From the transformation of AB to A’B’, it can be seen that an identical physical structure can be represented differently, depending on where it lies along the swath.

![Figure 1.3: Layover (Tortosa, 2008)](image)

Figure 1.3 shows another symptom of the radar imaging process, known as
layover, where the relative positions of A and B are actually reversed in the output image.

Knowing the relative heights of such features gives us the ability to calculate the distortion effects, and to some extent make corrections to the data. Different levels of processing offered by distributors of radar data take such effects into account.

Another limiting factor which can be seen from Figure 1.4 is radar shadow. It can be seen that objects over a certain height at a certain incidence angle will obscure adjacent areas below that incidence angle. These radar shadows can be “filled in” by combining two sets of data covering the exact same area, but taken from opposing sides. Such image pairs from, for example ALOS PALSAR, are available in one file. Another way to mitigate this problem to some extent is by choosing a smaller incidence angle, but this has effects on resolution and attenuation, as will be seen.

Figure 1.4: Radar Shadow (Tortosa, 2008)

1.2.1.2 Backscattering and Attenuation

The strength of a returned radar signal depends on a variety of response parameters, including physical structure, surface properties, electrical characteristics and attenuation. Figure 1.5 shows how different surfaces will produce different responses. Specular surface reflection will produce very little backscatter towards the sensor. ‘Rough’ surfaces cause varying degrees of backscattered signal, largely determined by the Rayleigh Criterion. Lillesand & Kiefer (2004) cite a preferred modified Rayleigh Criterion, which determines a surface to be rough when
1. Literature review

\[ h_{\text{rms}} > \frac{\lambda}{4.4 \times \cos \alpha} \]  (1.2)

and **smooth** when

\[ h_{\text{rms}} < \frac{\lambda}{25 \times \cos \alpha} \]  (1.3)

where \( h_{\text{rms}} \) is the rms surface height variation, \( \lambda \) is the wavelength and \( \alpha \) is the local angle of incidence (Lillesand & Kiefer, 2004). The summary shown in Table 1.1 gives an idea of how this relates to X and L-band at, for example, \( \alpha = 45^\circ \).

There are alternative definitions—for example for the *Fraunhoffer* criterion for smoothness, replace 25 with 32 in equation 1.3.

<table>
<thead>
<tr>
<th>Root-Mean-Square Surface Height Variation (cm)</th>
<th>X Band</th>
<th>L Band</th>
</tr>
</thead>
<tbody>
<tr>
<td>((\lambda = 3.2 \text{cm}))</td>
<td>Smooth</td>
<td>Smooth</td>
</tr>
<tr>
<td>0.05</td>
<td>Smooth</td>
<td>Smooth</td>
</tr>
<tr>
<td>0.10</td>
<td>Intermediate</td>
<td>Smooth</td>
</tr>
<tr>
<td>0.5</td>
<td>Rough</td>
<td>Intermediate</td>
</tr>
<tr>
<td>1.5</td>
<td>Rough</td>
<td>Rough</td>
</tr>
<tr>
<td>10.0</td>
<td>Rough</td>
<td>Rough</td>
</tr>
</tbody>
</table>

Table 1.1: Synthetic Aperture Radar Roughness at a Local Incident Angle of 45° (Lillesand & Kiefer, 2004)

![Figure 1.5: Radar reflection from various surfaces: (A) specular reflector, (B) diffuse reflector, (C) corner reflector. (Tortosa, 2008)](image)

**Figure 1.5:** Radar reflection from various surfaces: (A) specular reflector, (B) diffuse reflector, (C) corner reflector. (Tortosa, 2008)

Attenuation of radar radiation through a medium is largely a function of the medium’s **complex dielectric constant**. For our purposes, it is sufficient to know that water has a dielectric constant over ten times that of most dry minerals (Lillesand & Kiefer, 2004, p671).
Microwave radiation is refracted through media with higher dielectric constants in much the same way as optical light through optically denser media. It is slowed down, its path is refracted towards the normal (thus usually made shorter) and its wavelength is actually increased according to

\[ \lambda = \lambda_0 \sqrt{\epsilon_0 \epsilon_a} \]  \hspace{1cm} (1.4)

where \( \lambda_0 \) and \( \epsilon_0 \) are the wavelength and permittivity through free space, and \( \epsilon_a \) is the average dielectric constant through the penetrated medium (Schaber, 1999).

Radar response specifically to water will be dealt with in section 1.2.1.9.

### 1.2.1.3 Noise

Noise appears on radar images mainly in the form of speckle, and this is largely due to constructive and destructive interference of incident and reflected radiation. It therefore stands to reason that cross-polarised data (where the incident and reflected light are polarised in orthogonal planes) will be less subject to this form of interference. This and other forms of speckle are mitigated by processing *multilook* images, where multiple images of the same area are added, and in which such noise is averaged out (Lillesand & Kiefer, 2004). Despeckling filters, both spatial and temporal, will be looked at in relation to GM data in section 1.3.1.

### 1.2.1.4 Radar Wavebands

Radar wavelengths are grouped into ranges, or wavebands, identified by a single letter. Table 1.2 shows all of the common wavebands and their wavelengths and frequencies. Currently only L, C and X-bands are available from active sensors aboard satellites.

A more detailed look at response characteristics of the wavebands is taken in Section 1.2.1.9.
1. Literature review

<table>
<thead>
<tr>
<th>Band Designation</th>
<th>Wavelength ( \lambda ) (cm)</th>
<th>Frequency ( \nu = c\lambda^{-1} ) [MHz (10^6 cycles sec(^{-1}))]</th>
</tr>
</thead>
<tbody>
<tr>
<td>( K_a )</td>
<td>0.75–1.1</td>
<td>40,000–26,500</td>
</tr>
<tr>
<td>K</td>
<td>1.1–1.67</td>
<td>26,500–18,000</td>
</tr>
<tr>
<td>( K_u )</td>
<td>1.67–2.4</td>
<td>18,000–12,500</td>
</tr>
<tr>
<td>X</td>
<td>2.4–3.75</td>
<td>12,500–8,000</td>
</tr>
<tr>
<td>C</td>
<td>3.75–7.5</td>
<td>8,000–4,000</td>
</tr>
<tr>
<td>S</td>
<td>7.5–15</td>
<td>4,000–2,000</td>
</tr>
<tr>
<td>L</td>
<td>15–30</td>
<td>2,000–1,000</td>
</tr>
<tr>
<td>P</td>
<td>30–100</td>
<td>1,000–300</td>
</tr>
</tbody>
</table>

Table 1.2: Radar band designations (Lillesand & Kiefer, 2004)

1.2.1.5 Polarisation

Emitted radiation may be polarised either horizontally (H) or vertically (V). The received signal may be similarly polarised. There are, therefore, four possible polarisation configurations — HH, HV, VH, VV. Reflection by different materials at certain incidence angles and at certain wavelengths has a polarising effect, and the use of polarisation therefore has a filtering effect, which can either be beneficial or detrimental. Lillesand & Kiefer (2004) note that due to the complex nature of the relationship between polarity and the parameters at play, it is difficult to know whether using cross-polarised or co-polarised radiation is going to be best for a particular application until they are tried.

1.2.1.6 Local incidence angle

The local incidence angle \( \alpha \) is the angle between the radiation vector and the normal of the object surface. Whilst the look angle (or angle of elevation), which is the angle between the sensor direction and the nadir, is known, the local incidence angle is usually not, and unfortunately it is this latter parameter which has a direct bearing on how incident radiation will be scattered. It has complex effects. An increased angle gives us worse layover, foreshortening and shadow, as discussed in Section 1.2.1.1. However, according to Equation 1.2, a given surface’s ‘roughness’ will increase with an increased incidence angle (and therefore look angle), as found by Robinson et al. (2006) (Section 1.2.1.1 above refers). It will also be seen later that greater angles of incidence allow structural characteristics
to have an enhanced effect on backscatter, which is particularly important in the investigation of biomass (section 1.2.2.3 refers).

Schaber et al. (1997) tell us that a reduced incidence angle results in less speckle for co-polarised images, whereas not much difference is observed with cross-polarised images.

It should also be noted that increasing the incidence angle improves the range resolution, according to the following equation, adapted from Lillesand & Kiefer (2004):

$$R_r = \frac{c\tau}{2\sin \alpha}$$

where $\tau$ is the pulse duration (discussed next).

1.2.1.7 Pulse Frequency

Radiation is emitted from radar sensors in bursts, or pulses, of energy at a given frequency. Two object points very close together may return pulses which overlap, or merge, making the two points indistinguishable. Therefore, a shorter pulse duration (i.e. a higher pulse frequency) serves to improve range resolution. This may not seem important for our purposes, as range resolution is always quoted with available data products, but it serves to illustrate that increasing the look angle will also improve range resolution.

1.2.1.8 Resolution

Having discussed the range resolution (perpendicular to the azimuth), there remains the azimuth, or line-of-flight, resolution. Synthetic Aperture Radar (SAR) devices use a very wide broadcast angle, which would normally result in a very poor azimuth resolution. However, spatial position of an object parallel to the azimuth is finely resolved by measuring the frequency of the return signal, which is increased ahead, and reduced behind, the position of the sensor (due to the Doppler effect). Thus the azimuth resolution depends on the smallest discernible difference in frequency that can be detected by the sensor. Effectively, it seems that with most data products, the azimuth and range resolutions are coordinated towards a similar figure.
1. Literature review

1.2.1.9 Radar Response to Water

Open water, in the absence of resonance effects, or roughness due to weather causing waves, has a relatively smooth surface which causes radar radiation to be reflected away from the sensor, resulting in a low return signal (Henderson & Lewis, 2008).

Schaber et al. (1997) describe a phenomenon which is akin to Bragg Resonance in optics, where regularly spaced scatterers can cause resonant amplification of a signal, even if the scatterers and their distances apart are tiny with respect to larger undulations on which they are superimposed, and that this signal can “dominate” the response in the case of satellite sensors. Resonance conditions take the form

\[ \frac{2\Lambda}{\lambda \sin \theta} = n \quad n = 1, 2, 3 \cdots \] (1.6)

where \( \Lambda \) is the spatial wavelength, \( \theta \) the incidence angle and \( n \) any integer. This is relevant where any regularly-spaced deposits are found. The phenomenon allows for the possibility of tiny ripples caused by a breeze on the surface of still water to produce a high backscatter signal, where we would expect specular reflection to cause a low signal. In a large body of water, these capillary waves (on a scale of centimetres) are combined with the gravity waves (which occur on the scale of metres) to form a range of waveforms called the wave spectrum (Woodhouse, 2006).

Given that wind gives this structural component to water, and that the orientation of the ripples vary, it follows that the resultant backscatter is also a function of the relationship between the horizontal azimuth look angle and the orientation of the waves. The resultant backscatter coefficient \( \sigma^0 \) can be represented by the idealised function

\[ \sigma^0 = A + B \cos(\phi - \phi_R) + C \cos^2(\phi - \phi_R) \] (1.7)

where \( \phi \) is the azimuth look angle, \( \phi_R \) the orientation of the waves with respect to the same reference, \( A \) is the mean backscatter, \( B \) the upwind/downwind variation and \( C \) the asymmetry of the ripples (Ulaby et al., 1982; Woodhouse, 2006).
1. Literature review

1.2.1.10 Hydrological features and parameters

In the absence of water, hydrological features can be determined mainly by their structural characteristics. In a single, uncalibrated image, a feature can be determined if it contrasts with its surroundings. In the case of radar imaging, this contrast is likely to be due either to dielectric or scalar/roughness inhomogeneities between the feature and its surroundings. If a structure such as a drainage channel is significant enough in scale, it may have edges facing the radar sensor which appear very bright, with dark bands behind them due to layover, foreshortening and shadow effects as discussed in Section 1.2.1.1.

Vining & Wiseman (2006) record that relict meander scars appear as low-return “radar dark” arcs in Radarsat data, which is attributed purely to different textural qualities. This is corroborated by Schaber et al. (1997), who point out that non-calcified small-gravel alluvium is specular, and hence radar-dark. Schaber et al. (1997, p344) tell us that co-polarised data are more sensitive to surface roughness on the scale of the particular band’s wavelength, whereas cross-polarised data may be more sensitive to geometry or texture.

Surface infiltration capacity is a parameter in flood prediction, and surface moisture an indicator of saturation. In fact, soil moisture, surface roughness and incidence angle together combine to produce large variations in backscatter (Pathe et al., 2009). Surface roughness in itself is an important parameter in flood prediction, as it controls the speed of surface run-off (Bonn & Dixon, 2005).

1.2.2 Contemporary radar research

Henderson & Lewis (2008) carried out a review of the use of radar remote sensing in the detection of wetland ecosystems, consolidating prior reviews and updating the contemporary status of such research. This serves as an excellent source of information.

In his review of the past and potential use of Synthetic Aperture Radar (SAR), Gens (2007) gives us a statistical comparison of the volume of journal publications on the use of SAR, against their subject, which is reproduced in figure 1.6.

As mentioned previously, there are many examples of the use of radar remote sensing to map flooding. Some examples are the assessing of the extents of flooded
1. Literature review

Figure 1.6: Journal papers published on SAR techniques and applications between 1985 and 2006 (reproduced from Gens (2007))


Research using radar remote sensing can be classified according to the attributes of the data which effect the scope of detection capabilities: wavelength, polarisation, incidence angle, modulation and the use of complimentary data. These attributes are considered in the following sections.
1.2.2.1 Wavelength

It has been seen that certain parameters which effect backscattering by a responder are a function of wavelength — refraction (equation 1.4), Bragg resonance (equation 1.6) and roughness (equation 1.2), for example. These dependencies conspire to dictate that structural scale of responders effect backscatter from radiation of different wavelengths, differently. It therefore follows that the choice of waveband when considering the use of radar remote sensing depends on the predominant structure of the object under investigation, or that of its environment. When attempting to map the extents of flood inundation in a river catchment, it is the aquatic, riparian and terrestrial vegetation which must be considered in addition to physical structure and orientation, as it is these which interact with the incident radiation and scatter it.

What soon becomes clear when reviewing research in this field is that when a vegetated area becomes flooded, one cannot consider the expected response of the water and that of the vegetation separately, as their effect is synergistic. As we have seen, water, when smooth, is a specular reflector. Vegetation, depending on its type, size and structure, and on the wavelength of the radiation, is a volume backscatterer and will therefore tend to return a higher signal to the sensor. During inundation, therefore, in cases where such vegetation is totally submerged, the net effect is a drop in backscatter to some degree. During partial submersion, however, radiation which is scattered by the vegetation towards the water (which under dry conditions would most likely be absorbed by the soil or substructure) undergoes specular reflection from the surface of the water and enters back into the volume scatterer. This process is often called “double-bounce” (Parmuchi et al., 2002), or “dihedral scattering” (Rosenqvist et al., 2002). This will be discussed further in section 1.3.4.

Generally, C-band is considered most suitable for the study of smaller-structured vegetation components (Noernberg et al., 1999)—leaf-off low biomass conditions of deciduous trees, for example (Henderson & Lewis, 2008), aquatic plants (Henderson & Lewis, 2008) and wetlands dominated by herbaceous vegetation (Parmuchi et al., 2002). It is also to be expected that growth of a plant will see its relative scattering effects on C-band and L-band, for example, change. Pope
et al. (1997) found that a rise in water levels in a marsh following burning show no increase in backscatter in L-band, but a 6 dB increase in C-band response, presumably from dihedral reflection. In fact, Pope et al. (1997) go on to say that an increase in C-band backscatter can sometimes correspond with a decrease in that of L-band.

For reasons of scale, it is not surprising that C-band is more sensitive to ripples on water surfaces (Hostache et al., 2009a), and we would expect L-band to better respond to slightly larger wavelets. Of course, for the purposes of this study, it is not the observation of such waves that concerns us, so much as their elimination, as their presence only serves to make water bodies less distinguishable from the surrounding environment.

It may generally be considered that longer wavebands, such as L-band, produce a higher response in forested areas, and C-band in areas of shorter, sparser vegetation (e.g. Henderson & Lewis (2008)). Multiple wavebands would obviously be the optimal acquisition. On an encouraging note for the use of the C-band GM data, many consider C-band to be the most sensitive to flooding. Pope et al. (1997), whose data included phase information (section 1.2.1.1 refers), found that in a comparison of C and L-band co-polarised and cross-polarised configurations (amplitude modulated) and C and L-band phase modulated data, the most sensitive to flooding was C-band phase difference. Perhaps unfortunately for GM data users, the same study found that C-band HH data was the worst for detecting some flooded marshes, but that, once again, this was very much dependent on the size and structure of the vegetation—Cattail and Sawgrass marshes, in the Yucatan Peninsula, showed a significant increase in C-HH backscatter when flooded.

1.2.2.2 Polarisation

As discussed previously, it is difficult to predict the best cross-polarised or co-polarised configuration for a particular purpose until they have been tried and tested on the particular area under investigation. This notwithstanding, Henderson & Lewis (2008) tells us of parallels drawn by Ramsey (1998) “between the interaction of radar parameters (especially polarization) with vertically oriented
flooded trees at L-band and vertically oriented stalks of flooded herbaceous vegetation at shorter wavelengths.” Interestingly, Ramsey (1998) found that variation in attenuation of C-band signal with incidence angle was not dependent on polarisation, whereas this was not the case at L-band (Henderson & Lewis, 2008).

Noernberg et al. (1999) show us that like-polarised signals (HH and VV) give us clear separation of open water with various vegetation types in the area of -25dB to -31dB, but with cross-polarised signals, the water response may be confused with that of dead trees. This is perhaps intuitive, in that both water and tree trunks present a dominant structure which is orthogonal to one of the two axes of polarisation.

In their evaluation of the suitability of C-band SAR data for the mapping of wetlands, Baghdadi et al. (2001) classified 3 images at various polarisations as a function of incidence angle, reporting accuracies of 74%, 76% and 59% for HH, HV/VH and VV respectively (based on error methods which will be discussed in section 1.4.4).

Bourgeau-Chavez et al. (2001) found HH polarisation to be better than VV for the discrimination of wetland while mapping riparian ecosystems in Virginia using C and L-band data.

1.2.2.3 Incidence angle

The theoretical contribution of the local incidence angle to radar backscatter was briefly discussed in section 1.2.1.6. A more detailed look at this relationship will be taken when considering the proposed preprocessing of GM data. In terms of contemporary research, the consideration depends, once again, largely on the use to which the data is put.

There is general consensus that, where canopy needs to be penetrated in order to experience the maximum dihedral scattering to discriminate flooded from non-flooded woodland or forest, a low angle of incidence is best (Bourgeau-Chavez et al., 2001; Henderson & Lewis, 2008; Ramsey, 1998; Sanyal & Lu, 2004; Töyrä & Pietroniro, 2005; Töyrä et al., 2001). This, as Ramsey (1998) points out, is due to the increased signal/canopy interaction at higher incidence angles. Later in this review, when considering the preprocessing of GM data, the issue
over the possible normalisation of pixel values in terms of incidence angle will arise. Given the potential for dihedral scattering and other interactions that vary with incidence angle and with the structural nature of the scatterer that are discussed here, it is reasonable to suggest that no such normalisation is possible without a precise knowledge of the underlying structure and orientation. Given how quickly vegetation can change, especially in the tropics, this may never be practicably achievable in the scale at which this review wishes to look. Section 1.4.1.5 deals with this issue.

If, indeed, a lower angle of incidence increases dihedral scatter in flooded forests, then it follows that a ratio between data from a higher incidence angle and that from a lower angle will be an excellent indicator of flooding in wooded environments. As discussed, with smooth open water, the predominant response results from specular reflection away from the sensor. For this reason, Kandus et al. (2001) recommend large incidence angles to discriminate the interface between land and water (Henderson & Lewis (2008), Kandus et al. (2001)). A ratio, therefore, of high incidence angle with low would seem a good method for distinguishing open water. In fact, such a combination was found to have significantly higher accuracy when classifying open water, as compared to the use of single images by Töyrä & Pietroniro (2005).

1.2.2.4 Modulation

Potentially, any of the parameters mentioned in sections 1.2.1.1 and 1.2.1.2 which effect the scattering of incident radar radiation on a subject may be used in detection. The Rayleigh and Fraunhoffer criteria were discussed, which govern the property roughness. As part of their research into flood forecasting by estimating surface run-off risk, Bonn & Dixon (2005) make use of this relationship to attempt to measure surface roughness using Radarsat data. Incidentally, with reference to the previous section, Bonn & Dixon (2005) find that greater incidence angles provide greater radiometric resolution in terms of surface roughness.

Due to the fact that water tends to reflect most of the incident signal away from the sensor, a simple threshold is the most common method used to classify water in the research studied (Abhyankar et al. (2007); Baldassarre et al. (2009))
and Hostache et al. (2009a) are examples). There are variations on this theme—Chandran et al. (2006), for example, use a threshold on the variance of an $11 \times 11$ moving window to discriminate flooded areas. Shifting in phase in HH and VV polarisations at C and L-bands may be used to discriminate flooded from non-flooded forests (Henderson & Lewis, 2008).

Grings et al. (2009) use electromagnetic models of structural and dielectric properties of soil and vegetation to predict backscatter response of marshland locations and to extract water level estimates, with some success, particularly when using interferometric techniques. Frappart et al. (2005) also use radar altimetry to gauge heights and estimate water storage volumes in the Negro basin in South America.

1.2.2.5 Ancillary data

Having established that there are a variety of parameters at play affecting how an object or area scatters or absorbs radar radiation, the inversion of the signal response necessarily involves having a qualitative and quantitative understanding of properties which cannot be extracted from the radar data. These may include the following:

**Terrain and elevation:** Governed by both the requirement for orthorectification and by the dependence of backscatter on local incidence angle. A Digital Elevation Model (DEM) is an essential component of the orthorectification process (section 1.4.1.3 refers). Further, any corrections due to local incidence angle or height calculation will necessarily require an accurate DEM. Waisurasingha et al. (2007), for example, use a DEM (unidentified) and Radarsat-derived flood extents to estimate flood depth to a high degree of accuracy. Usually an exceptional level of height accuracy is required of a DEM to derive water levels, such as that attainable using LIDAR (Töyrä & Pietroniro, 2005).

**Soil moisture:** The dielectric properties of water changing the signal response significantly, and a valuable tool in flood prediction, as discussed.
1. Literature review

**Regolith:** The dielectric properties of the surface material having a direct result on the signal attenuation.

**Soil types:** As regolith, but in addition may play a part in soil infiltration capacity and flood propensity estimates.

**Meteorological data:** For covariant analyses of flood-related events.

**In situ hydrological data:** Data collected in the field is an essential requirement for calibrating and verifying the results derived from remote sensing.

**Other:** Many other satellite data products were used in the literature researched to verify derived results or to play an integral part in class segmentation. Zhou *et al.* (2000) use multi-temporal NOAA AVHRR and Radarsat data to monitor flooding in China. Landsat data was used by Bartsch *et al.* (2008) to compare inundation mapped with GM data. Seiler *et al.* (2009) use ASTER data to compare texture analysis methods with ASAR. Töyrä & Pietroniro (2005) use multi-temporal SPOT and Radarsat data over a five year period to accurately map wetland variations. In the latter paper, the importance of spatio-temporal databases built up over time, with the understanding of relationships between vegetation patterns and flood dynamics over an entire delta, are stressed.

1.2.3 **Envisat ASAR Global Monitoring mode**

ENVISAT ASAR’s Global Monitoring mode uses ScanSAR technology, incorporating *beam steering*, to cover wide swaths with varying look angles (ESA, 2007a), and has the potential to operate continuously, making it unique amongst SAR instruments aboard satellites (Pathe *et al.*, 2009). The resulting product has an excellent temporal coverage (figure 1.1 refers), at the expense of the spatial resolution at around 1km, and a significant noise component to the signal (Sabel *et al.*, 2008). Figure 1.7 shows the swath configurations against their look angles. Table 1.3 shows the specifications for GM data.
1. Literature review

Table 1.3: ASAR Global Monitoring Mode Image product summary, taken directly from ESA’s ASAR Product Specifications (ESA, 2007b). Note that ENL refers to Equivalent Number of Looks (ESA, 2007a).

<table>
<thead>
<tr>
<th>PRODUCT NAME</th>
<th>ASAR Global Monitoring Mode Image</th>
</tr>
</thead>
<tbody>
<tr>
<td>DESCRIPTION</td>
<td>ASAR product generated from data collected when the instrument is in global monitoring mode. It is a multi-look coarse resolution image.</td>
</tr>
<tr>
<td>APPLICATIONS</td>
<td>It is for users wishing to perform applications-oriented analysis of large scale phenomena, where high resolution is not needed.</td>
</tr>
<tr>
<td>COVERAGE</td>
<td>Up to 400 km across track by up to 40000 km along track.</td>
</tr>
<tr>
<td>THROUGHPUT</td>
<td>1 product per orbit</td>
</tr>
<tr>
<td>PRODUCT SIZE</td>
<td>Stripline Max: 139 MB (80000 MDSRs × (850 samples × 2 bytes/sample + 17 bytes header)). Extracted Scene Max: 1.40 MB (800 MDSRs).</td>
</tr>
</tbody>
</table>

Continued on next page...
### Table 1.3: ASAR Global Monitoring Mode Image product summary (contd.)

<table>
<thead>
<tr>
<th><strong>Totals include all aux. data.</strong></th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>GEOMETRIC</strong></td>
</tr>
<tr>
<td><strong>SAMPLING</strong></td>
</tr>
<tr>
<td>pixel spacing 500 m by 500 m</td>
</tr>
<tr>
<td><strong>GEOMETRIC</strong></td>
</tr>
<tr>
<td><strong>RESOLUTION</strong></td>
</tr>
<tr>
<td>approximately 1000 in ground range by 1000 m in azimuth</td>
</tr>
<tr>
<td><strong>GEOMETRIC</strong></td>
</tr>
<tr>
<td><strong>ACCURACY</strong></td>
</tr>
<tr>
<td>absolute location accuracy: 1000m + orbit data error</td>
</tr>
<tr>
<td><strong>RADIOMETRIC</strong></td>
</tr>
<tr>
<td><strong>RESOLUTION</strong></td>
</tr>
<tr>
<td>Product ENL &gt; 15 (TBC); Rad. resolution = $10 \log(1 + \frac{1}{\sqrt{ENL}})$</td>
</tr>
<tr>
<td><strong>AUXILIARY</strong></td>
</tr>
<tr>
<td>Orbit State Vectors, Time correlation parameters, Main Processing</td>
</tr>
<tr>
<td><strong>DATA INCLUDED</strong></td>
</tr>
<tr>
<td>Parameters ADS, Doppler Centroid ADS, Chirp ADS, Antenna Elevation, Pattern ADS, Geolocation Grid ADS, PQS ADS.</td>
</tr>
<tr>
<td><strong>ALGORITHMS</strong></td>
</tr>
<tr>
<td><strong>USED</strong></td>
</tr>
<tr>
<td>data decompression;</td>
</tr>
<tr>
<td>raw data correction;</td>
</tr>
<tr>
<td>replica construction and power estimation;</td>
</tr>
<tr>
<td>calibration pulse processing;</td>
</tr>
<tr>
<td>antenna elevation gain function calculation;</td>
</tr>
<tr>
<td>noise power estimation;</td>
</tr>
<tr>
<td>image formation (SPECAN);</td>
</tr>
<tr>
<td>geolocation.</td>
</tr>
<tr>
<td><strong>NOTES</strong></td>
</tr>
<tr>
<td>Produced systematically from the GM Level 0 product.</td>
</tr>
<tr>
<td>The product covers a continuous area along the imaging swath.</td>
</tr>
<tr>
<td>User extracts child product of region of interest, subject to minimum scene size</td>
</tr>
</tbody>
</table>

The ASAR instrument operates at a wavelength of 5.6cm (C-band), potentially with HH or VV polarisation. All GM data covering the Australian mainland acquired to date has used HH polarisation.
1.3 Methodological challenges with radar data

1.3.1 Speckle

As with all measurement processes using coherent radiation (e.g. Pathe et al. (2009)), GM data suffers noise in the form of speckle, resulting from additive and subtractive interference processes. Although the worst of this is removed by averaging over 7–9 equivalent numbers of looks (ENL) (ESA, 2007a), further filtering is generally required. There are many tried and tested spatial filters used to mitigate the effects of speckle, possibly the most popular of which is the Lee Adaptive filter (Henderson & Lewis, 2008).

Essentially all spatial filters are identifying a pixel as speckle by some criterion based on its variation from its neighbourhood. They are then adjusting the value of this pixel, again, according to some function of the values of the neighbourhood. This function may simply be an average—though the median is often used (Bourgeau-Chavez et al., 2001), as this ensures that the value used belongs to the domain of actual values received.

1.3.2 Temporal resolution

The importance of adequate temporal frequency of data in respect of the monitoring of inundation dynamics has been touched upon in the introduction and elsewhere. Oberstadler et al. (1997) express frustration at the time taken to access data when a flood event has occurred, a sentiment agreed with by Sanyal & Lu (2004). Rosenqvist et al. (2002) consider adequate time series of data to be the “prime factor” governing the reliability and accuracy of flood models. Limitations of single frequency images are identified by Parmuchi et al. (2002), who also anticipate the emergence of greater availability of multi-temporal data and some of its possibilities.

When considering a particular region over a period of time, for which multi-temporal data of high frequency is available, there exists the opportunity to establish parameters which are very specific to each square kilometre of the region by analysing data taken at varying incidence angles and at various times under different environmental conditions.
1. Literature review

1.3.3 Spatial resolution

Most applications of radar remote sensing would benefit from data with as high a spatial resolution as possible, but this must be played off against constraints in terms of data transfer practicality, storage and processing time. Increasingly, data storage and processing time become less of an issue. This review is concerned with the use of GM data, which has a nominal spatial resolution of 1km. The constraint that this imposes is fairly straightforward, but the consequences may be more complex. Certainly, the ability to derive details about the environment within a pixel is reduced, thus increasing the requirement for more auxiliary data against which to analyse the radar signal.

The obvious limitation of coarse resolution data is the minimum scale at which classes can be distinguished. With multi-temporal data, this can be mitigated to a limited extent for structures that are stable over a long period, which can be discriminated by temporal averaging at a finer resolution—an example of this may be seen in figure 1.8, which shows a mean of many GM images over Cape York at a resolution of 16” (≈ 500 m).

However, for the purposes of analysis, the minimum size for a homogeneous class such as a river channel to be discernible will be a function of pixel size and the contrast between the backscatter value of the interfacing classes we wish to separate. The following section will have some bearing on the minimum discernible features.

1.3.4 Ambiguity of response

Simplistically, if water presents itself with a low backscatter value, and dry land with a high value then a mid-range value will tell us we have some of both within a pixel, and the actual value may give us an indication of the proportion of each. Further to this, if we know the terrain precisely, we may know the pattern adopted by the rising stage of water and in theory could even predict the extents within the pixel. Unfortunately we are not looking at a region that is made up of water and dry homogeneous land.

During a flood, in fact, the high moisture content of the surrounding soil will result in an increased backscatter (e.g. Pope et al. (1997)). The flooded area itself
may also be returning an almost lossless signal due to dihedral scattering from the interaction of water with emergent vegetation, as discussed in section 1.2.2.1. Some researchers approach the complexity introduced by vegetation by masking it out (Bonn & Dixon, 2005). However, many wetlands and floodplains are covered with aquatic vegetation, which renders this approach counter-productive.

Therefore, having detailed vegetation mapping which takes into account the type and structure of the vegetation, allowing us to predict its response, and its corresponding contribution to the received signal, is important. There are various ways to do this. Bracaglia et al. (1995) develop an electromagnetic model for the analysis of crops using radiative transfer theory and matrix doubling. Grings et al. (2009) adapt this model for a technique to monitor flooded marshes, using cylinders and discs to model different species. The role of vegetation is not a
static one, as it is dynamic in both spatial extent and in size, and the emergence of vegetation and differential leaf geometry in different seasons can lead to a different (and, in fact, opposite) response (Bartsch et al., 2009).

Data in known conditions, together with empirical data, may allow us to establish a baseline with which to analyse backscatter during periods of inundation, enabling the pixel values to be inverted and to derive a measure of the extent of inundation. Knowing the degree of effects such as dihedral scattering is important. There seems to be a minimum scale of vegetation for which the phenomenon occurs—Pope et al. (1997) tell us that during their study of marshes, this effect was only observed with C-band HH data in tall marshes with less than one third spatial coverage. Henderson & Lewis (2008) tell us of the interaction between specular reflection, volume scattering and dihedral scattering, and that the ratio between them changes as the water level varies. Martinez et al. (2001) observe peaks in backscatter from regrowth interaction with water in marshland. Bartsch et al. (2009) find that inundated areas with emergent vegetation are difficult to distinguish from areas of soil with a high moisture content.

The effect of wind on water has been discussed. Parmuchi et al. (2002) observe large variations in water response even with low wind conditions, whereas Wilson & Rashid (2005) experience near perfect (specular) reflection from water in the absence of wind. The fact that wind does effect the ability to distinguish water is a problem discussed by various researchers (Hostache et al., 2009a; Sanyal & Lu, 2004; Töyrä et al., 2001). Mitigation of this problem can, apparently, only come by recording wind conditions (and direction—equation 1.7 on page 21 refers) against all data. Additionally, we may consider our choice of incidence angle from the available data during times of strong wind conditions. Töyrä & Pietroniro (2005) point out that higher incidence angles provide data that is less subject to this effect. A good visual example of the effects of certain wind conditions on the radar response from water can be seen in Figure 3.11, later in Chapter 3.

Of course, having discussed only those ambiguities which are peculiar to radar data, we must also consider the simple fact that many different responders produce a similar brightness value or DN value for any single-band electromagnetic sensor.

It is felt that the uncertainties inherent in this evaluation of radar data should
form an integral part of the segmentation analysis when seeking the demarcation of water. Binary decisions early in the process propagate and magnify errors through to the final output. This project seeks to improve accuracy over single-image thresholds, by effectively adding data channels using displacements in respect of time and of incidence angle.

Multi-temporal data lends itself well to classification based on decision tree analysis, examples of which will be dealt with in section 1.4.2. Such an approach allows for a series of logical steps, each of which may involve different evaluation techniques.

### 1.3.5 Local incidence angle

ESA’s *ASAR Product Handbook* identifies the potential of high incidence angles and HH polarisation to map flooding, but concedes that “*Combined use of images acquired with different incidence angles poses a new set of challenges*” (ESA, 2007a).

The theoretical relationship between backscatter and local angle of incidence was covered in section 1.4.1.5. The main consideration is, in fact, that the local incidence angle is not known, and when dealing with pixel sizes of 500m, the aggregate contribution of all of the structural orientations within a pixel would be too complex to calculate, even with an accurate DEM, as the effects of radar shadow, foreshortening, layover (section 1.2.1.1 refers) and corner reflection etc. would vary in relation to each other with the changing look angle.

GM data gives us the time series capability that allows us to establish the relationship between local incidence angle and relative backscatter by observation. An example of this process using GM data is given to us by Sabel *et al.* (2008) in their investigation into the detection of surface soil moisture in Australia. In this case, the full time series was used to fit a function to the incidence angle and then the data was normalised to an angle of 30°. This was modelled on a linear fit, as described by Pathe *et al.* (2009), who demonstrated the relationship between backscatter and local incidence angle determined using a scatterometer in the Grasslands of Oklahoma—their plot is reproduced in figure 1.9. The two vertical lines in the plot highlight the region used in GM data, and demonstrate
the suitability of a linear approximation under these conditions (Pathe et al., 2009).

![Graph: Backscatter Coefficient vs. Incidence Angle](image)

**Figure 1.9:** Dry and wet backscatter reference curves for winter conditions with minimum vegetation cover and summer with maximum vegetation cover (Pathe et al., 2009).

There is an approximate 50% split between data taken from an ascending orbit and from a descending orbit. Backscatter values from an area taken from the same look angle but opposing directions may, where structural components of the backscatter are significant, bear little relationship to each other. However, the flat nature inherent in floodplains and open water make orbit direction less consequential for these cases.

The varied incidence angles used in GM data provide an opportunity to better distinguish water from land. Ahmed (2006), for example, uses local incidence angle as part of the classification process, and Töyrä & Pietroniro (2005) describe the use of the low incidence angle Radarsat S1 data with the high incidence angle S7 data to obtain a “significant increase” in the classification of open water.
Certainly the experimentation with images having a close temporal proximity but having different incidence angles is worthy of pursuit.

1.3.6 Data acquisition

One of the strengths of GM data is the fact that it is generated systematically and made available over the Internet very easily. The process requires a Category-1 fast registration with the European Space Agency through its Earth Observation Principal Investigator Portal (ESA, 2009a). Data is downloaded directly from one of two processing stations, at Kiruna in Sweden (ascending orbits) and at the ESA Centre for Earth Observation (ESRIN), at Frascati in Italy (descending orbits). Other satellite data is available from ESA and other space agencies through the submission of project proposals.

In order to make full use of the abundance of GM data available globally, it will be necessary to consider the software and hardware required to download and process a high volume of such data.

1.4 Data analysis

1.4.1 Preprocessing

1.4.1.1 Despeckling

The occurrence and cause of speckle have been discussed in section 1.3.1. When confronting its mitigation there are several options to consider. Firstly is the type of filter. In the broadest terms, and in respect specifically of GM data, averaging or sampling can take place spatially or temporally. For each of these there are then many choices of method. Whether applying a single filter, or a combination of filters, there is also sequence and staging to consider.

There are some good reasons to consider despeckling measures prior to the georectification process. The grid axes of the raw data bear a direct relationship to the azimuth direction and therefore the planar axes parallel and orthogonal to the incidence angle. If any terrain corrections due to orientation are to be taken into account, then maintaining this relationship would be a distinct advantage.
1. Literature review

Despeckling would necessarily need to take place prior to such corrections. Also, as speckle is largely the result of interference due to the coherence of the electromagnetic waves (Pathe et al., 2009), analysis using, for example, wavelets (Amirmazlaghani et al., 2009) would benefit from knowledge of the emission angles. Finally, applying despeckling measures prior to the interpolation process necessary for georectification would avoid “contaminating” any neighbouring pixels with the erroneous speckle value. Practically speaking, only spatial filters may be applied prior to georectification, as the latter process is what ties the temporal snapshots together.

Notwithstanding such arguments in favour of despeckling prior to georectification, there is also a strong reason to leave this process until after orthorectification. This is due to the fact that, as will be seen later in Section 2.2.2, terrain effects (foreshortening, for example) can cause backscatter values from several locations to be added to a single pixel, where the terrain causes the slant range distances to differ less than a critical amount, depending on the resolution. For example, using GM data, taking a central look angle of 30°, and assuming a satellite height of 800km, a difference in elevation of adjacent pixels of around 300m would cause the entire value of a pixel to be added to that of its neighbour. Lesser elevation differences incur value transfers to lesser degrees, but they occur nonetheless. Speckle filters seek to “iron out” contrasting pixel values such as would be created by this effect. Thus, if orthorectification, which reverses the terrain effect described, were carried out after despeckling, then we could end up with artificially low, or even negative, pixel values as a result.

Choice of spatial filter can be done in a testing phase. There is much information available on other comparisons of filters for different purposes, from which to draw. Henderson & Lewis (2008), for example, describe a test involving 18 spatial despeckling filters which found that the Lee Sigma filter, with three iterations of different window sizes, increased the accuracy of brackish marsh detection from 79% to 95%.

It is envisaged that the processing required to combat other distorting effects will involve raster algebra with more than one data file. This, in itself, must have an averaging effect which can only serve to diminish outlying values. Also, a strong contender as a despeckling approach, as we are intending ultimately to
produce a binary determination of whether a pixel does or does not represent flood waters, would be a post-classification clustering algorithm, which accounts for texture rather than individual value; alternatively, a simple modal filter would be appropriate.

1.4.1.2 Georeferencing

The geolocation of GM data makes use of the *Geolocation Grid* which is included within Annotation Data Set Record (ADSR) of the binary data file. This grid includes the incidence angle and a geodetic latitude and longitude for a number of data points throughout the image. These data points occur at a nominal distance of 40km in azimuth, and at 11 locations across the swath, the spacing of which depends on the incidence angle (ESA, 2007a). Typical root-mean-square (RMS) errors, in the third-order polynomial fit used in the georeferencing process during this research, were calculated at around 0.16 pixels ($\approx 80m$). The corresponding incidence angles, which are the calculated incidence angles with respect to the WGS84 ellipsoid (ESA, 2004), are required for further processing.

1.4.1.3 Orthorectification

Orthorectification serves to minimise the geometric distortions in an image caused by terrain. Methods specific to ScanSAR images are documented (Low & Mauser, 2003). The fact that the terrain under investigation (flood plains) is usually flat means that relative distortion and therefore displacement will be minimum. Bartsch et al. (2009) orthorectify GM data over flat terrain using GTOPO30 data. Open source software, the *NEXT ESA SAR Toolbox*, which is being developed on ESA’s behalf (Array S. C. Inc., 2009), will orthorectify GM data (amongst many others), and will download SRTM3 DEM data automatically in the process. The software also includes speckle filtering and many analysis tools.

In the case where it is intended to download and pre-process all GM data systematically, benefits of an alternative means to orthorectify must be considered. In order for data to be registered, as will be discussed in the next section, bespoke scripts are used to extract header information from the data files. Querying the data files in this way using an abstraction layer from the Geospatial Data
Abstraction Library (GDAL)\textsuperscript{1}, gives us all the information we need to carry out the orthorectification ourselves, giving us control over the whole process. Using open-source programming languages such as Perl\textsuperscript{2} or C\textsuperscript{++}\textsuperscript{3}, could allow us to run our scripts on the university’s High Performance Computer (HPC), allowing us to take advantage of high-speed parallel processing capabilities. To this end, the scripts will need to be written with multi-threading in mind.

1.4.1.4 Database registration

The binary GM data files contain a header in text format. From this header, much salient details of the file can be extracted prior to any further processing, and fed through to a database. This data includes corner coordinates which allow the image outlines to be mapped without any georectification (which were used, for example, for the production of figure 1.1 on page 13), and this, together with other information, allow the efficient selection of data for further analysis prior to full processing. The data file may then be temporally co-registered with other data such as wind, rainfall and other remotely sensed data to build an environmental context for analysis.

1.4.1.5 Sigma-nought versus beta-nought

Most satellite radar data comes with DN values which are a function of radar brightness $\beta$. Conversion of GM data to $\beta$ involves squaring the DN and dividing by a calibration constant. Most analysis of radar data converts the raw DN values to the coefficient of backscatter, $\sigma^0$, which involves further multiplication by the sine of the local incidence angle of the transmitted signal to the surface reflecting it. This common approach is questioned (and answered) by Raney \textit{et al.} (1994) in their article entitled “Plea for radar brightness”, and by others (David \textit{et al.}, 1998). By way of clarification, it is first necessary to examine exactly what the GM data pixel values (DN) represent.

The ASAR sensor on board the satellite is able to measure, in raw terms, power. Knowing the angle from which the power is received and the distance it

\begin{itemize}
\item \textsuperscript{1}http://www.gdal.org
\item \textsuperscript{2}http://www.perl.org
\item \textsuperscript{3}http://www.cplusplus.com
\end{itemize}

41
has travelled, the system uses an algorithm to estimate the incidence angle (based 
on the WGS84 ellipsoid) and attribute this power to an area, giving a measure of 
power per unit area (called, in most fields, *intensity*, though this term has some 
ambiguity, as shall be seen). Knowing the dimensions of the receiving antenna 
and (approximately) the relative displacements of the satellite and the scatterer, 
allows the determination of the solid angle over which the backscattered radiation 
is received, giving a measure of the power per unit area (of the backscatterer) 
per unit solid angle (or steradian). This, in radar terms, is also sometimes called 
intensity (Woodhouse, 2006), but from here on shall be called *brightness*, which 
has units $Wm^{-2}sr^{-1}$.

As mentioned, most articles using radar remote sensing for some sort of clas-
sification refer to backscatter values in terms of the *normalised radar backscatter 
coefficient*, $\sigma^0$, usually converted to decibels. This is due to the fact that, in order 
to describe the backscatter response of a feature, a quantity which is independent 
of the sensor is sought. $\sigma^0$ is derived from the *radar cross-section*, $\sigma$, where

$$\sigma = \frac{I_{\text{received}}}{I_{\text{incident}}} \cdot 4\pi R^2 \quad (1.8)$$

where, $I$ is the intensity in the classical sense of power per unit area (Woodhouse, 
2006). $\sigma$ has units $m^2$. However, as a greater incidence angle $\theta$ causes the power 
to be spread over a larger area, $\sigma$ is dependent on $\theta$. This gives rise to $\sigma^0$, which 
is the radar cross-section per unit ground area, or

$$\sigma^0 = \frac{\sigma}{A} \quad (1.9)$$

(Woodhouse, 2006). This makes $\sigma^0$ dimensionless. ESA’s ASAR calibration 
guidelines (ESA, 2004) tell us that $DN$, $\beta^0$ and $\sigma^0$ are related by

$$DN^2 = C \cdot \beta^0 = K \cdot \frac{\sigma^0}{\sin \alpha} \quad (1.10)$$

where $C$ is some constant and $K$ is known as the *absolute calibration constant*. 
$K$ is described as being

“...processor and product type dependent, and might change be-
between different beams for the same product type” (ESA, 2004).

For the GM data in this study, it has been found that $K$ has had the same value throughout ($2.19 \times 10^{7}$), although the preprocessing script has been designed to extract the value from each individual file during every calibration.

For a volume scatterer, as $\alpha$ increases, the incident power is distributed over a larger area (Woodhouse, 2006), making $\sigma^0$ a function of $\cos \alpha$. To take this into account, it therefore may become more convenient to talk in terms of $\gamma$ (or the range independent backscatter coefficient (Shimada, 2010a)), such that

$$\gamma = \frac{\sigma^0}{\cos \alpha} \quad (1.11)$$

However, it must be understood (and indeed it is pointed out in the guidelines (ESA, 2004)) that $\alpha$ is the local incidence angle, which is not usually known. In fact, the derivations of $\gamma$ and $\sigma^0$ above assume volume scatterers and a flattish, smooth surface (based on the WGS84 ellipsoid). Water, being largely a specular reflector as discussed, and wet surfaces, are more greatly effected by varying incidence angle, as shall be seen. Also, as pointed out by Raney et al. (1994) and David et al. (1998), the actual incidence angle is not one angle, but rather a complex quality made up of all of the orientations of structural components within a pixel. This is especially true when considering what was discussed in section 1.3.5, that scatterers will respond differently in an ascending orbit and a descending one due to irregular structural orientation (effectively making the real incidence angles different, where the nominal ones are the same).

Therefore the DN (or perhaps $\beta^0$, simply reversing the output scaling operation (Parmuchi et al., 2002)), should be taken on its own merits, and processed according to the intention of the analysis, rather than being automatically converted to an erroneous $\sigma^0$.

In fact, the effects of incidence angle are greater than the gamma conversion in equation 1.11 accounts for. This fact is illustrated in figure 1.10.

Baghdadi et al. (2001), in their evaluation of C-band SAR data for mapping
1. Literature review

Figure 1.10: Two GM data images of the Mitchell catchment using GM data. The first has undergone a conversion to $\sigma^0$ using equation 1.10, and the second to $\gamma$ using equation 1.11. It is clear that the effects of the geodetic “local” incidence angles have not yet been fully factored out, as viewed by the tendency for higher values (seen here in red) on one side of the swath.

Wetlands, normalise $\sigma^0$ using a function $F(\alpha)$ where

$$\sigma_n^0 = \frac{\sigma_\alpha^0}{F(\alpha)}$$

(1.12)

where $F(\alpha)$ is the mean angular dependence for a given date and given polarisation. This is shown as a cosine function, which for HH typically $\approx \cos^{2.5} \alpha$.

Baghdadi et al. (2001) derive this function by assuming the relationship

$$\sigma_\alpha^0 \approx \sigma_n^0 \cos \psi \alpha$$

(1.13)

(adapted from Ulaby et al. (1982)), where $\psi$ is the slope of the linear fit of $\sigma^0$(dB) against $10 \log(\cos \alpha)$ (Baghdadi et al., 2001).

Pathe et al. (2009) model radar backscatter as a function of incidence angle, soil moisture and time, where

$$\sigma^0(\theta, t) = \sigma_{\text{dry}}^0(30) + \beta(\theta - 30) + S m_s(t)$$

(1.14)

where $m_s$ is the relative soil moisture content and $S(t)$ the sensitivity of the radar backscatter to relative soil moisture changes at time $t$. Note that Pathe et al. (2009) assume a linear change with respect to incidence angle $\theta$, which varies only
1. Literature review

between 20° and 40°. Sensitivity $S$ for each 30’ tile is estimated as

$$S = \sigma_{wet}^0(30) - \sigma_{dry}^0(30)$$  \hspace{1cm} (1.15)

Wet and dry references are established through time series analysis.

The research following this review will seek to fit a function similar to that in equation 1.13 to observed values in the time series. As a single region is being examined, it is felt that rather then attempting to fit one parameter across the (heterogeneous) area, parameters specific to subregions of a few pixels can be mapped and applied more discriminatingly.

Further, much analysis will be on the basis of change detection, in which case the operative values are relative to a subregion or pixel, and normalising becomes unnecessary. Here even using the linear DN values are an option, eliminating some steps in the preprocessing of data, an approach adopted by, for example, Rosenqvist et al. (2002) when modelling inundation patterns in the Amazon, and as advocated by Raney et al. (1994) and David et al. (1998).

1.4.2 Classification

Radar images, with essentially a single band, do not allow the same comparison of signatures of different responders at different wavelengths possible with optical and multi-spectral images. However, there are other comparisons which may be made in the same way by choosing different modes of image as channels in place of wave bands. The ASAR product Alternating Polarisation (AP) modes, for example, offer two images taken at the same time, being one of the following combinations: HH–VV, HH–HV, VV–VH (ESA, 2009c). There are techniques to classify single radar images. The most simple of these in the discrimination of water from land is a threshold. Abhyankar et al. (2007), for example, recommend a simple threshold of -15dB for this purpose, but point out the problems already discussed, where the presence of waves on the water causes an increase in backscatter value. Hostache et al. (2009a) take a less deterministic, more probabilistic approach by classifying in terms of likelihood of values representing flooding. A minimum threshold is established by choosing the lowest backscatter value found in the definite non-flooded regions. Below this threshold, it is as-
1. Literature review

assumed that values represent flooding. A maximum threshold is then established by selecting the highest values observed in the non-flooded water bodies (lakes, permanent river channels, etc.). Values above this threshold are assumed to represent non-flooded areas. Values between the thresholds are then classified as potentially flooded (Parmuchi et al., 2002). Outliers would need to be taken into consideration with this approach, and in the case of strong winds, for example, it might be possible for most of the flooded areas to fall only within the “potentially flooded” classification, making the exercise of limited use.

A common method for classifying single radar images is to take into account textural relationships between pixels. The Grey Level Co-occurrence Matrix model (GLCM), for example, is a two-dimensional histogram of grey levels for a pair of pixels which are separated by a fixed spatial relationship (Antoniol et al., 2005). Various different textural measures are calculated from this histogram, and depend on the choice of window size and pixel displacement. These include entropy, a measure of the similarity of values within the local window, correlation, a measure of linear dependency of values of neighbouring pixels, and angular second moment, a measure of local homogeneity (Antoniol et al., 2005).

Seiler et al. (2009) use variance and entropy from an ASAR Image Mode product, together with a Normalised Difference Vegetation Index (NDVI) and Normalised Difference Water Index (NDWI) using ASTER data, to good effect, in deriving eight classes around a flood plain. The improvements found by introducing textural components to the segmentation process are reproduced in figure 1.11.

Töyrä & Pietroniro (2005) note that their GLCM texture analysis increased the accuracy of their classification from 22% with a single Landsat image to 74%, based on a Kappa coefficient ($\kappa$) test (section 1.4.4 refers). Arzandeh & Wang (2002) use outputs from a GLCM analysis as channels for the supervised classification of a single Radarsat image. An interesting comparison of the use of different window sizes and the resulting accuracy was carried out (see figure 1.12). Whilst many categories such as swamps and forests seem to be more accurately identified with larger window sizes, water remains with a reasonably constant (and high) level of accuracy, which seems to benefit from smaller window sizes (Arzandeh & Wang, 2002).
One classification method which exploits the use of a spectral class model known as a Gaussian mixture distribution is the Sequential Maximum A Posteriori (SMAP) method. The model works by segmenting the image at various scales or resolutions and using the course scale segmentations to guide the finer scale segmentations (Bouman & Shapiro, 2004), and is available with some remote sensing software such as GRASS (GRASS Development Team, 2009).

Grandi et al. (2009) use wavelet variance analysis to derive textural signatures from radar data, and in particular use the Fischer Criterion (Grandi et al., 2009) to demonstrate the feasibility of such a tool as a supervised classifier. Among the five methods described by Baldassarre et al. (2009) is an active contour model, which uses a region growing algorithm to minimise an energy function which seeks to encircle as many “good” pixels as possible.

The segmentation processes discussed are pertinent to a single radar image. With the benefit of a high frequency time series, classification methods can take into account backscatter responses of an area over a range of environmental conditions and incidence angles, giving a more complex choice of signatures to analyse.
1. Literature review

Figure 1.12: Influence of window size on the average overall accuracy of the classification results of texture features for different land cover types (Arzandeh & Wang, 2002)

This offers a much greater potential for accurate classification of a single image, as well as behavioural classification of the whole series. A number of approaches to dealing with this complexity present themselves in the available literature, two of which are Artificial Neural Networking (ANN) (Ahmed, 2006; Quan et al., 2008) and Decision Tree Analysis (DTA) (Baghdadi et al., 2001; Parmuchi et al., 2002). In some cases both approaches are adopted, either together, or for comparison (Zhou et al., 2000).

A good example of a decision tree based on individual thresholds in a hierarchical step-wise process, is reproduced in figure 1.13. It is conceivable that certain decisions may be made, not only on the basis of a cross-image threshold, but also on some pre-classification process, that effectively allows a different threshold for a different geographical location. Also, the threshold can be made dependent on incidence angle, and orbit direction. Any analysis method can be contained within such a decision tree, including an index, for example, such as that combining the response to high and low incidence angles, as used by Töyrä & Pietroniro (2005) and discussed in section 1.3.5. Also discussed in this section was the likely difference between two backscatter signals representing the same
1. Literature review

Figure 1.13: Decision tree design from multi-temporal RADARSAT SAR images (Parmuchi et al., 2002)
area with the same angle of incidence, but with opposing orbit directions, due to structural asymmetry. One responder that is structurally closest to the ellipsoid estimation, and thus would suffer less from this effect, is water.

Ahmed (2006) provides a comparison of accuracies for a $K$ nearest neighbour analysis and that of an artificial neural network (ANN) in the classification of land cover, including water. Overall it is determined that the ANN provides the most robust results, with water being well distinguishable using both techniques. Interestingly, temporal comparisons are made only with like incidence angles.

Bourgeau-Chavez et al. (2001) describes a semi-automatic hierarchical classification process that uses a cost reduction algorithm to produce an optimal model for classification of, in this case, SIR-C data, in which an accuracy of $\approx 90\%$ is measured for the open water classification.

Oberstadler et al. (1997) use the Evidence-Based Interpretation of Satellite images (EBIS) algorithm to perform a supervised classification of water and non-water classes. The classifier is based on the “Dempster–Shafer theory”, assigning a pixel to a class based on the relative probability of the value belonging to that class, the probabilities being established for all values in the initial training (or parametric) stage, using a multinominal distribution model. The conclusions which were relevant to flood mapping emphasised the fact that the resolution of the ERS-1 data was too coarse for the purpose of flood mapping in Germany, and that temporal frequency was insufficient.

Noernberg et al. (1999) use coefficients of variation for different polarisation configurations to discriminate classes, and find open water easily distinguishable in all cases—this is, however, using C-band to a resolution of 6m.

Quan et al. (2008) combine Probabilistic Neural Network (PNN) with Multi-scale Auto Regressive (MAR) models to extract multi-scale features of SAR images. The MAR is used to train the PNN. Interestingly, the MAR is a borrowed time-series analysis technique, used here for spatial analysis of the same image. The scale recursion device used in this algorithm seems akin to the SMAP model discussed above. This combination of techniques is interesting, as the scaling method can identify and account for speckle; the method appears to produce particularly good results from homogeneous areas returning heterogeneous values (see figure 1.14).
Figure 1.14: (Top) Original SAR image, (left) Segmented image obtained using PNN algorithm and (right) Segmented image obtained using MAR and PNN (Quan et al., 2008)
Parmuchi et al. (2002) conclude their research with two findings which fit well with the aims and methodology of the research proposed in this review:

“(1) the need for multi-temporal SAR data acquired under different environmental conditions for mapping wetlands, and (2) the advantages and flexibility of physically based reasoning classifiers for synthetic aperture radar (SAR) data classification.” (Parmuchi et al., 2002)

1.4.3 Output

Figure 1.15 shows a basic schematic representation of the processing and output of GM data proposed. The normalisation function $F(DN, K, r, \alpha)$ represents the function derived from the time series data describing the relationship between orbit direction, computed local incidence angle $\alpha$ and the backscatter, together with the reversal of the scaling operation at the sensor, being a function of $\alpha$ and the absolute calibration constant $K$, as discussed in section 1.4.1.5. Final output might be, for example, single flood extent maps, relative soil moisture maps, or flood duration maps (Rosenqvist et al., 2002). An important compliment to such data will be a measure of their accuracy.

1.4.4 Analysis of accuracy

Each stage of processing changes values or assigns categories based on assumptions, and the magnitude of error of each stage should be considered independently.

The GM data itself is understood to have a radiometric accuracy of 1.2dB (Pathe et al., 2009; Sabel et al., 2008). As stated previously, georectification and orthorectification are expected to contribute low spatial errors in the case of GM data in flood plains (see sections 1.4.1.2 and 1.4.1.3).

For automated classification processes involving training areas, pixels within the training areas may be randomly assigned to either training or testing, following classification (Bourgeau-Chavez et al., 2001). A common method of analysis against a known accurate classification is the use of error matrices, where differ-
1. Literature review

**Figure 1.15:** Flow diagram outlining one possible processing sequence using methods discussed.
ences are quantified by the kappa coefficient $\kappa$ (Ahmed, 2006; Arzandeh & Wang, 2002; Töyrä et al., 2001; Waisurasingha et al., 2007), simplified by the estimation $\hat{\kappa}$ as follows (Campbell, 2007):

$$
\hat{\kappa} = \frac{\text{observed} - \text{expected}}{1 - \text{expected}} \quad (1.16)
$$

As this represents a sample, some researchers go on to test the statistical significance of the results using a $z$-test (Waisurasingha et al., 2007). In the case of the mapping of inundation (a binary decision), accuracy can be stated in overall terms, or separately in respect of errors of omission ($\epsilon_o$) and commission ($\epsilon_c$) (Rosenqvist et al., 2002).

### 1.5 Gaps in existing research

**Computing power to manage multi-temporal time series analysis**

Much research has been done on river systems using radar remote sensing (Bonn & Dixon, 2005; Chandran et al., 2006; Frappart et al., 2005; Waisurasingha et al., 2007; Wilson & Rashid, 2005). However, *multi-temporal* analysis has been confined to a handful of images, perhaps taken in different seasons. Where data needs to be ordered in advance, it is often done post facto, and it is highly unlikely that a pre-flood image will be obtained unless the flood itself is highly predictable. GM data affords us the opportunity to analyse data from a high-frequency time series with near-global coverage, from which data may be drawn hours after a flood event.

The challenges which face any attempt to separate flood water from unflooded areas, as described, are amplified when dealing with such a large dataset, particularly in terms of computing constraints. Ever-increasing advances in data storage capacity and processing speed make the task of managing the computation required to pre-process and analyse GM data feasible. It is, however, a necessary prerequisite for such analysis to set up a robust, automated system that can download, register and pre-process GM data to suit the needs of this
research, and that can, preferably, make full use of the HPC distributed network available to researchers at James Cook University. This involves choices of software and languages which are sympathetic to the HPC server environment, with which tasks may be multi-threaded and assigned to some of several hundred processing nodes. Apart from the orthorectification and geolocation operations as described, consideration must be given to our ability to customise the process, to drive the system towards different types of output, depending on the operation being performed. In certain instances, for example, it may not be necessary for the processing of a single data file to result in an individual image output, where the information gleaned from the data is merely required to contribute to an aggregate. In such cases, there exists the opportunity to save on input/output procedures and on memory, making the overall calculation process more efficient. It seems clear that the research to follow would benefit from (and, indeed, may prove to be conditional upon) a bespoke pre-processing set-up which concedes control of each part of the process for the sake of flexibility.

**Pixel-specific normalisation for incidence angle**

The broad-brush standard for “normalising” for incidence angle effects which are commonly adopted, have been discussed. The limitations in relying on differential measurements across a swath to apply linear corrections to the radiometric value of a pixel include the potential loss of differences due to flooding. The large time series available with GM data allows us to test for localised parameters against previously described models, to use in normalising for incidence angle. This involves calculations for local incidence angle based on a DEM and on satellite coordinates extracted from the data file, and on regression analysis of backscatter against local incidence angle (or functions of these) over the time series. Such calculations therefore build upon the ground-work done in setting up the computational system described above.

The resultant measure of the behaviour of a target area with respect to incidence angle can be correlated with various environmental parameters such as ground cover, regolith and soil types in order to determine the drivers behind particular radar signatures.
Limitations in the use of radar remote sensing due to land cover

Usually, the ground in the region surrounding flood waters has been saturated, by surface flow, by direct rainfall or by flood waters themselves where the level has dropped. For this reason, the backscatter values seen in these areas are usually very high, as characterised by wet soil, contrasting well with the expected low return values from open water. As discussed in Section 1.3.4, a similar high backscatter response is received from areas where vertical structures, such as tree trunks or dense grasses, protrude through the surface of the water, causing dihedral backscattering. For a particular radar data product, at its spatial frequency, wavelength and at a particular incidence angle, there must be environmental conditions such that, as flooding progresses over a certain proportion of a pixel, there will never be sufficient open water to drop the aggregate pixel value significantly enough to detect flooding. This means that there are land cover conditions for which, for example, C-band radar data with 500m pixels could not be used to monitor flooding under (almost) any circumstances. Given the overall objective of this thesis, an understanding of these conditions, in respect of GM data needs to be achieved.

Ambiguity of low backscatter response due to absorption

It may be for the same reason as described above, i.e. that flood waters are usually surrounded by wet soil, that ways to overcome ambiguity due to radar attenuation and absorption by dry soil or sand have not been investigated to any significant degree. The perception of the problem may not have any prominence in the field, as it may not be expected to be a problem at all. However, in a country such as Australia, where we see how heavy rainfall in Queensland causes surface flows between two deserts to flood the Eyre basin in South Australia, hundreds of kilometres away, we can perhaps see this phenomenon more clearly as a potential problem. This is, of course, a potential issue for all rivers running through arid regions, and the question as to whether the advantages brought by radar remote sensing to the mapping of floods can be applied to these cases, in light of this problem, deserves investigation.
1. Literature review

1.6 Conclusion

The mapping of inundation using GM data seems, given certain constraints, feasible. The difficulties faced by this task are worth attempting to overcome, due to the benefits that GM data can offer. Indeed, GM data may provide the only means of monitoring large scale inundation and recession which occur very quickly, in hydrological systems which need, increasingly, to be understood. The challenges are complex. Firstly, given the coarse spatial resolution of GM data, its use is limited to large flood plains with continuous coverage of a few square kilometres at least. In fact, it is the large, flat, quickly-changing flood plains that benefit most from the use of GM data, due to difficult timely access to gather data from the ground. Secondly, it is important that the effects on radar response of parameters such as local terrain, vegetation, soil moisture and incidence angle are understood, so as to be able to interpret the radar data to a known and sufficient degree of accuracy. It is envisaged that the opportunity to compare signal responses from an area using two or more incidence angles in very close temporal proximity will allow the segmentation of water to a degree not possible from a single angle. Also, there are opportunities to take advantage of the high temporal frequency of GM data to explore different segmentation methods beyond simple thresholding of single radar images. More specifically, the following questions arise from this review:

1.6.1 Research questions

1. How does radar backscatter vary with incidence angle for different surface conditions? How does this affect the segmentation of open water?

2. Under what vegetation conditions (vegetation type, size, density, orientation) does multihedral backscatter distort the radar signal so as to make flood water indistinguishable from its surroundings?

3. How can we separate dry soil/sand from floods through arid regions? How significant is this problem?

4. How can the processing of such a large dataset be managed? How can we
automate the download, registration and orthorectification of a high volume of GM data files to allow efficient analysis?

These questions represent the primary thrust of what is to follow. If we can answer them to some degree of success, then the contribution of this thesis will be the demonstration that GM data, and its foreseen successors, can be counted as valuable tools with which to map large floods, crucially within a time-frame in which immediate action may be taken to assist the many people whose lives may be affected. Any contribution to such a cause, however small, must surely be significant.
Chapter 2

Practical high volume preprocessing of Envisat ASAR Global Monitoring Mode data over a distributed network

Chapter context
2. High-volume preprocessing

Abstract

A means to pre-process a large volume of Envisat ASAR Global Monitoring Mode (GM) radar data, using open source software and abstraction layers, is described. The preprocessing includes orthorectification, in which case the Digital Elevation Model (DEM) is projected into the local frame of reference of the data file and header parameters are extracted, to calculate the local incidence angle and to make along-swath adjustments against terrain displacement effects.

The objective is to develop a practical method to pre-process GM data using the tools available on James Cook University’s High Performance Computing facility. The method allows for control throughout the whole preprocessing stage, so that adjustments can be made early on to eliminate unnecessary calculations in cases where, for example, parameters may be factored out, to avoid undue overhead. Spatial accuracy tests well against alternative software. Tools to extract and interpolate tie-point coordinates, incidence angles and slant-range times into Grass ASCII data files from GM data files will be made available for contribution to the GRASS GIS development community.

2.1 Introduction

From the inception of Geographic Information Systems several decades ago, advances in GIS and remote sensing instrumentation and methodologies have run concurrently with increasing computation and data storage capabilities required to take advantage of the expanding field. Satellite radar data has been widely recognised as having an important role in the application of remote sensing, due largely to its independence from solar radiation and for its abilities to penetrate cloud cover (Alsdorf et al., 2007; Badji & Dautrebande, 1997; Leblanc et al., 2011). Within the Radar field, ScanSAR technology has allowed increased temporal frequency of radar coverage, enabling its use to track rapid events such as floods, where the use of other data is precluded by the presence of cloud (Baup et al., 2007), and has provided temporally denser time series data sets, increasing the scope of methods available for, for example, temporal speckle filters (Ciuc et al., 2001; Trouvé et al., 2003), and change detection (Colesanti & Wasowski,
2. High-volume preprocessing

This increase in data volume requires increased computing capacity to meet preprocessing requirements. Radar data, for example, commonly requires georectification, terrain correction and usually some form of recoding or normalisation, together with some form of spatial filtering to reduce effects such as speckle. Prior to the recent demise of Envisat (see Section 6.2.1 towards the end of this thesis), GM data was made available in near real-time for download to parties with at least a Category-1 fast-track agreement with ESA. The data was available quickly because it was preprocessed at the sensor, before being transmitted down to one of two ground stations in Europe. This was made possible by keeping file sizes and processing requirements low, by using a coarse resolution (pixel size: 500m, spatial resolution: 1km) (ESA, 2007b). The main use to which GM data was perceived to be put was the monitoring of large-scale changes in polar ice extents, with ice and open water having readily distinguishable backscatter characteristics (Zink et al., 2001), but the frequency of coverage of GM data (see Table 2.1) extend the possibilities to other variables such as soil moisture (Pathe et al., 2009) and flood extents (O’Grady et al., 2011).

<table>
<thead>
<tr>
<th>Latitude (+/-)</th>
<th>0°</th>
<th>45°</th>
<th>60°</th>
<th>70°</th>
</tr>
</thead>
<tbody>
<tr>
<td>Frequency</td>
<td>5</td>
<td>7</td>
<td>11</td>
<td>16</td>
</tr>
</tbody>
</table>

Table 2.1: Average revisit frequency capability per 35-day orbit cycle as a function of latitude, for descending orbit path only (ESA, 2007a)

There are a couple of software tools available to preprocess GM data, such as NEST (Next ESA SAR Toolbox) and BEAM (Basic ERS & Envisat (A) ATSR and Meris Toolbox) as made available for download by ESA via their website\(^1\). Partly in recognition of the access and computing requirements for the near real-time processing of large data sets, ESA has developed a web-accessible grid environment called Earth Observation Grid Processing On Demand (G-POD), which gives fast access to data and computing resources and allows verification of algorithms (Cossu et al., 2009). However, as our analysis requires a greater control over the georectification and terrain correction process, and as a high volume of data is required to be transferred, it was considered beneficial to write bespoke

---

\(^1\)https://earth.esa.int/web/guest/pi-community/toolboxes
code to handle the process (e.g. Rees & Steel (2001)). This chapter outlines a practical method of preprocessing large volumes of GM data, employing parallel processing over James Cook University’s (JCU) high performance network using open-source software and part procedural, part object-oriented code.

2.2 Theory

2.2.1 Theoretical basis

The adopted algorithms adjust for range shifts and ignore azimuth shifts. The consequences of this are tested against the results of tools provided by ESA, using data files covering areas of mixed topography. The algorithms adjust for cross-track displacement caused by the fact that the process used by the ASAR sensor to geo-locate the source of a backscatter value assumes that the source lies on the surface of the WGS84 ellipsoid, without taking into account its height above the datum. The algorithm therefore requires an elevation for each pixel, and for this purpose, a SRTM 250m DEM (Jarvis et al., 2008) is resampled to the 500m pixel spacing of the GM data. In the raw GM data file, the position of a pixel within a row corresponds to its position across the ground range swath. For this reason, the processing is done in the local x-y grid of the raw data prior to rectification to geographical coordinates. In order to convert the raw GM data values into backscatter values, the local incidence angle $\alpha$ is required, which in turn requires the (nominal) incidence angle $\theta$ and the local orientation of the target pixel with respect to the incident radar beam. This latter parameter is established using the DEM values adjacent to the target pixel.

2.2.2 Orthorectification

2.2.2.1 Establishing the displacement error

This algorithm works within the frame of reference of the raw data, whose rows and columns lie parallel to the swath and azimuth respectively. Consider Figure 2.1, where $R$ represents the distance of the sensor from the centre of the WGS84 ellipsoid, $r$ is the radius of the ellipsoid at the target coordinates, $h$ is height.
2. High-volume preprocessing

Figure 2.1: Geometry used in the calculations for orthorectification

above WGS84 of the target, (i.e. the DEM value) and $s$ is the slant range, or the direct distance from the target to the sensor.

The target pixel is positioned in the data according to its distance along the ground-range swath as calculated using the slant range $s$. This calculation assumes a target positioned precisely on the WGS84 ellipsoid. Where the target has height $h$, it can be seen that the calculated distance has a displacement $\Delta$ closer to the nadir than the actual target position. In Figure 2.1, two triangles can be seen for which the length of all of the sides are known. From this, we can calculate angles $\Omega$ and $\omega$, which in turn gives us the displacement arc $\Delta$.

$$\Delta = r(\Omega - \omega) \quad (2.1)$$

where $\Omega$ and $\omega$ are in radians. This assumes that $r$ is constant over the displacement $\Delta$, despite the ellipsoidal, rather than spherical model. This is perfectly acceptable, as the difference would only amount to a few millimetres over an entire pixel width (e.g. Rees & Steel (2001)). Angles $\Omega$ and $\omega$ are calculated as follows.
2. High-volume preprocessing

According to the cosine rule, in any euclidean triangle,

\[ a^2 = b^2 + c^2 - 2bc \cos A \]

where \( A \) is the angle subtended by the side of length \( a \), and \( b \) and \( c \) are the lengths of the other two sides.

Rearranging gives us

\[ A = \cos^{-1} \left( \frac{b^2 + c^2 - a^2}{2bc} \right) \]

If we equate \( A \) with \( \Omega \), then

\[ \Omega = \cos^{-1} \left( \frac{(r + h)^2 + R^2 - s^2}{2R(r + h)} \right) \]  \hspace{1cm} (2.2)

and similarly for \( \omega \)

\[ \omega = \cos^{-1} \left( \frac{r^2 + R^2 - s^2}{2Rr} \right) \]  \hspace{1cm} (2.3)

Equations 2.2 and 2.3 may then be substituted into Equation 2.1 to obtain the displacement \( \Delta \).

2.2.2.2 Redistribution of cell value

Consider the cell (the yellow cell in Figure 2.2) in column (or range address) \( C \)
2. High-volume preprocessing

of our desired terrain-corrected final image, whose terrain correction displacement
is $\Delta$. Due to the difference in height between the datum ellipsoid and the target,
the value from this cell has been wrongly assigned, partly to the cell at $C - n$,
and the remainder to the cell at $C - n - 1$. Let $\eta$ be the number of cells within
displacement $\Delta$, where

$$\eta = \frac{\Delta}{P} \quad \text{and} \quad n = \lfloor \eta \rfloor \quad (2.4)$$

We cannot simply reassign the values from the cells in columns $C - n$ and $C - n - 1$, as their overall brightness value may be the sum of contributions of values
from more than one cell experiencing foreshortening or layover. It is therefore
necessary first to establish how many contributions to each cell value were made,
and then to determine what proportion of the overall value was contributed by the
target cell. We cannot determine the precise contribution from various locations
along the swath to an aggregated value. The best we can do is to give equal
weight to each contribution, effectively redistributing a contribution average to
each correct location. In practice, this requires a two-pass algorithm. Let $K_i$ be
the total number of cells contributing to the final value of the cell in column $i$. Then, given Equation 2.4:

$$K_{C - n} \leftarrow K_{C - n} + 1 - \eta + n$$
$$K_{C - n - 1} \leftarrow K_{C - n - 1} + \eta - n$$

The second pass of the algorithm may then assign a new value $V_i$ to the cell
in column $i$, such that

$$V_C = (\eta - n) \cdot \frac{U_{C - n - 1}}{K_{C - n - 1}} + (1 - \eta + n) \cdot \frac{U_{C - n}}{K_{C - n}} \quad (2.5)$$

where $U_i$ is the original value of the cell in column $i$.

If desired, we can record the $K$ values and output a raster to give a measure
of the degree of foreshortening encountered. Whether layover has occurred is
trivial, as it is simply a matter of the relative value of the displacement $\Delta$ of a
cell with respect to that of its neighbours. What is of far greater importance is
that no contributions are allocated to cells from which no backscatter has reached
the sensor due to radar shadow.

2.2.2.3 Establishing Radar Shadow

With reference to the diagram in Figure 2.3, if $\theta$ is the incidence angle of the
target at the WGS84 ellipsoid, then $\phi$ is the incidence angle at the target height,
or the Off-nadir angle, being the angle between the zero-Doppler slant range $S$
and the normal to the ellipsoid at the target pixel, in the plane containing the
sensor, the target and the ellipsoid center. Let $\phi_i$ be the Off-nadir angle for a cell
with range address $i$, where $i$ is zero closest to the nadir. If there exists a range
address $j$ closer to the nadir for which $\phi_j > \phi_i$, then $i$ is obscured by $j$ and is
therefore in shadow:

$$\text{Shadow} = \{ i \mid \exists j \ni (i > j) \land (\phi_i < \phi_j) \}$$ (2.6)

2.2.3 Local incidence angle

Figure 2.4 shows the signing convention used here. The X and Z axes lie in the
plane containing the sensor, target and swath. The Y axis comes out of the
page. $\Lambda$ and $\theta$ are the Look angle and Incidence angle respectively. $\hat{n}$ is the unit
vector normal to the target surface, and $\hat{r}$ of incident radiation returning to the
2. High-volume preprocessing

Figure 2.4: Signing convention for the following calculations. The X axis lies on the plane containing the satellite and the target, and is tangential to the WGS84 ellipsoid. Λ, θ and α are the look angle, incidence angle and local incidence angle respectively. The direction of \( \hat{n} \) is dependent on topography, and is arbitrary in this diagram.

sensor. Using this convention, the X and Z components of \( \hat{r} \) are \(-\sin \theta\) and \(\cos \theta\) respectively:

\[
\hat{r} = \begin{pmatrix}
-\sin \theta \\
0 \\
\cos \theta
\end{pmatrix}
\] (2.7)

We are interested in finding the local incidence angle \( \alpha \), which is the angle between \( \hat{n} \) and \( \hat{r} \).

Let \( \hat{n} = \begin{pmatrix} A \\ B \\ C \end{pmatrix} \)

We know that \( \mathbf{a} \cdot \mathbf{b} = ab \cos \alpha \) where \( \alpha \) is the angle between vectors \( \mathbf{a} \) and \( \mathbf{b} \).

\[
\Rightarrow \hat{n} \cdot \hat{r} = \begin{pmatrix}
-\sin \theta \\
0 \\
\cos \theta
\end{pmatrix} \cdot \begin{pmatrix} A \\ B \\ C \end{pmatrix} = C \cos \theta - A \sin \theta
\]
2. High-volume preprocessing

⇒ Local Incidence Angle

\[ \alpha = \cos^{-1}(C \cos \theta - A \sin \theta) \]  

(2.8)

as \( \hat{n} \) and \( \hat{r} \) are both of unit length.

2.2.3.1 To establish normal vector

The vector \( \mathbf{n} \) can be established from the cross product of two vectors contained in the target surface. Figure 2.5 represents the target pixel in the centre, with DEM height \( e \), where the X axis from our local frame of reference described in Figure 2.4 points upwards, with the origin at the centre of the target pixel at height \( e \). Thus a vector \( \mathbf{v}_1 \) from the origin to the top-right pixel would be

\[ \mathbf{v}_1 = \begin{pmatrix} p \\ -p \\ a - e \end{pmatrix} \]

where \( p \) is the pixel spacing, and a vector \( \mathbf{v}_2 \) from the origin to the bottom-left pixel would be

\[ \mathbf{v}_2 = \begin{pmatrix} -p \\ p \\ i - e \end{pmatrix} \]

A diagonal vector \( \mathbf{d}_1 \) from the bottom-left to the top-right pixel would there-
2. High-volume preprocessing

fore be

\[ \mathbf{d}_1 = v_2 - v_1 = \begin{pmatrix} 2p \\ -2p \\ a - i \end{pmatrix} \]

Similarly, another diagonal vector

\[ \mathbf{d}_2 = \begin{pmatrix} 2p \\ 2p \\ g - c \end{pmatrix} \]

We can derive a vector normal to the plane containing vectors \( \mathbf{d}_1 \) and \( \mathbf{d}_2 \) by calculating their cross-product using the determinant of the formal matrix

\[
\mathbf{d}_1 \times \mathbf{d}_2 = \text{det} \begin{bmatrix} \mathbf{i} & \mathbf{j} & \mathbf{k} \\ 2p & -2p & (a - i) \\ 2p & 2p & (g - c) \end{bmatrix} \\
= \mathbf{i}(-2p(a - i) - 2p(g - c)) \\
- \mathbf{j}(2p(g - c) - 2p(a - i)) \\
+ \mathbf{k}(4p^2 + 4p^2)
\]

giving us the normal vector

\[ \mathbf{n}_1 = \begin{pmatrix} 2p(i - a + c - g) \\ 2p(a - i - g + c) \\ 8p^2 \end{pmatrix} \quad (2.9) \]
2. High-volume preprocessing

Similarly two other vectors contained with the target plane are

\[
d_3 = \begin{pmatrix} p \\ 0 \\ d - e \end{pmatrix} - \begin{pmatrix} -p \\ 0 \\ f - e \end{pmatrix} = \begin{pmatrix} 2p \\ 0 \\ d - f \end{pmatrix}
\]

and

\[
d_4 = \begin{pmatrix} 0 \\ p \\ h \end{pmatrix} - \begin{pmatrix} 0 \\ -p \\ b \end{pmatrix} = \begin{pmatrix} 0 \\ 2p \\ h - b \end{pmatrix}
\]

whose cross product

\[
d_3 \times d_4 = \det \begin{bmatrix} i & j & k \\ 2p & 0 & (d - f) \\ 0 & 2p & (h - b) \end{bmatrix}
\]

giving us another “normal” vector

\[
n_2 = \begin{pmatrix} 2p(f - d) \\ 2p(b - h) \\ 4p^2 \end{pmatrix}
\]

(2.10)

Of course, neither of the vectors in equations 2.9 nor 2.10 represent the true normal to the target plane (which only really exists as a mathematical concept), but they are approximations to the normal based on the surrounding terrain. It seems reasonable that our best approximation of the normal could be taken as the bisector of these two vectors, found by averaging \( n_1 \) and \( n_2 \) to give us

\[
n = \begin{pmatrix} p(i - a + c - g + f - d) \\ p(a - i - g + c + b - h) \\ 6p^2 \end{pmatrix}
\]

To obtain the unit vector we divide \( n \) by its length
2. High-volume preprocessing

\[ \hat{n} = \frac{1}{|n|} \begin{pmatrix} p(i - a + c - g + f - d) \\ p(a - i - g + c + b - h) \\ 6p^2 \end{pmatrix} \]  \hfill (2.11)

where

\[ |n| = \sqrt{p^2(i - a + c - g + f - d)^2 + 36p^4 + p^2(a - i - g + c + b - h)^2} \]

\[ = p \cdot \sqrt{(i - a + c - g + f)^2 + 36p^2 + (a - i - g + c + b - h)^2} \]

### 2.2.3.2 Final algorithm

This gives us the final algorithm for the local incidence angle \( \alpha \). We recall from equation 2.8 that

\[ \alpha = \cos^{-1}(C \cos \theta - A \sin \theta) \]

Into this we can now substitute the X and Z components of equation 2.11, such that

\[ A = \frac{S}{\sqrt{S^2 + 36p^2 + T^2}} \]

and

\[ C = \frac{6p}{\sqrt{S^2 + 36p^2 + T^2}} \]
\[ T = a - i - g + c + b - h \]
\[ S = i - a + c - g + f - d \]
Thus $\alpha$ is fairly easily calculated from DEM values, the pixel spacing $p$ and $\theta$, which is the angle of incidence with respect to the WGS84 ellipsoid. These angles are provided within the raw data files for sample tie points along each swath line. Keeping the calculations in the frame of reference of the raw data (ie prior to georectification) allows us to derive such values for all intermediate pixels by interpolation, along with the slant range times which are used in the orthorectification process, described in Section 2.2.2.

2.2.3.3 Methodological errors

There is an error caused by the use of the nominal incidence angle $\theta$, as opposed to the off-nadir angle $\phi$ (see Figure 2.3), to arrive at the vector $\hat{r}$ representing incident radiation in Equation 2.7. This error is plotted in Figure 2.6. In extreme cases, for example where $\theta = 44^\circ$ and with an elevation of 4000m, the this error would amount to less than $0.2^\circ$. This is considered negligible for our purposes.

The SPECAN algorithm used by the ASAR sensor in GM mode corrects for antenna elevation gain using functions derived from the periodic calibration data (ESA, 2007a). The correction is updated when a new periodic calibration cycle occurs according to continually shifting relationship between the sensor location and the datum ellipsoid. The error in incidence angle described above has a consequence on the output data, as the Antenna Elevation Pattern (EAP) is
2. High-volume preprocessing

indirectly dependent upon the incidence angle. A typical relationship between the elevation angle and the EAP is shown in Figure 2.7.

The relationship between the ground-projected backscatter $\gamma$ before and after orthorectification, the incidence angle $\theta$ and the off-nadir angle $\phi$ is as follows:

$$\gamma_{ORTH} = \gamma_{SRT} \left( \frac{G_{ele}^2(\theta)}{G_{ele}^2(\phi)} \right)$$

where $G_{ele}^2$ is the two-way antenna elevation pattern gain (Shimada, 2010b). Thus when converting to decibels, this gives us

$$\Delta \gamma \ (\text{dB}) = \Delta G_{ele}^2 \ (\text{dB})$$

The greatest rate of change of $G_{ele}^2$ with respect to elevation angle observed was below $-6 \ \text{dB degree}^{-1}$. This would result in an extreme case (4000m target height above datum, 44° look angle) in a difference of 1.2 dB to the final $\gamma$ figure. Whilst this amount is significant, it is within the stated radiometric accuracy of GM data (1.54–1.74 dB, ESA (2007a)), and in most cases the error would be considerably less, and it is not accounted for in this procedure.

Figure 2.7: Antenna elevation pattern against Elevation angle, for each of the five predetermined overlapping antenna beams operated in the ScanSAR modes, extracted from a GM data file centred over Australia.
2. High-volume preprocessing

2.3 Method

2.3.1 Choice of software and programming languages

The decision to use the *Geographic Resources And Support System* (GRASS) GIS package to manage raster files in a prototype image preprocessing module is not a difficult one to make (GRASS Development Team, 2009). GRASS adheres to the *Unix Philosophy* (Kernighan & Pike, 1984), ensuring that its modules can be combined into shell scripts or can be easily incorporated into other modules. It is open source, meaning that its source code is readily available, and that any work done to build functionality that is not currently available, such as is intended by this research, may contribute source code in turn that is then available to the wider community, and in this way GRASS, which has been around for nearly three decades, maintains currency in the ever-broadening field of GIS. GRASS can be used in a multi-user environment (Neteler & Mitasova, 2008), lending itself to be set up for parallel tasks, and it is able to handle large data sets with low memory load (Huang et al., 2011; Jasiewicz, 2010).

Although GRASS is written in C++, Perl\(^2\) was used for the prototype for a variety of reasons. Perl is known for its strength as a rapid prototyping language, being high level and having lazy memory management. It is compiled just-in-time, rather than being interpreted, and allows pre-execution syntax and sanity checks. Perhaps the three most valuable benefits of Perl are its ready availability on almost all Unix servers, important when considering the use of High Power Computing (HPC), its ability to run inside the GRASS shell environment, and its powerful REGEX engine. This latter virtue is of particular importance in handling output streamed from GRASS and from auxiliary software to extract header information from binary data files, and allows rapid data exchange between arrays and data stored in text files, buffering output to reduce expensive system calls.

\(^1\)http://www.cplusplus.com  
\(^2\)http://www.perl.org
2.3.2 Structure

The code was organised in a hybrid procedural/object-oriented manner. It was designed so that the task of preprocessing a large number of data files could be split among multiple nodes in an HPC. For each node, a unique GRASS environment is set up, including the automated output of a GRASS batch script, the setting of the GRASS_BATCH_JOB variable and the creation of a unique MAPSET within which to work, to avoid any locking conflicts. Each node is set to run a procedural batch script, the arguments of which include a target output directory and the allocated file input list. This script then calls GRASS commands and initiates Perl object instances, ultimately to output a GeoTIFF file, containing a minimum of two bands, representing the orthorectified DN values and the local incidence angles. A radar-shadow mask and a representation of layover/foreshortening can easily be produced (as these are calculated), but in this first instance this explicit output is avoided, to reduce I/O system calls and free up memory. Radar shadow is manifest in the output image in the form of null values.

Figure 2.8 shows the salient features of the principle classes used. The pivotal class is GrassAscii, which comprises the attributes and methods necessary to write to and read from an ASCII raster file compatible with GRASS GIS. Classes Alpha and Ortho are themselves child classes of GrassAscii, Alpha containing the method by which to calculate local incidence angles, and Ortho the method to perform the terrain correction. The GM class represents the raw data file, and comprises the methods required to extract the Annotation Data Set Records (ADSR) within the GM file and to interpolate the Slant Range Times (SRT), Latitudes and Incidence angles. These, together with the DN values and the recoded DEM, are exported as GrassAscii objects.

Figure 2.9 shows the operation flow and exchange of data between the process_gm procedure and the various modules. Operations on the left side are carried out in geographical lat-long coordinates, while those on the right occur within the x–y grid local to the raw data file. Data exchanges via ASCII grid files are shown in yellow.

Algorithm 1 shows pseudo-code demonstrating the orthorectification proce-
2. High-volume preprocessing

dure carried out by the Ortho class. For each row (each azimuth address) in the Digital Number (DN) array, two passes are made through the pixels, effectively travelling across the radar swath in the direction away from the nadir. The first pass (lines 5–19) determines the off-nadir angle $\phi$. If there has been a previous value of $\phi$ higher than the current value on the same swath, the pixel is marked as shadow; otherwise, the misplaced contribution value of the current pixel, based on its SRT and incidence angle, is added to the relevant address in the CONTRIBUTIONS array. In the second pass (lines 20–24), now knowing the total number of contributions that were aggregated to produce the pixel values in the raw data, these values are then redistributed according to the displacement $\Delta$ recorded in the first pass.

The function `RedistributeDNValue` in line 22 then apportions values in accordance with Equation 2.5 and as shown in Figure 2.2 in Section 2.2.2.2, as
2. High-volume preprocessing

Figure 2.9: Flowchart showing preprocessing carried out by one node running process_gm.pl follows:

\[
\begin{align*}
\text{ODN}[c] & \leftarrow \left[ (\Delta - \lfloor \Delta \rfloor) \cdot \frac{\text{DN}[c - \lfloor \Delta \rfloor - 1]}{\text{CONTRIBUTIONS}[c - \lfloor \Delta \rfloor - 1]} \right] + \\
& \quad \left[ (1 - \Delta + \lfloor \Delta \rfloor) \cdot \frac{\text{DN}[c - \lfloor \Delta \rfloor]}{\text{CONTRIBUTIONS}[c - \lfloor \Delta \rfloor]} \right]
\end{align*}
\]

where \( c \) is the current pixel range address. Pseudo-code representing the interpolation method provided in the GM class is shown in Algorithm 2. The algorithm interpolates firstly between tie-points across the swath to produce a full raster row (lines 3–7), and then interpolates down the columns for each pixel.
2. High-volume preprocessing

Algorithm 1 Orthorectify()

1: Fetch into arrays (DEM, DN, LAT, SRT, THETA)
2: for each row in DN do
3:   Zero-fill arrays (SHADOW, CONTRIBUTIONS, DELTA)
4:   $\phi_{\text{max}} \leftarrow 0$ # Max Off-nadir angle so far
5:   for each pixel in row do
6:     $s \leftarrow$ Calc Slant Range Distance (SRT)
7:     $\phi \leftarrow$ Calc Off-nadir Angle ($s$, THETA, DEM)
8:     if $\phi_{\text{max}} > \phi$ then
9:       SHADOW $\leftarrow$ true
10:     else
11:       $\phi_{\text{max}} \leftarrow \phi$
12:     $r \leftarrow$ Calc Radius At Latitude (LAT)
13:     $\Omega \leftarrow$ Calc Omega ($r$, DEM, R, s)
14:     $\omega \leftarrow$ Calc Little Omega ($r$, R, s)
15:     $\Delta \leftarrow$ Calc Delta ($r$, $\Omega$, $\omega$)
16:     Add contribution (CONTRIBUTIONS)
17:     Update Arrays (DELTA, SHADOW)
18:   end if
19: end for
20: for each pixel in row do
21:   if NOT SHADOW then
22:     RedistributeDNValue (DELTA, DN, ODN, CONTRIBUTIONS)
23:   end if
24: end for # Next Pixel
25: end for # Next Row

between the current raster row and that corresponding to the previous tie-point row (lines 8–10).

2.3.3 Parameters

The GM binary data contains within it a data set pertaining to a series of tie points, numbering precisely eleven across the swath (approximately 37 km apart) the number of rows in the azimuth direction depending on the length of the frame strip, which varies. For each tie point is given geodetic latitudes and longitudes, allowing ground control points to be derived. These, together with SRT, incidence angles and 2-way antenna elevation gains are extracted from the ADSR by calling
Algorithm 2 Interpolate()

1: for each Pointfile Row do
2:   RasterRow ← Get Corresponding Raster Row (PointfileRow)
3: for each Tiepoint in PointfileRow do
4:   lastTPval ← Get Tiepoint Value (This)
5:   nextTPval ← Get Tiepoint Value (Next)
6:   lastTPcol ← Get Raster Column of Tiepoint (This)
7:   nextTPcol ← Get Raster Column of Tiepoint (Next)
8: for each col ← raster column between lastTPcol and nextTPcol do
9:   Pixelvalue ← lastTPval + ( (nextTPval - lastTPval) \times (col - lastTPcol) / (nextTPcol - lastTPcol) )
10: end for # Next col
11: end for # Next Tiepoint
12: end for # Next Pointfile Row

2. High-volume preprocessing

the command line tool Pds2Ascii, provided within EnviView by ESA (ESA, 2010).

<table>
<thead>
<tr>
<th>Parameter</th>
<th>Function</th>
<th>Dataset</th>
<th>Field</th>
</tr>
</thead>
<tbody>
<tr>
<td>Absolute Calibration Constant K</td>
<td>Conversion to $\sigma^0$</td>
<td>2</td>
<td>calibration_factors</td>
</tr>
<tr>
<td>Orbit State Vectors X,Y,Z</td>
<td>Calculation of $R_{\text{satellite}}$</td>
<td>2</td>
<td>orbit_state_vectors</td>
</tr>
<tr>
<td>Tie Points</td>
<td>Lat, Long, SRT, $\theta$</td>
<td>7</td>
<td>*_line_tie_points</td>
</tr>
<tr>
<td>AEP</td>
<td>EAP, $\Lambda$</td>
<td>6</td>
<td>elevation_pattern,*</td>
</tr>
</tbody>
</table>

Table 2.2: Parameters extracted from GM data file using Pds2Ascii

Table 2.2 shows the parameters extracted from the GM data file, together with the particular data set and field in which they are contained. Data set 2 refers to the Main Processing Parameters ADS, 6 is the MDS1 Antenna Elevation Pattern ADS and 7 is the Geolocation Grid ADS. $R_{\text{satellite}}$ is the distance of the satellite from the center of the WGS84 ellipsoid, used in Equations 2.2 and 2.3 in Section 2.2.2.1. The radius $r$ (ellipsoid radius at target location) used in the same equations is calculated as follows (e.g. Groten (2004)):

$$r_{\Theta} = a(1 - f \sin^2 \Theta)$$

where $a = \text{Equatorial Radius} = 6378137m \ (\text{WGS84})$

$$f = \text{flattening} = (a - c)/a$$

$c = \text{polar radius} = 6356752.3m$

$\Theta = \text{Geocentric Latitude}$
2. High-volume preprocessing

The Geocentric Latitude $\Theta$ relates to the Geographic Latitude $\Phi$ by

$$\tan \Theta = (1 - f)^2 \tan \Phi$$

2.3.4 Multi-node network distribution

The key benefit of having control over the scripts which manage the preprocessing of a large number of GM data files is the ease with which the task can be split and allocated to a High Power Computing network (HPC). The HPC used in this instance managed multiple nodes on a Portable Batch System (PBS\textsuperscript{1}). The perl script which manages the distribution of work to the HPC may be found in Appendix A, Listing 6. The essence of the workings of the script is described in Algorithm 3 below.

The script looks for all GM data files under the current directory and separates them into batches, creating a directory for each batch and symbolic links within each to those folders. Each batch will then be processed concurrently. Two batch scripts are then created by the main script for each batch - one being the instructions for Grass GIS, the other being the PBS batch script. The latter is submitted to the HPC server. The start of each PBS script sets the GRASS\_BATCH\_JOB variable to point to the individual Grass batch script, and starts a new instance of GRASS GIS, creating a unique temporary MAPSET within which to work.

2.4 Results and discussion

The decision to develop a bespoke method for the preprocessing of GM data was based on the need to have control over the process, partly due to the sheer numbers of data files to be managed, but partly also to allow for interruption to parts of the process where appropriate - to allow for deviation from the usual methods under certain circumstances. Examples of this would be the case of work done using image differencing techniques involving data taken from the same orbit track, where certain incidence angle multipliers may be factored out

\textsuperscript{1}http://www.OpenPbs.org/
2. High-volume preprocessing

Algorithm 3 PBS algorithm
1: for each Batch do
2: Create Batch Directory
3: Link allocated GM files to Batch Directory
4: end for # Next Batch
5: for each Batch Directory do
6: Write to ThisGrassBatch “process gm.pl ThisGrassBatch ThisDirectory”
7: Make ThisGrassBatch executable
8: Write to ThisPBSBatch “Load Grass”
9: Write to ThisPBSBatch “Set GRASS_BATCH_JOB = ThisGrassBatch”
10: Write to ThisPBSBatch “Start Grass using MAPSET name MThisDirectory”
11: Make ThisPBSBatch executable
12: Submit ThisPBSBatch to PBS Server
13: end for # Next Batch Directory

of the calculations early in the processing (O’Grady et al., 2011), and in the case of regression calculation applied over hundreds of images, where the individual results of image processing are not kept, but rather, following preprocessing, their pixel-wise contribution to running sums used in the final calculations is aggregated (in the case of linear regression, for example, aggregate images would store $\Sigma x$, $\Sigma x^2$, $\Sigma xy$, $\Sigma y^2$, $\Sigma x^2y^2$, and $n$).

Adequacy of the methods described here, in terms of fitness for purpose, accuracy and speed are an issue. Once adequate accuracy is established, fitness for purpose is gleaned from the use of the methods, the worth of their results and the ability for the process to be run in a reasonable time. In this regard, the speed requirement for the purposes of our research was actually a binary condition of fitness for purpose. Using the HPC in the manner described, hundreds of GM data files were able to be processed in a matter of minutes. Having the data-download and preprocessing of the (typically) 70 or so daily images being automated remotely overnight by the cron (time-based scheduler) daemon, this fact meant an easy tick for the speed criterion. Speed benchmarking tests were not carried out for this reason.

Accuracy was tested by comparing outputs from our process with that for the same data files processed by what is perhaps the most suitable contender -
The well resolved *NEST* (Next ESA SAR Toolbox) package developed by Array Systems Computing \(^1\) and made available online through ESA’s website \(^2\).

The tests were confined to the range of latitudes and elevations for which our methods were to be used in Australia. A number of bodies of open water were chosen around the region at various heights above sea level. Transects were sampled across the centre of the water bodies in azimuth and range directions. The resultant values from three images were overlain against each other in plot outputs; MODIS Band 6 (SWIR $\lambda \approx 1620$ nm), radar backscatter ($\sigma^0$) from our preprocessed data and that from the data processed using NEST. The MODIS plot shows us a good indication of the spatial demarcation between land and water for comparison with the responses from the two separately derived radar images. The differences observed, in terms of pixels, are tabulated and can be seen in Table 2.3. Accuracy for both of the radar images was worse in the azimuth direction, with displacement averaging at 2.56 pixels for NEST and 2.33 for our method. In the range direction, average displacement for NEST was 1.63, against our 1.31. The apparent greater accuracy in the azimuth direction using our method comes as somewhat of a surprise, due to the fact that our calculations work solely in the range direction, but it is considered likely that the results would be reversed in more extreme terrain, which would be outside the scope of our research with this method.

Figure 2.10 shows some of the output of the analysis as it appeared graphically. The conclusion was that our method produced a level of spatial and radiometric accuracy that was certainly comparable with a very worthy third-party alternative.

\(^1\)http://www.array.ca/
\(^2\)http://nest.array.ca/web/nest
2. High-volume preprocessing

Figure 2.10: Transect value profiles in Azimuth (left) and Range (right) directions across three bodies of water in Australia. The red profile corresponds to MODIS Terra or Aqua Reflectance in thousands (scale on the right of the plots), the green corresponds to $\sigma^0$ output using our algorithm, and the blue using NEST software, both scaled in dB on the left axes.
2. High-volume preprocessing

<table>
<thead>
<tr>
<th>Ref</th>
<th>Azimuth</th>
<th>Range</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Our diff from MODIS</td>
<td>NEST diff from MODIS</td>
</tr>
<tr>
<td>1</td>
<td>1</td>
<td>2</td>
</tr>
<tr>
<td>2</td>
<td>2</td>
<td>2</td>
</tr>
<tr>
<td>5</td>
<td>2</td>
<td>1.5</td>
</tr>
<tr>
<td>6</td>
<td>3</td>
<td>2.5</td>
</tr>
<tr>
<td>7</td>
<td>6</td>
<td>5</td>
</tr>
<tr>
<td>8</td>
<td>1.5</td>
<td>3.5</td>
</tr>
<tr>
<td>9</td>
<td>1.5</td>
<td>1.5</td>
</tr>
<tr>
<td>10</td>
<td>0</td>
<td>3</td>
</tr>
<tr>
<td>11</td>
<td>4</td>
<td>2</td>
</tr>
<tr>
<td>MEAN</td>
<td>2.33</td>
<td>2.56</td>
</tr>
</tbody>
</table>

**Table 2.3:** Spatial displacement (“diff”) in pixels between edges of sampled water bodies as identified in outputs from our described process, ESA’s NEST software and MODIS Band 5. Values were obtained by comparing profiles for transects across the water bodies in the azimuth and range directions. Water thresholds were chosen as $\sigma_0 = -16\text{dB}$ for the radar data, against a reflectance of 0.2 for MODIS Band 5.
2. High-volume preprocessing

2.5 Conclusion

Preprocessing of GM data was performed using GDAL functions, GRASS raster manipulation techniques and Perl scripting, allowing control and transparency over the whole procedure, and allowing a very high volume of data to be processed quickly on JCU’s HPC. The accuracy of the process was tested to the extent that enabled us to determine the method’s fitness for purpose. Throughout its subsequent use, many “one-off” comparisons were possible with occasional data processed by ESA’s NEST software, and in this way the ongoing satisfactory performance of the method described above was further verified.

The philosophy of collaboration, transparency and sharing which enables the development of tools that are so valuable to the GIS and remote sensing community, such as GDAL and GRASS, as well as the hugely important and ubiquitous tools such as Perl and C++, ties in very well with the philosophy of collaborative research in the scientific field. Whilst many of these tools pre-date much of the commercially-developed software available, the open source model is becoming increasingly prominent, and we envisage it playing a vital role in meeting the increasing needs of emerging technologies in remote sensing.
Chapter 3

Relationship of incidence angle with satellite radar backscatter for different surface conditions

Chapter context
Abstract

This chapter aims to exploit the use of multiple angles of incidence, and the high temporal frequency of Envisat’s Advanced Synthetic Aperture Radar (ASAR) Global Monitoring Mode data to observe the relationship between the radar backscatter response from different land surfaces with varied angles of incidence. The primary objective is to establish a parameter surface for a region against a model relating the variation in backscatter with incidence angle, in order that backscatter values across the swath in a GM image can be normalised. Backscatter response characteristics to varied incidence angles are examined against various surface categories using pixel-level regression across a large time series, with regolith displaying the strongest correlation. The potential use of the slope of the linear model approximating the relationship of incidence angle with backscatter, to directly map open water accurately, is observed, with the method’s possible reduction of Bragg Resonance effects in C-HH data being apparent.

3.1 Introduction

Much research has been carried out to investigate the value of satellite radar data in the classification of ground surface properties (Bindlish et al., 2009; Bonn & Dixon, 2005; Chandran et al., 2006; Frappart et al., 2005; Waisurasingha et al., 2007; Wilson & Rashid, 2005). Where data taken from differing look angles is used for comparison or change detection, the elimination of the effects of differing incidence angle by normalisation to some common angle is usually attempted alongside despeckling, as a necessary part of preprocessing.

However, change in backscatter with respect to incidence angle is a function of the structural and dielectric properties of the target surface, being sensitive, therefore, to differences in regolith, soil types, vegetation and land use, for example. As these are the very parameters by which researchers mean to classify the surface, it is important that normalisation functions applied in the preprocessing stage are specific to these properties, in order that changes in backscatter observed at a particular location can be confidently attributed to an environmental change, rather than to a difference in the incidence angles used to obtain the
3. Incidence angle effects

respective values.

The high frequency of coverage provided by Envisat ASAR’s Global Monitoring Mode (GM) products affords us a time series across multiple incidence angles that gives us sufficient data to carry out regression, across the time series, for all locations, in order to derive a relationship of backscatter with incidence angle specific to the ground conditions in each location.

3.2 Theoretical Context

Recapping a little from Section 1.4.1.5, GM data, as made available by the European Space Agency (ESA), comprises digital numbers (DN) which are related to the local incidence angle $\alpha$ by

$$DN^2 = K \cdot \beta^0 = K \cdot \frac{\sigma^0}{\sin(\alpha)} \quad (3.1)$$

where $\sigma^0$ is the radar backscatter coefficient and $K$ the absolute calibration constant (ESA, 2004). For a volume scatterer being comprised of random scatterers of varying sizes and orientations, we expect

$$\gamma_\alpha = \frac{\sigma^0}{\cos \alpha} \quad (3.2)$$

where $\gamma_\alpha$ is the range independent backscatter coefficient at incident angle $\alpha$ (Shimada, 2010a).

3.2.1 Local incidence angle

The quantity of GM data available affords us the opportunity to directly investigate the relationship between the incidence angle and $\sigma^0$ or $\gamma$, as derived from equations 3.1 and 3.2.

Recalling the assumption by Baghdadi et al. (2001) that
3. Incidence angle effects

\[ \sigma_0^\alpha \approx \sigma_0^0 \cos^\Psi \alpha \quad (3.3) \]

(adapted from Ulaby et al. (1982)), where \( \Psi \) is the slope of the linear fit of \( \sigma^0 \) (dB) against \( 10 \log(\cos \alpha) \) (Baghdadi et al., 2001).

From Equations 3.2 and 3.3 we can say

\[ \gamma_\alpha \approx \sigma_0^0 \cos^\psi \alpha \quad (3.4) \]

where \( \psi = \Psi - 1 \).

![Figure 3.1: Function \( f(\alpha) = 25 \log(\cos \alpha) \) for \( \psi = 2.5 \) and \( \psi = 10 \) within the range of incidence angles \( \alpha \) relevant to GM data.](image)

Some researchers have assumed a near linear relationship between \( \sigma^0 \) and \( \alpha \) (e.g. Pathe et al. (2009)), in cases where incidence angles are confined to a narrow range. Based on equation 3.4, typical values of \( \psi \) were found to be between 2.5 over land and 10+ over water. Figure 3.1 shows that a linear approximation of these powers of cosine within the range of incidence angles found in GM data (15°–44°) would be reasonable, especially in the context of the 1.2 dB radiometric resolution of GM data (Pathe et al., 2009; Sabel et al., 2008).
3. Incidence angle effects

3.2.2 Physical models

Simple physical models class radar scatterers as surface scatterers, volume scatterers and heterogeneous combinations of the two. The backscatter of surface scatterers depends on the dielectric properties of the surface for scattering strength, and on the surface roughness for the distribution of scatter with respect to incidence angle. With volume scatterers, dielectric discontinuities and heterogeneous densities dictate scattering strength, with the dependence on incidence angle being governed by boundary surface roughness, average dielectric constant and structural scale within the medium (Ulaby et al., 1982).

The effects on backscatter of a variety of different types of surface must be considered, especially where pixels are 0.5km in size. The response of bare soil is governed mainly by surface roughness and soil moisture content, with volume scattering contributing in the case of extremely dry soil. Studies have found that for bare soil at a given angle of incidence, radar backscatter (in decibels) has a near-linear relationship with the soil moisture content / moisture capacity ratio (Ulaby et al., 1982). Where the soil is covered by vegetation, the relationship between the structural dimensions of the components of the vegetation and the wavelength of radar signal, together with the extent of cover, add to the response characteristics of the soil. C-band radar is more sensitive to smaller-structured vegetation components (Noernberg et al., 1999), such as grasses and herbaceous vegetation (Parmuchi et al., 2002) and aquatic plants (Henderson & Lewis, 2008).

3.3 Methodology

3.3.1 Data

GM data were downloaded systematically from ESA’s Earthnet Online Portal through a Category 1 Fast Registration agreement (ESA, 2009b). This data comprised all GM acquisitions for the year 2009 between latitudes 10° S and 29° S and from longitude 140° E to 155° E.

MODIS data were downloaded via NASA’s GSFC web portal (NASA, 2011). A comprehensive list of data used may be found in Appendix B.
3. Incidence angle effects

3.3.2 Procedure

For each GM data file, a raster surface representing local incidence angles was created with 18” pixel spacing, using satellite geo-location parameters, digital elevation models and tie-point data provided within the data header files.

Firstly, DN values were converted to backscatter according to equation 3.1, to obtain $\sigma_0^\alpha$, and orthorectified. An investigation was carried out assuming the relationship described in equation 3.4. This equation may be written

$$\gamma_\alpha(dB) \approx \psi \cdot 10\log(\cos\alpha) + \sigma_0^0(dB)$$ (3.5)

the slope of $\gamma_\alpha$, as a function of $10\log(\cos\alpha)$, being $\psi$. Regression was carried out for each pixel over the time series for each quarter of 2009.

To analyse the variance of $\gamma_{30}$ and $\Delta\gamma/\Delta\alpha$ for different physical parameters, 20,000 point sites were randomly chosen within a region of Queensland. For each of the points, values of the following parameters were extracted from the ancillary data:

1. Dominant vegetation species
2. Vegetation growth form
3. Soil type
4. Geology
5. Lithology
6. MODIS Band 6 ($\lambda \approx 1.6\mu m$)

The value of $\psi$ derived for the whole of 2009 over Queensland compared reasonably well with those of Baghdadi et al. (2001) for horizontally co-polarised C-band data over different land covers. Coefficient of Determination $R^2$ showed a poor fit over land, averaging 0.1, but averaged around 0.8 over the sea, where values of $\psi$ were typically above 10.

Further regression was carried out on the same dataset based on the relationship described for $\sigma^0$: 
3. Incidence angle effects

\[ \bar{\sigma}_0^0 = \bar{\sigma}_0^0 e^{-\alpha/\alpha_0} \]  

by Ulaby et al. (1982) after Moore & Fung (1979), who found the e-folding angle \( \alpha_0 \) to be 5.35° for HH polarised 13.9 GHz Skylab scatterometer data over the sea (Ulaby et al., 1982). In our case a comparable e-folding angle over sea was found to have a mean of 5.1°, with a standard deviation of 0.6°. The same regression resulted in an average e-folding angle over land of 48.5°. However analysis of the \( R^2 \) showed goodness-of-fit to be confined once again to open water.

The second round of analysis was based on the assumption that, given the narrow range of incidence angles from which the GM data was derived (15° – 44°), and the radiometric resolution of 1.2 dB (see Figure 3.1), the range independent backscatter coefficient \( \gamma \) could be normalised with respect to \( \alpha \) based on a linear function whose slope and intercept were derived for each pixel over the time series.

Due to the great influence of surface wetness and soil moisture on backscatter values, and considering the seasonal variation in rainfall over Queensland, regression was carried out over each of the quarters of 2009 separately.

3.4 Results and discussion

3.4.1 Relationship of backscatter with incidence angle

Figure 3.2 shows the range independent backscatter coefficient \( \gamma_{30} \), calculated by feeding 30° into the linear best-fit function for each pixel, over each quarter. Figure 3.3 shows the slope of the same function, \( \Delta \gamma/\Delta \alpha \). It can be seen that \( \gamma_{30} \) and \( \Delta \gamma/\Delta \alpha \) show quite different features over the surface of northern Queensland. It is also apparent that the significant difference seen between the \( \gamma_{30} \) values in the first (wettest) quarter and the other drier quarters are not matched with quite such a difference in the values of \( \Delta \gamma/\Delta \alpha \). This is to be expected, as while wet soil and vegetation return a higher backscatter value than their dry equivalent, they remain diffuse reflectors.

Figure 3.4 shows a comparison of the \( \gamma_{30} \) and \( \Delta \gamma/\Delta \alpha \) values for the third
3. Incidence angle effects

Figure 3.2: Range independent backscatter coefficient $\gamma_{30}$ (dB), calculated from the linear best-fit function at 30°, for each quarter of 2009. High backscatter values caused by the wet season in the tropical far north of Queensland dominate the whole of Cape York Peninsula in the first quarter.

The greater separability between terrestrial values and those of the sea in the $\Delta \gamma / \Delta \alpha$ image over the $\gamma_{30}$ image are clear. Additionally, the $\gamma_{30}$ image shows bright areas of increased value in the sea along the east coast. These are the result of wind effects due to Bragg resonance on the surface of the water (Schaber et al., 1997). The prevalence of these effects is confirmed by their presence in this image despite the fact that, rather than being instantaneous, the image is a composite modelled on some 40 data files. It is significant that these effects are not seen, at least to the same degree, in the $\Delta \gamma / \Delta \alpha$ image.

The difference in separability of land and water between $\gamma_{30}$ and $\Delta \gamma / \Delta \alpha$ can be seen by comparing Figures 3.5 and 3.6. The density plot in the former shows that the separation of land from water would be problematic, whereas the latter shows us that, given sufficient data surrounding a flood event by which to establish $\Delta \gamma / \Delta \alpha$, the separability between water and land is greatly increased. The con-
3. Incidence angle effects

Figure 3.3: Slope $\Delta \gamma / \Delta \alpha$ (dB per degree) from linear regression for each quarter of 2009 sequence of this is that, given a sufficient temporal data frequency, $\Delta \gamma / \Delta \alpha$ may give us a means to accurately separate open water from land, despite the presence of Bragg resonance effects.

3.4.2 Variation with land surface properties

In the categorical analysis to determine, if applicable, the driving factor behind $\Delta \gamma / \Delta \alpha$ characteristics in terms of surface conditions, separability between categorical groups was examined both between each group pair, and over each categorical group as a whole. In the latter case, the singular values (SV), which give the ratio of the between- and within-group standard deviations on the linear discriminant variables, were calculated (the SV is the square root of the canonical F-statistic), with the results shown in Table 3.1.

The results of Table 3.1 tell us that the factors *regolith* and *dominant vegetation species* in Queensland have demarcations which fall closer into line with the variation of both $\gamma_{30}$ and $\Delta \gamma / \Delta \alpha$, in comparison with the other classifiers.
3. Incidence angle effects

Figure 3.4: Comparison of $\gamma_{30}$ (above) and $\Delta \gamma/\Delta \alpha$ (below) for the third quarter of 2009
Figure 3.5: (Top) Scatter plot showing MODIS Band 6 reflectance against $\gamma_{30}$, which clearly separates water from land, and (bottom) the corresponding density plot of $\gamma_{30}$ values, demonstrating their separability (or lack thereof)
Figure 3.6: (Top) Scatter plot showing MODIS Band 6 reflectance against $\Delta \gamma / \Delta \alpha$, which clearly separates water from land, and (bottom) the corresponding density plot of $\Delta \gamma / \Delta \alpha$ values, demonstrating their separability.
3. Incidence angle effects

### Table 3.1: Singular values indicating categorical separability among land cover classifiers, using $\gamma_{30}$ and $\Delta \gamma / \Delta \alpha$

<table>
<thead>
<tr>
<th>Classifier</th>
<th>SV $\gamma_{30}$</th>
<th>SV $\Delta \gamma / \Delta \alpha$</th>
</tr>
</thead>
<tbody>
<tr>
<td>Soil type</td>
<td>18.99</td>
<td>66.24</td>
</tr>
<tr>
<td>Gross rock Descriptor</td>
<td>19.08</td>
<td>48.24</td>
</tr>
<tr>
<td>Growth form</td>
<td>23.79</td>
<td>79.27</td>
</tr>
<tr>
<td>Dominant species</td>
<td>29.64</td>
<td>109.08</td>
</tr>
<tr>
<td>Regolith</td>
<td>30.07</td>
<td>109.14</td>
</tr>
</tbody>
</table>

From Figures 3.7 and 3.8 it can be seen that there is some spatial correlation between regolith and species. The clue as to which factor influences backscatter lies with the SV result for the classifier *growth form*, which groups the $\gamma_{30}$ values according to the following forms:

- Hummock Grasses
- Low shrubs < 2 metres
- Low trees < 10 metres
- Medium trees 10–30 metres
- Other herbaceous plants
- Tall Shrubs > 2 metres
- Tall trees > 30 metres
- Tussocky or tufted grasses

If the relatively high separability of vegetation classes using $\gamma_{30}$ were due to the structural properties of the vegetation itself, we might expect to see a similar SV value for growth form as we see for vegetation (the *Dominant Species* classifier), but this is not the case. This suggests that, of the classifiers studied, regolith has the highest correlation between incidence angle and backscatter.

Figure 3.9 shows us dominant vegetation species, displaying standard deviation (boxes), 25% and 75% percentiles (outer tics) and median values (central tics) of $\gamma_{30}$ and $\Delta \gamma / \Delta \alpha$ for each class. From the $\gamma_{30}$ plot on the left, it is clear that the characteristically low $\gamma_{30}$ values of water are shared by some vegetation classes, particularly *Chenopodiaceae* (Saltbush and Bluebush), *Astrebla* (Mitchell Grass) and *Acacia*. The plot on the right, showing $\Delta \gamma / \Delta \alpha$ values against the
same vegetation species, separates water from everything else due to its relatively high change in backscatter with respect to incidence angle.

The mean and standard deviations for $\gamma_{30}$ and $\Delta\gamma/\Delta\alpha$ against regolith classes are shown in Table 3.2. Figure 3.10 shows the corresponding box plots. In the $\gamma_{30}$ plot on the left of Figure 3.10, the backscatter values for water have the biggest overlap with classes representing aeolian sand and residual material. As can be seen from the regolith class map in Figure 3.7, these classes occupy the dry interior in the south-west of Queensland, and the low backscatter represents absorption and attenuation of the signal. Perhaps the only specular reflectors other than water represented here are lacustrine sediments. All of these classes become clearly separable from water in the $\Delta\gamma/\Delta\alpha$ plot.

If we return once again to the vegetation species which are less easily separable from water in terms of their $\gamma_{30}$ values (i.e. Chenopodiaceae, Astrebla and Acacia), and look at their distribution over the region (Figure 3.8), we find that they are, to a large extent, confined to those areas of the western interior where absorption dominates the nature of backscatter response.
3. Incidence angle effects

Figure 3.8: Map showing dominant vegetation species of Queensland (source Geoscience Australia)

3.5 Conclusion

We have established the slope for a linear model describing the location-specific variation of backscatter with incidence angle, by which we may normalise backscatter values to mitigate the effects of the variance of incidence angle across a swath. This is an important precursor to the appliance of a threshold in the classification of water.

Surface soil moisture and Bragg resonance have less of an effect on the change of backscatter with respect to incidence angle than on absolute backscatter values, providing a possible means to mitigate their hindering of land-water segmentation. The potential for the use of multiple images, at multiple incidence angles, to observe a quasi-instantaneous rate of change of backscatter with incidence angle in order to determine the presence of water with a greater degree of confidence is apparent. The potential use of such a method may be considered in the event of future C-Band ASAR missions providing sufficiently frequent coverage. This is
3. Incidence angle effects

Figure 3.9: Box plots showing median, 25% and 75% percentiles (tics) for values of ground range projected backscatter $\gamma$ (dB, left) and $\Delta\gamma/\Delta\alpha$ (right) against dominant vegetation species, based on 20,000 randomly-selected sites in Queensland. Box widths are proportional to number of sites within each category.
3. Incidence angle effects

Figure 3.10: Box plot showing ground range projected backscatter $\gamma$ (dB, left) and $\Delta\gamma/\Delta\alpha$ (right) against lithology classes, based on 20,000 randomly-selected sites in Queensland.
3. Incidence angle effects

<table>
<thead>
<tr>
<th>REGOLITH</th>
<th>$\gamma_{30}$ MEAN</th>
<th>STD</th>
<th>$\Delta \gamma/\Delta \alpha$ MEAN</th>
<th>STD</th>
</tr>
</thead>
<tbody>
<tr>
<td>Aeolian sand</td>
<td>-13.19</td>
<td>1.98</td>
<td>-0.21</td>
<td>0.11</td>
</tr>
<tr>
<td>Alluvial sediments</td>
<td>-11.73</td>
<td>1.58</td>
<td>-0.14</td>
<td>0.09</td>
</tr>
<tr>
<td>Coastal sediments</td>
<td>-11.69</td>
<td>1.97</td>
<td>-0.18</td>
<td>0.08</td>
</tr>
<tr>
<td>Highly weathered bedrock</td>
<td>-10.60</td>
<td>1.75</td>
<td>-0.11</td>
<td>0.07</td>
</tr>
<tr>
<td>Lacustrine sediments</td>
<td>-12.43</td>
<td>1.10</td>
<td>-0.13</td>
<td>0.05</td>
</tr>
<tr>
<td>Moderately weathered bedrock</td>
<td>-10.30</td>
<td>1.13</td>
<td>-0.08</td>
<td>0.07</td>
</tr>
<tr>
<td>Residual clay</td>
<td>-12.24</td>
<td>1.14</td>
<td>-0.18</td>
<td>0.05</td>
</tr>
<tr>
<td>Residual material</td>
<td><strong>-12.12</strong></td>
<td><strong>2.30</strong></td>
<td><strong>-0.18</strong></td>
<td><strong>0.07</strong></td>
</tr>
<tr>
<td>Residual sand</td>
<td>-10.93</td>
<td>1.38</td>
<td>-0.08</td>
<td>0.07</td>
</tr>
<tr>
<td>Soil on bedrock</td>
<td>-9.54</td>
<td>1.28</td>
<td>-0.07</td>
<td>0.07</td>
</tr>
<tr>
<td>Terrestrial sediments</td>
<td>-11.20</td>
<td>1.79</td>
<td>-0.12</td>
<td>0.08</td>
</tr>
<tr>
<td>Very highly weathered bedrock</td>
<td>-11.45</td>
<td>1.49</td>
<td>-0.14</td>
<td>0.08</td>
</tr>
<tr>
<td>Water</td>
<td><strong>-13.47</strong></td>
<td><strong>1.38</strong></td>
<td><strong>-0.78</strong></td>
<td><strong>0.13</strong></td>
</tr>
</tbody>
</table>

Table 3.2: Mean and standard deviation $\gamma_{30}$ and $\Delta \gamma/\Delta \alpha$, for regolith classes

especially the case if, as seen, the high backscatter observed as a result of Bragg Resonance varies to a similar degree as would be observed from water where the phenomenon were not present. This would provide a possible means to eliminate an effect which is a significant barrier to the broad use of C-HH radar for the classification of water (Figure 3.11 demonstrates the results of Bragg Resonance well).

From the above analysis, we may conclude as follows:

1. For GM signals, the underlying regolith is the most important contributor to the level of backscatter under dry conditions.

2. Where there are sufficient contemporary GM images of a water body at varied incidence angles, the rate of change of backscatter with incidence angle may provide a far clearer means to map the water extents than a single threshold on backscatter alone.

3. Where absolute backscatter values are to be used as a threshold in the detection of water, absorption in very dry soils and specular reflection on lacustrine sediments are potential sources of commission errors.
3. Incidence angle effects

**Figure 3.11:** GM images of the Aral Sea on 18 June (above) and 21 June 2010 (below). The contrast is all but eliminated by what can only be the effects of Bragg Resonance due to particular wind conditions on 21 June.
Chapter 4

Use of ENVISAT ASAR Global Monitoring Mode to complement optical data in the mapping of rapid broad-scale flooding: case study of the 2010 Indus flood

Chapter context
Abstract

Envisat ASAR Global Monitoring Mode (GM) data were used to produce maps of the extent of the 2010 flooding in Pakistan which were made available to the rapid response effort within 24 hours of acquisition. The high temporal frequency and independence of the data from cloud-free skies makes GM data a viable tool for mapping flood waters during those periods where optical satellite data is unavailable, which may be crucial to rapid response disaster planning. Image differencing techniques were used, with pre-flood baseline image backscatter values being deducted from target values to eliminate regions with a permanent flood-like radar response due to volume scattering and attenuation, and to highlight the low response caused by specular reflection by open flood water. The effect of local incidence angle on the received signal was mitigated by ensuring that the deducted image was acquired from the same orbit track as the target image. Poor separability of the water class with land in areas beyond the river channels was tackled using a region-growing algorithm which sought threshold-conformance from seed pixels at the center of the river channels. The resultant mapped extents were tested against MODIS SWIR data where available, with encouraging results.

4.1 Introduction

4.1.1 Pakistan floods

Over the 2010 monsoon season, Pakistan saw extensive flooding of the Indus river and its tributaries, which affected over 20 million people, damaging over 2 million hectares of crop land and causing the loss of 1,985 lives (NDMA, 2011). Heavy rainfall in the northern regions of the Khyber Pakhtunhwa, reaching 280 mm on 29 July, damaged major irrigation headworks on the Swat River at Munda, which were built to a discharge capacity of 4.5 Mls$^{-1}$ and which were damaged by the peak discharge of 8.5 Mls$^{-1}$. Further rainfall in Gilgit and Jammu and Kashmir and further south in Balochistan, contributed further to the huge body of water which flooded irrigation channels and agricultural land covering tens of
thousands of square kilometers. The UN Food and Agricultural Organization estimate losses of wheat stocks at around 450,000 tonnes (Fair, 2011). The Damage Needs Assessment conducted by the World Bank estimated that the recovery from the floods would cost between $8.7 and $10.9 billion (WBG, 2011).

To facilitate the international relief effort in such crises, maps are made available in near real time by facilities such as NASA’s MODIS Rapid Response System, which makes use of the Moderate Resolution Imaging Spectroradiometer (MODIS) on board NASA’s Aqua and Terra satellites. Such instruments are, however, limited by the cloud cover that is often present for some time after such flood events. It is largely for this reason that, in recent years, satellite-borne radar instruments have attracted much research into their viability as a means to map flooding, due to their ability to penetrate cloud cover and to their independence from the relative position of the sun (Rosenqvist et al., 2007; Waisurasingha et al., 2007; Wilson & Rashid, 2005).

The classification of water with satellite radar data is problematic and error-prone, particularly when some of the main environmental factors affecting the result, such as wind speed and direction and soil moisture, are unknown. The clear advantage of the independence from cloud cover may result in the availability of radar data where more reliable optical data is not available. Revisit time is an important feature of remotely sensed data used for disaster management. The COSMO-SkyMed constellation\(^1\) comprises a cluster of four X-band SAR sensors on the same orbit path which are capable of a revisit time of less than 12 hours. Most radar sensors such as the Advanced Synthetic Aperture Radar (ASAR) aboard the European Space Agency’s (ESA) Envisat satellite, and the Phased Array type L-band Synthetic Aperture Radar (PALSAR) on the Japanese Aerospace Exploration Agency’s (JAXA) ALOS satellite\(^2\) have repeat orbit cycles of more than a month, but are able to provide a higher repeat coverage thanks to operation modes which overlap regions at different incidence angles on adjacent orbits. Here we take a closer look at data from the ASAR sensor operating in Global Monitoring (GM) mode, which is systematically acquired when data in other configurations is not required. GM data is made available in near real-time.

\(^1\)http://www.cosmo-skymed.it/

\(^2\)both of which satellites have unfortunately become defunct recently
for download to parties with at least a Category-1 fast-track agreement with ESA. The data is available quickly because it is preprocessed at the sensor, before being transmitted down to one of two ground stations in Europe. This is made possible by keeping file sizes and processing requirements low, by using a coarse resolution (pixel size is 500m, spatial resolution is 1km). Such a coarse resolution over the full range of incidence angles ($12^\circ$ to $44^\circ$) may prohibit the data’s viability as an alternative means to map flooding, but we believe that its temporal frequency and ready availability give the data potential advantages that warrant further investigation. The following questions emerge:

- Is the frequency of GM coverage such that it can provide an alternative source of data to map flooding when cloud cover prohibits the use of optical data?
- Can effects due to factors such as incidence angle be eliminated?
- Will the radiometric uncertainty of signals received from partially and totally inundated areas and non-flooded areas allow the classification of flooding to a level of accuracy sufficient to produce useful maps, given the coarse resolution of GM data?

4.2 Study Area

The flood plain of the Indus River occupies nearly half of Pakistan’s area. Bounded by the Karakoram, Hindu Kush and Pamir mountain ranges to the north and the Balochistan Plateau to the west, the Indus and its tributaries flow southwards from the northern ranges to the Arabian Sea more than 1000km to the south (see Figure 4.1).

Having left the ranges, the river falls only a few hundred metres across this distance. Outside of the major cities of Lahore in the north and Karachi in the south, much of Pakistan’s population lives close to the Indus and Chenab rivers, farming wheat, cotton, rice and other crops to sustain its population of some 170 million. The 2010 monsoon season saw higher than usual rainfall (Figure 4.2 refers), contributing to floods which started affecting populated areas in the north.
4. Pakistan floods

Figure 4.1: Pakistan and the Indus-Chenab flood plain

in July and progressed southwards throughout the following months, remaining
in some areas throughout October and beyond.
Figure 4.2: Daily rainfall (mm) across the Pakistan region (Lat. 28N–35N, Long. 70E–74E) in July/August 2009 (above) and 2010 (below). Reproduced from GESDISC (2011).
4. Pakistan floods

4.3 Theoretical Basis

Fundamental choices of radar data rest on wavelength, polarisation configuration, spatial and temporal frequency. Radar signals will interact with their target in a manner dictated by structural, textural and dielectric properties of the target surface. Structural and textural properties are matters of scale, and must be considered in relation to the wavelength of the radar signal. When considering the detection of flood water, we are interested in the radar response from water itself, from the surrounding land cover and, where partial inundation occurs, a combination of the two. The radar response from each of these three categories is a complex combination of effects. In the case of vegetation, C-band radar (wavelength $\lambda = 3.75–7.5$ cm) will tend to interact with small branches and leaves, and L-band ($\lambda = 15–30$cm) with larger branches and trunks, each in a scale of order comparable to their own wavelength. Where partial inundation occurs, the radar signal may interact multiple times between the emergent structure (vegetation or buildings, for example), in a phenomenon known as dihedral scattering (or “double-bounce”), resulting in a very high return signal. The extent of this occurrence is dependent, therefore, on the relative scale of the emergent structure with respect to the wavelength of the signal. Where open water is found, if the surface is smooth, much of the radar signal is reflected away from the sensor, resulting in a low backscatter response. The extent of the return signal in this case has a sinusoidal relationship with the angle that the radar is incident to the surface of the water. However, where there are regular waves on the surface of the water, Bragg resonance can result in a very high return signal (e.g. Schaber et al. (1997)). The degree to which this occurs depends once again upon the relationship between the scale of the wave and the wavelength of the radar signal. C-band radar is sensitive mainly to small capillary waves and L-band to larger “chop”, as might be expected. The alignment of the waves with respect to the direction of the incident radar wave is also important here. The effect is greatest when the incident signal is orthogonal to the alignment of the wave (or parallel to the wind direction), and may not occur at all when the radar signal and the wave direction are parallel (Liebe et al., 2009). Finally, the strength of the radar signal returned to the sensor is reduced by an increasing incidence angle (Monsiváis et al., 2006).
Whilst simple geometry allows us to calculate the theoretical degree of this effect, its precise value is the combination of various target characteristics averaged over a pixel-space, and is therefore not readily known for a given time. The change in radar signal for a given degree increase in incidence angle is greater for a specular reflector such as smooth water, than for a diffuse or volume scatterer such as bare soil or vegetation. The particular environment encountered in Pakistan presents a further complication in the detection of flood waters using radar data. Desert areas which remain dry during the flood event absorb and attenuate microwave radiation (Robinson et al., 2006; Schaber et al., 1997), returning a low signal which encroaches into the range of that expected by open water. The difference with the desert response is its relative permanency, and so by deducting values from a GM image taken from the same orbit track prior to the flood, we are able to discern the water from the desert. Digital Numbers (DN) in ASAR detected products correspond to brightness amplitude. The radar backscatter coefficient $\sigma^0$ may be calculated from the DN values by:

$$\sigma^0 = \frac{DN^2}{K} \cdot \sin \alpha$$  \hspace{1cm} (4.1)

where $K$ is the absolute calibration constant (ESA, 2004). However, the received backscatter is further dependent on $\alpha$ by some function $F$ which is peculiar to the target environmental conditions (Baghdadi et al., 2001; Ulaby et al., 1982) such that

$$\sigma^0_0 = \sigma^0_\alpha \cdot F(\alpha)$$  \hspace{1cm} (4.2)

Given a reasonably close temporal separation of images, and in the absence of flooding, the environmental conditions, and therefore the nature of $F$, are similar for a given pixel in the target image to the corresponding pixel in the dry baseline offset.

Converting to decibels and deducting the base backscatter values $\sigma^0_0$ from the
target values $σ^0_1$ gives

$$\Delta σ^0 = 10 \left[ \log(σ^0_1) + \log(F(α)) - \log(σ^0_2) - \log(F(α)) \right]$$

(4.3)

Substituting 4.1 into 4.3 gives

$$\Delta σ^0 = 20 \cdot \log \left( \frac{DN_1}{DN_0} \right)$$

(4.4)

This assumes that the difference in $α$ between the two images is negligible. This difference between local incidence angles for the image pairs was found to have a mean value of -0.02°, with a standard deviation across the region of interest of 0.08°. Taking an extreme case of MEAN $- 3 \times$ STDDEV gives a difference of -0.26°, which would result in an error in the final difference image of 0.02 dB at 44° and 0.07 dB at 15°. Given the radiometric uncertainty in the GM data of 1.54–1.74 dB (ESA, 2007a), such differences may be disregarded for our purposes.

**4.3.1 Expected values**

![Key map to describe range of $Δσ^0$ values derived from image differencing process](image)

*Figure 4.3:* Key map to describe range of $Δσ^0$ values derived from image differencing process

The range of values encountered in the resultant $Δσ^0$ image are represented graphically in Fig. 4.3, and may be categorised as follows:
4. Pakistan floods

1. Areas with values common to both the target and the baseline image are shown in red (values close to zero), and include permanent water and desert, shown in Box 1, together with all other unchanged values between Boxes 4 and 5.

2. Would occur where water were present in the baseline image and not in the target image (which is unexpected). Such values are far more likely to represent the surface of desert or very dry radar-dark soils becoming wet, which greatly increases backscatter (Robinson et al., 2006). These values can also represent the occurrence of Bragg Resonance due to wind effects on permanent water (Schaber et al., 1997).

3. Flooding. Mid-high values in the baseline image have become mid-low.

4. These have values below the threshold which may therefore be rightly or wrongly classified as flood water. It is with the intent to capture such errors that the region-growing algorithm, making spacial association relevant to the classification decision process, is adopted here.

5. Mid-high baseline values which undergo a small-large increase in backscatter values between the baseline and target images. Again, this can represent a dry surface becoming wet, but could also encompass open flood waters where conditions are right for wind-induced Bragg resonance to return a high backscatter signal.

In the image differencing process, permanent water bodies, which are common to the baseline and target images, are removed. However, where water bodies are permanent and semi-permanent, their extents are easily mapped and overlain if required. We are interested here in mapping inundation outside of the current river coarse. The filling-in of permanent water after the flood classification and prior to testing would increase accuracy, but the results here are left as they are, with such limitations remaining exposed. Part of the reason for this is that the flood dynamics of the Pakistani rivers are complex, punctuated by sudden effects such as the breaching of the many levees which regulate their flow, and we therefore prefer in this instance to make no assumptions as to any possible deviations to the normal channel flow.
4.4 Method

4.4.1 Data acquisition
Data was acquired systematically via download from ESA’s Kiruna and ESRIN ground stations, made available in a two-week moving window through the Category 1 Fast Track Registration agreement.

4.4.2 Coverage
Area and frequency of cover of the region of interest was compared with MODIS Aqua and Terra data. In order to test comparative coverage, a mask was created by buffering the Indus river by 50km. For each day of August, the sum of pixels covered by GM data for that day were recorded along with the number of cloud-free pixels contained within the MODIS Terra and Aqua images, as a percentage of the total pixels for the masked region.

4.4.3 Image Preprocessing
Incidence angles ($\theta$), slant-range times (SRT) and geographical coordinates are provided within the raw GM data file, corresponding to tie points that form a grid with a spacing of 80–85 pixels across the image. Interpolation of $\theta$ and SRT was carried out first across the swath, and then the azimuth direction, to obtain values for each pixel. Terrain correction was also carried out in the frame of reference of the raw data file, due to the fact that columns and rows run parallel to the azimuth and swath respectively, making the geometry involved in the calculation much simpler. For this purpose, SRTM 7.5 arc-second Digital Elevation Model (DEM) data (Jarvis et al., 2008; Reuter et al., 2007) were projected into the local x–y coordinate system. The incidence angles $\theta$ and the DEM were then used to calculate local incidence angles ($\alpha$) for each pixel. Both the orthorectified Digital Number (DN) and $\alpha$ surfaces were then transformed to geographic coordinates by third order polynomial transformation.
4.4.4 Image Differencing

For each target image acquired over the study area, a matching image from the same orbit track was chosen as a baseline, being the latest available image covering all of the azimuth extent of the target and occurring prior to the commencement of the flood event. Details of the data used are shown in Table 4.1. The raw data files were registered in a database, at which time tie point data, including coordinates, incidence angles and slant-range times were extracted.

Table 4.1: GM Data used in this study. The Baseline Cycle refers to the orbit cycle corresponding to the deducted baseline data

<table>
<thead>
<tr>
<th></th>
<th>Date</th>
<th>Orbit Cycle</th>
<th>Orbit Track</th>
<th>Baseline Cycle</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>2010-04-01</td>
<td>88</td>
<td>134</td>
<td>-</td>
</tr>
<tr>
<td>2</td>
<td>2010-04-17</td>
<td>88</td>
<td>363</td>
<td>-</td>
</tr>
<tr>
<td>3</td>
<td>2010-05-13</td>
<td>89</td>
<td>234</td>
<td>-</td>
</tr>
<tr>
<td>4</td>
<td>2010-05-16</td>
<td>89</td>
<td>277</td>
<td>-</td>
</tr>
<tr>
<td>5</td>
<td>2010-05-26</td>
<td>89</td>
<td>420</td>
<td>-</td>
</tr>
<tr>
<td>6</td>
<td>2010-06-01</td>
<td>90</td>
<td>5</td>
<td>-</td>
</tr>
<tr>
<td>7</td>
<td>2010-06-06</td>
<td>90</td>
<td>84</td>
<td>-</td>
</tr>
<tr>
<td>8</td>
<td>2010-06-07</td>
<td>90</td>
<td>91</td>
<td>-</td>
</tr>
<tr>
<td>9</td>
<td>2010-06-12</td>
<td>90</td>
<td>170</td>
<td>-</td>
</tr>
<tr>
<td>10</td>
<td>2010-06-15</td>
<td>90</td>
<td>213</td>
<td>-</td>
</tr>
<tr>
<td>11</td>
<td>2010-06-17</td>
<td>90</td>
<td>234</td>
<td>89</td>
</tr>
<tr>
<td>12</td>
<td>2010-06-19</td>
<td>90</td>
<td>270</td>
<td>-</td>
</tr>
<tr>
<td>13</td>
<td>2010-06-20</td>
<td>90</td>
<td>277</td>
<td>89</td>
</tr>
<tr>
<td>14</td>
<td>2010-06-22</td>
<td>90</td>
<td>313</td>
<td>-</td>
</tr>
<tr>
<td>15</td>
<td>2010-06-23</td>
<td>90</td>
<td>320</td>
<td>-</td>
</tr>
<tr>
<td>16</td>
<td>2010-06-28</td>
<td>90</td>
<td>399</td>
<td>-</td>
</tr>
<tr>
<td>17</td>
<td>2010-07-01</td>
<td>90</td>
<td>442</td>
<td>-</td>
</tr>
<tr>
<td>18</td>
<td>2010-07-03</td>
<td>90</td>
<td>463</td>
<td>-</td>
</tr>
<tr>
<td>19</td>
<td>2010-07-04</td>
<td>90</td>
<td>485</td>
<td>-</td>
</tr>
<tr>
<td>20</td>
<td>2010-07-05</td>
<td>90</td>
<td>499</td>
<td>-</td>
</tr>
<tr>
<td>21</td>
<td>2010-07-09</td>
<td>91</td>
<td>48</td>
<td>-</td>
</tr>
<tr>
<td>22</td>
<td>2010-07-11</td>
<td>91</td>
<td>84</td>
<td>90</td>
</tr>
<tr>
<td>23</td>
<td>2010-07-12</td>
<td>91</td>
<td>91</td>
<td>90</td>
</tr>
<tr>
<td>24</td>
<td>2010-07-15</td>
<td>91</td>
<td>134</td>
<td>88</td>
</tr>
<tr>
<td>25</td>
<td>2010-07-17</td>
<td>91</td>
<td>170</td>
<td>90</td>
</tr>
</tbody>
</table>

Continued on next page
### Table 4.1: GM Data used in this study (Contd.)

<table>
<thead>
<tr>
<th>Date</th>
<th>Orbit</th>
<th>Orbit Track</th>
<th>Baseline Cycle</th>
</tr>
</thead>
<tbody>
<tr>
<td>2010-07-22</td>
<td>26</td>
<td>91</td>
<td>348</td>
</tr>
<tr>
<td>2010-07-24</td>
<td>27</td>
<td>91</td>
<td>270</td>
</tr>
<tr>
<td>2010-07-25</td>
<td>28</td>
<td>91</td>
<td>277</td>
</tr>
<tr>
<td>2010-07-27</td>
<td>29</td>
<td>91</td>
<td>313</td>
</tr>
<tr>
<td>2010-07-28</td>
<td>30</td>
<td>91</td>
<td>320</td>
</tr>
<tr>
<td>2010-07-31</td>
<td>31</td>
<td>91</td>
<td>363</td>
</tr>
<tr>
<td>2010-08-02</td>
<td>32</td>
<td>91</td>
<td>399</td>
</tr>
<tr>
<td>2010-08-04</td>
<td>33</td>
<td>91</td>
<td>420</td>
</tr>
<tr>
<td>2010-08-05</td>
<td>34</td>
<td>91</td>
<td>442</td>
</tr>
<tr>
<td>2010-08-07</td>
<td>35</td>
<td>91</td>
<td>463</td>
</tr>
<tr>
<td>2010-08-08</td>
<td>36</td>
<td>91</td>
<td>485</td>
</tr>
<tr>
<td>2010-08-09</td>
<td>37</td>
<td>91</td>
<td>499</td>
</tr>
<tr>
<td>2010-08-10</td>
<td>38</td>
<td>92</td>
<td>5</td>
</tr>
<tr>
<td>2010-08-13</td>
<td>39</td>
<td>92</td>
<td>48</td>
</tr>
<tr>
<td>2010-08-15</td>
<td>40</td>
<td>92</td>
<td>84</td>
</tr>
<tr>
<td>2010-08-16</td>
<td>41</td>
<td>92</td>
<td>91</td>
</tr>
<tr>
<td>2010-08-19</td>
<td>42</td>
<td>92</td>
<td>134</td>
</tr>
<tr>
<td>2010-08-21</td>
<td>43</td>
<td>92</td>
<td>170</td>
</tr>
<tr>
<td>2010-08-24</td>
<td>44</td>
<td>92</td>
<td>213</td>
</tr>
<tr>
<td>2010-08-26</td>
<td>45</td>
<td>92</td>
<td>234</td>
</tr>
<tr>
<td>2010-08-28</td>
<td>46</td>
<td>92</td>
<td>270</td>
</tr>
<tr>
<td>2010-08-29</td>
<td>47</td>
<td>92</td>
<td>277</td>
</tr>
<tr>
<td>2010-08-31</td>
<td>48</td>
<td>92</td>
<td>313</td>
</tr>
<tr>
<td>2010-09-01</td>
<td>49</td>
<td>92</td>
<td>320</td>
</tr>
<tr>
<td>2010-09-04</td>
<td>50</td>
<td>92</td>
<td>363</td>
</tr>
<tr>
<td>2010-09-06</td>
<td>51</td>
<td>92</td>
<td>399</td>
</tr>
<tr>
<td>2010-09-08</td>
<td>52</td>
<td>92</td>
<td>420</td>
</tr>
<tr>
<td>2010-09-09</td>
<td>53</td>
<td>92</td>
<td>442</td>
</tr>
<tr>
<td>2010-09-11</td>
<td>54</td>
<td>92</td>
<td>463</td>
</tr>
<tr>
<td>2010-09-12</td>
<td>55</td>
<td>92</td>
<td>485</td>
</tr>
<tr>
<td>2010-09-13</td>
<td>56</td>
<td>92</td>
<td>499</td>
</tr>
<tr>
<td>2010-09-14</td>
<td>57</td>
<td>93</td>
<td>5</td>
</tr>
<tr>
<td>2010-09-17</td>
<td>58</td>
<td>93</td>
<td>48</td>
</tr>
<tr>
<td>2010-09-20</td>
<td>59</td>
<td>93</td>
<td>91</td>
</tr>
<tr>
<td>2010-09-25</td>
<td>60</td>
<td>93</td>
<td>170</td>
</tr>
</tbody>
</table>

Continued on next page
4. Pakistan floods

<table>
<thead>
<tr>
<th>Table 4.1: GM Data used in this study (Contd.)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Date</td>
</tr>
<tr>
<td>61</td>
</tr>
<tr>
<td>62</td>
</tr>
<tr>
<td>63</td>
</tr>
<tr>
<td>64</td>
</tr>
<tr>
<td>65</td>
</tr>
<tr>
<td>66</td>
</tr>
</tbody>
</table>

4.4.5 Baseline datasets from MODIS

MODIS data were chosen from the results of the cloud-cover study described in Section 4.4.2, for three dates, with which to establish GM data classification thresholds and by which to gauge the accuracy of the classification process.

4.4.6 Thresholding and Classification

In order to obtain a binary map of the inundated regions, a region-growing function provided as part of the GRASS GIS package (GRASS Development Team, 2009) was used. The `r.lake` function is primarily intended to fill a lake to a target water level from a given start point, or seed. This starting point can be a set of coordinates, or a raster map in which the seed points are represented by non-null values. The function will grow a region, starting at the seed points, until a specified water level is reached, as determined by a given DEM. In our case, the seed was a rasterised line-type shape file of Pakistan’s river channels, the “DEM” was the $\Delta \sigma^0$ image, and the “water level” was set to the various thresholds tested. This method allows the use of a threshold value that is well inside the standard deviation of values for non-flooded areas, with the provision that the selected pixels are adjacent to other selected pixels as grown from the river channels. In order to try to mitigate errors of commission on the outskirts of the selected regions, a $3 \times 3$ modal neighbourhood filter was then applied to the binary classification.

MODIS band 6 data, representing Short Wave Infra-Red (SWIR) radiation ($\lambda$
= 1628–1652 nm), were used to map flooding and to establish thresholds to use with the radar images. Light in this short-wave infra-red waveband is absorbed by all but the most turbid water, and is therefore often used to map water (e.g. Dheeravath et al. (2010); Ordoyne & Friedl (2008)). Whilst MODIS reflectance bands provided by USGS attempt to achieve surface reflectance values, it is the case that attenuators such as thin cloud, that vary from image to image, preclude the use of a single absolute threshold applied to MODIS data in order to establish the benchmark. Fig. 4.4 shows the density plot of MODIS Band 6 reflectance values over the flooded region from two images, taken two days apart. There is a clear full-range displacement of reflectance values of 0.03–0.05.

Figure 4.4: Density plot of MODIS Band 6 reflectance values over the flooded region on 27 and 29 August 2010. The peaks at reflectance values of 0.05 and 0.1 represent water, as is seen later in the kappa analyses.

To account for this uncertainty, a bivariate sensitivity analysis was carried out, matching conformance of a range of radar backscatter thresholds in the radar images against a range of thresholds in the contemporaneous MODIS images. A peak in cross-correlation outside of the extreme threshold values (which would classify the whole image as flooded or non-flooded) would only likely represent common optimal flooded/non-flooded thresholds, as beyond the low signal response to water common to SWIR and radar data, the characteristic responses
of each of the wavebands are largely independent. In order to gauge the performance of classifications under varied thresholds, it is insufficient to simply determine the percentage of coincidence of allocated classes, as this gives a distorted result. If, for example, a flooded area comprises 5% of the region under study, then a classification omitting all the flooding would, with such a method of assessment, be 95% correct. For this reason, Cohen’s kappa statistic is often used as a “coefficient of agreement” between two classification processes (Cohen, 1960; Foody, 2006; Hunt et al., 2010; Tolpekin & Stein, 2009). The kappa statistic \( \kappa \) is calculated as

\[
\kappa = \frac{p_0 - p_c}{1 - p_c}
\]

where \( p_0 \) is the proportion of pixels in which agreement is observed and \( p_c \) is the theoretical proportion expected by chance selection (Cohen, 1960). It is the latter parameter which is perceived in certain instances to be problematic, as the observed proportion of allocation to each class is used as a basis to calculate random expectation (in effect, assuming that the decision process will always allocate the correct proportion of pixels to each class, whether the specific allocations are correct or not). This is not considered an issue in our case. Firstly, the assumption of proportion is observed to be approximately correct. Secondly, we are, while in pursuit of an optimal threshold, seeking a relative measure of classification accuracy rather than an absolute one.

Where good clear MODIS data were available during the flood event, \( \kappa \) was calculated for a matrix of classifications made with SWIR reflectance upper thresholds ranging between 0.05 and 0.25, and GM (difference image) \( \Delta \sigma^0 \) upper thresholds ranging between -10 and 0 dB. From this, the optimal backscatter thresholds could be observed, together with their sensitivity and inter-image variability. A comparison was made using the single target GM image, with a ground-range projected backscatter \( (\gamma) \) upper threshold range between -20 and 0 dB. A threshold to be used where MODIS data were unavailable was decided in this way, and its suitability assessed on a further series of \( \kappa \) analyses on another date for which MODIS data were available.
4. Pakistan Floods

4.5 Results

4.5.1 Coverage

During the 98 days between 11 July and 17 October 2010, the average repeat coverage period by GM data over the region studied was around 9 days (see Figure 4.5). Though this would be insufficient for a complete time series of the flood dynamics, it can feasibly serve to fill the gaps in information gained from optical sensors such as MODIS caused by the presence of cloud cover. Figure 4.6 shows a comparison of percentage of the full flood extent captured independently by MODIS Terra, MODIS Aqua and GM data, for each day of August 2010. Cloud cover limited the use of MODIS data through the first week of August, during the build-up of flood waters north of Sukkur, whilst there were sufficient GM data to build a picture of the flood extents at this time. Much of the rain that caused the floods in Pakistan fell on the ranges to the north, and as such, there were significant periods free of cloud further downstream where most of the catastrophic flooding occurred, and so in this respect, as a “dry flood”, this event enjoyed better coverage than most similar events with optical data, and in many other flood events, it may be reasonably assumed that the difference in availability of data could be far greater.

4.5.2 Image Differencing

It can be seen from the probability density functions shown in Figure 4.7 that attempts to map flooding using a simple threshold would result in large errors of commission and omission, due to the range of overlap of values. In the non-flooded areas, as we have chosen a region close to the known flooding, the slight rise in average backscatter value may be due in part to increased surface soil moisture (Pathe et al., 2009) in the vicinity of the flood and possibly from dihedral scattering from vegetation emergent from flood waters at the boundary of the flooded class regions (Hess et al., 2003). In addition to native vegetation close to the main river channel, wheat, cotton, onion, sunflower, rice, pulses and dates are all grown in the region (Ashraf & Majeed, 2006). The large standard deviation of $\Delta \sigma^0$ values in the non-flooded regions can be explained by a couple of factors.
Firstly the propensity of radar data to contain noise, most of which is speckle. This is characterised by high and low valued pixels whose values represent interference arising from the use of a coherent electromagnetic radiation source, rather than having anything to do with the target. Speckle can lead to sharp differences in values between any two radar images. Secondly, variations in the immediate recent rainfall history can cause large differences in backscatter value. Not only moist soil, but also wet vegetation tends to give a high backscatter response at C-band (e.g. Ulaby et al. (1982)). A drop in radar value in non-flooded areas can be seen where smooth specular reflecting alluvial sediments, for example, dry out. Low backscatter also occurs where the signal is absorbed / attenuated by very dry sand (Robinson et al., 2006; Schaber et al., 1997), though such a fall from relatively wet to very dry conditions necessary to produce a low $\Delta \sigma^0$ value are unlikely to have occurred so close to the main channel.
The distribution of $\Delta \sigma^0$ values in the flooded region is perhaps best explained in terms of what may be observed in Figure 4.8. The images centre on the segment of the Indus river running between Sukkur and Dadu, where its course changes from a south-west to a south-east direction. The image on the left shows radar backscatter values acquired on 20 August 2010. The image on the right shows the same values, with those of a previous cycle deducted. The regions labelled D and F represent sections of the river characterised by a large flood channel superimposed with the meandering and anabranching main Indus channel (see Figure 4.9). At the time of acquisition, this large channel was completely flooded, and appears as radar dark in the first image. However, due to the fact that
alluvial sediment can also act as a specular reflector in the same way as water, much of the D region is punctuated with mid-range value pixels in the difference image, and region F is all but indistinguishable from non-flooded land. The large area at A with low backscatter values in the first image shows that part of the lowlands which protrudes into the Sulaiman mountains of Balochistan, comprising mainly the districts of Bolan and Sibi. This region is normally dry, with an annual rainfall of 200–250 mm. The low backscatter is considered to be the result of attenuation and absorption of the signal, rather than of specular reflection. The low backscatter values are clearly offset in the difference image, leaving only those low values representing the rivers that run west below A and then south towards Lake Manchar below B. Similarly, the dark region in the first image at C is at the western edge of the Thar Desert. The bright strip running north-south immediately to the right of C shows the relatively high backscatter from the vegetation bordering the Nara Canal and its irrigated hinterland. As with A, the response of both the irrigated strip and the desert are common to the consecutive orbit cycles, and hence do not appear in the difference image.
Figure 4.8: The region between Jacobabad and Nawabshah in mid August 2010. The image on the left shows backscatter values in decibels. Smooth open water is commonly represented by values of around -16dB or below. The image on the right shows the same data, with the values from the previous cycle along the same orbit track having been deducted. A better understanding of the true extent of flooding can be discerned by a difference of around -4dB in this image.
Figure 4.9: Landsat composite colour image of the Indus and its floodplain southwest of Sukkur. The Nara canal is seen running north–south to the right of the image
4.5.3 Thresholding and Classification

A comparison of value profiles of $\Delta \sigma^0$ and MODIS Band 6 across a section of the flooded Indus, as at 10 August 2010, is shown in Figure 4.10. It can be seen that, at this scale, the choice of threshold of Band 6 to classify water is not particularly sensitive between around 0.2 and 0.15 units, where the profile crosses the flooded area, with relatively few pixels taking intermediate values. There is little doubt that where SWIR reflectance values fall close to zero on all legs, there is open water. In these areas along legs 1, 2 and 3, the corresponding $\Delta \sigma^0$ values fall below -2 dB, corresponding to a fall in backscatter values caused by increased specular reflection, due to the increased presence of water. Along legs 4 and 5, however, there are large fluctuations of $\Delta \sigma^0$ values. This is mainly due to the fact that, as mentioned before, the alluvial sediment can also act as a specular reflector in the same way as water, thus the dry baseline low pixel values are offset from the target image, producing mid-range difference values. Areas where a slight rise in SWIR reflectance coincides with a sharp rise in $\Delta \sigma^0$ (such as at 40km and 120km on the x-axis) are believed to represent partial inundation with emergent vegetation, the high $\Delta \sigma^0$ values being the result of dihedral scatter.

These profiles demonstrate the volatility of $\Delta \sigma^0$ values, especially in those areas that show a low backscatter response under non-flooded conditions, as discussed above. It was found that the choice of a simple $\Delta \sigma^0$ threshold to suit conditions in the main river channel would result in many regions mapped incorrectly as flooding in areas well away from the river channels. For this reason it was decided to make contiguousness with other flooded pixels adjacent to the river channels a condition of the flood class, in addition to the satisfaction of the $\Delta \sigma^0$ threshold. Therefore flooded regions were mapped by growing contiguous areas that satisfied the threshold criterion from pixels at the centre of the river channel outwards, using the technique described in Section 4.4.6.

4.5.3.1 Bivariate sensitivity analysis to determine threshold

$\kappa$ statistic values calculated in the sensitivity analysis described in section 4.4.6 are shown in Figs. 4.11, 4.12 and 4.13. It can be seen that, while the optimal SWIR reflectance threshold varies between the dates, the optimal $\Delta \sigma^0$ threshold
of around -2dB is common to the three instances. The reasons for the differences in the MODIS thresholds was discussed in section 4.4.6 and the difference between the optimal reflectance thresholds of 0.07 and 0.11 on 27 and 29 August respectively are manifest in the shift in distribution of values between the two MODIS images that was shown in Fig. 4.4.

Fig. 4.14 shows flood extent estimates from MODIS (top), the single contemporaneous GM image (centre) and the GM Difference image (bottom), each using thresholds optimized from the process described above.

With the single image in the centre, there are two processes resulting in the low backscatter response. To the north-west of the dashed line, the low response is dominated by specular reflection from the surface of flood waters. To the south-east of the dashed line, the low backscatter response is caused by absorption in desert sands. The wrongly classified desert area is eliminated in the third image, as this low response from the desert areas is common to both the target image.
and its baseline partner, and is therefore subtracted out.

When comparing the MODIS (top) image with the Difference image (bottom), it can be seen that whilst the boundaries of the flood are well defined, areas of permanent water or radar-dark flood plain regions are also eliminated. In the central image derived from the single GM data set, such areas which do fall within the flooded region are more completely defined. With a priori knowledge of terrain and environmental conditions, one can mask out desert areas and achieve a more accurate classification using the single image. Masking must be very precise, however, as some absorption areas can lie extremely close to the flooded region, as can be seen from the area encircled in red in the middle image. Assuming sufficient information is available within the time frame allowed, the higher accuracy which may be achieved by such masking is demonstrated in Fig. 4.15, where a $\kappa$ value of 0.7 is achieved. Note that the precedence of MODIS SWIR thresholds matches that seen in Fig. 4.12, as expected.

Where a fast indicator of the extent of flooding through otherwise dry land is urgently required, we propose that the image differencing technique offers a
Figure 4.12: $\kappa$ statistic calculated for individual classifications of flooding on August 27, 2010, based upon $\Delta \sigma^0$ thresholds ranging from -10 to 0 dB. The series represents corresponding MODIS Band 6 reflectance thresholds used in the reference image, ranging from 0.05 to 0.25.

reasonably stable means to identify those extents between periods where optical data are unavailable, enabling a broad scale view of the flood dynamics with a better temporal resolution than could otherwise be achieved.
Figure 4.13: $\kappa$ statistic calculated for individual classifications of flooding on August 29, 2010, based upon $\Delta \sigma^0$ thresholds ranging from -10 to 0 dB. The series represents corresponding MODIS Band 6 reflectance thresholds used in the reference image, ranging from 0.05 to 0.25
Figure 4.14: Flood extent estimates from MODIS (top, $\kappa = 1$), the single contemporaneous GM image (centre, $\kappa = 0.3$) and the GM Difference image (bottom, $\kappa = 0.6$), using optimised thresholds.


4. Pakistan floods

Figure 4.15: $\kappa$ statistic calculated for individual classifications of flooding on August 27, 2010, based upon $\gamma$ thresholds ranging from -20 to 0 dB, following precise masking of radar-dark dry land established from the image differencing process. The series represents corresponding MODIS Band 6 reflectance thresholds used in the reference image, ranging from 0.05 to 0.25.
4.5.4 Inundation Dynamics

Figure 4.16 shows instances from the resultant time series of binary flood maps. The first two images show the build up of the upper reaches of the Indus. By August 7, the Chenab has flooded and the main flood has reached Kashmore. The image at August 12 shows the situation following the breaching of a bund at Thori in Kashmore. By August 20 the flood has reached the Hyderabad district. The August 29 image shows the results of two significant breaches, one at Sukkur, allowing the flood to split and inundate the Jacobabad region to the north of the Indus, and another at Sarjani to the south, where part of a dyke collapsed on August 26. This resulted in the extensive flooding in the south which is evident on September 11. The final two images, set three weeks apart, show the flooded region covering over 7,000 km² between Jacobabad and the Manchar Lake in Dadu, which remained for many weeks. Beyond 20 October 2010, there followed a period in which the Envisat satellite underwent a scheduled program change, during which time GM data was unavailable.

Figure 4.17 shows the duration of flooding over the full extents, up to 17 October 2010, as derived from the GM data. Many regions remained inundated for several weeks, the greatest duration being observed in the area around Jacobabad described above. The greatest flood duration shown at 97 days represents the enlarged Lake Manchar to the southwest of this region.

An idea of the propagation speed of flood waves can be gained from Figure 4.18, which shows the distance along the Indus river channel of the head of the main flood, and the receding tail end. The initial advance covered some 500km in 5 days (about 4 km per hour), with the flood reaching the southern extents towards the end of the first week in September. The greatest length of flooding occurs where the recession curve is flattest at around 21 August. The recession rate is seen to increase at the end of August, following the bund breaches at Sukkur and Sarjani, flattening off once more in early September, when the front of the flooding is seen to retreat back to localised areas.
Figure 4.16: Selected instances from the time series showing the build-up of flooding and much of its recession. Flooding is still evident in the third week of October 2010, at which time data temporarily ceased to be available, due to ESA’s scheduled preparations for Envisat’s project extension program.
Figure 4.17: Map showing the extent and duration of inundation surrounding the Indus and Chenab rivers as derived from satellite radar data acquired between July and October 2010.
Figure 4.18: Distance of the flood head and tail along the Indus channel from the foot of the northern ranges at 71° N, 32° S

Figure 4.19 shows the flooded area over the time series, derived from the combination of GM and MODIS data. The lower curve shows only the flooding around the main Indus Channel. The upper curve shows the total area, including the near-static flooding between Jacobabad and Dadu, which remained well into October 2010.

4.5.5 Accuracy

A measure of the classification accuracy of this method was known throughout from the $\kappa$ statistic values used to ascertain the optimal thresholds to use. The consequence of fixing a $\Delta \sigma^0$ threshold of -2 dB based on the $\kappa$ tests done on 10 and 27 August was further tested on the classification done for 29 August, the results of which are tabulated in Table 4.2 and shown graphically in Figure 4.20. The $\kappa$ statistics over the full range of MODIS Band 6 reflectance values and $\Delta \sigma^0$ values were also seen in Fig. 4.13.
4. Pakistan floods

Figure 4.19: Area of inundation over time, of the Indus Channel and the total flood.

The following factors are considered to be the main contributors to inaccuracy:

- As discussed in Section 4.5.2, much of the flooded region covers the immediate flood plain which ordinarily contains large meanders, anabranching and ox bow lakes. Low backscatter returns from these semi-permanent water bodies contribute to a lower value in a larger area of the baseline image when averaged to a pixel size of 500m. This is offset from similar values in the target image, resulting in mid-range values, incorrectly interpreted as land.

- Semi-submerged vegetation can cause high backscatter values due to multi-hedral scattering as discussed, which, again, when averaged with low values from open water may return a mid-range value.

- Wind conditions may be such that Bragg Scattering occurs, causing relatively high return values. Determination of the extent of this effect would
require detailed wind speed and direction data, which were not available for this period.

4.6 Discussion

4.6.1 Natural disaster response

For regions which face a high risk of flooding that may be ever increasing (Schiermeier, 2011), mitigating the impact of flooding can fall within two broad categories: planning and organisation based on predicted scenarios, and reactive response during and after an event. Action under the first of these requires an
4. Pakistan floods

Table 4.2: Error matrix and $\kappa$ statistic for the flood map on 29 August 2010 when compared with MODIS flood classification.

<table>
<thead>
<tr>
<th>Category</th>
<th>MODIS</th>
<th>Row Sum</th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Flooded</td>
<td>Non-Flooded</td>
<td>Col Sum</td>
</tr>
<tr>
<td>$\Delta$ Flooded</td>
<td>34761</td>
<td>10464</td>
<td>45225</td>
</tr>
<tr>
<td>$\sigma^0$ Non-Flooded</td>
<td>18662</td>
<td>228902</td>
<td>247564</td>
</tr>
<tr>
<td>Col Sum</td>
<td>53423</td>
<td>239366</td>
<td>292789</td>
</tr>
</tbody>
</table>

Cats | % Commission | % Omission | Est. $\kappa$ |
Flooded | 23.1 | 34.9 | 0.71 |
Non-Flooded | 7.5 | 4.4 | 0.59 |

$\kappa$ $\kappa$ Variance
0.65 0.000004

Obs Correct | Total Obs | % Observed Correct |
263663 | 292789 | 90.1 |

understanding of processes which govern the magnitude and extent of possible floods. Such an understanding cannot rely exclusively on historical data where land use and climate are changing, but must instead require predictive modelling.

Remote sensing has played an increasingly important role in this process in recent years. The establishment of parameters in hydrological models have called upon, for example, leaf area index calculations or on surface water extents using data from optical sensors such as Landsat Thermal Mapper (Chen et al., 2004; Milzow et al., 2009a; Stisen et al., 2008) and have increasingly incorporated soil moisture values gained from Envisat ASAR in such modelling (Decharme et al., 2009; Liu et al., 2010; Saux-Picart et al., 2009). Radar-derived DEM’s have been assessed and used in hydrological modelling widely (e.g. Ludwig & Schneider (2006)). More directly, hydraulic processes involved in flooding have been modelled to estimate flood magnitudes (Conesa-Garcia et al., 2010; Hostache et al., 2009b) and to develop flood inundation models (Schumann et al., 2007). Coupled models using SAR data have been employed in the last few years to useful effect. Montanari et al. (2009) investigate the usefulness of SAR data to gauge flood extents and stage heights in deriving soil saturation values. Milzow et al.
4. Pakistan floods

(2009b) seek to verify hydrological models by comparing simulated flood patterns with flood maps derived using AVHRR and ASAR data. Pauwels et al. (2009) calculate soil hydraulic conductivity values through a combination of SAR-based moisture maps and land surface modelling.

Predictive modelling has significant limitations in certain instances. An example of this can be observed in the case of the floods in Pakistan. There are limitations to ascertaining water volumes for the modelling process in large areas of very low gradient, necessitating a spatially and radiometrically high resolution DEM (Sanyal & Lu, 2004). Further to this, the use of levees on a large and small scale is widespread throughout the floodplain, many of which are built “privately” and therefore remain unmapped \(^1\). In this case, therefore, the importance of flood mapping based on observation, increases. So too does the significance of data availability, speed of acquisition, spatial coverage and temporal resolution.

4.6.2 Use and limitations of the GM data for flood mapping

The swath width of GM data is \(\sim 405\) km and permits a synoptic assessment of large flood events at the basin scale. Capturing the onset of a flooding event as early as possible is critical for emergency response. ENVISAT ASAR GM was one of the few sensors capable of capturing the full extent of the flooding in Pakistan during the first week and a half (e.g. MODIS data were unavailable from 2 to 9 August 2010 due to high cloud cover). ENVISAT ASAR GM acquisition is not systematic but will depend on other modes being switched off. Hence the coverage of a particular region can be variable. Across Asia, we found from September 2009 to May 2011 an average of 2–3 weekly observations at a single location. Over the same period there were no GM data acquired over New Zealand and 8 per week in parts of North America. Average frequency of land coverage per week by GM data over this 600 day period is shown in Figure 4.21. The temporal distribution is not evenly spread. South-east Asia, for example, received virtually no coverage for the first four months of the study period.

\(^1\)see http://tribune.com.pk/story/219602/private-dykes-on-public-land-may-lead-to-another-bout-of-floods/, for example
4. Pakistan floods

In Pakistan, we found the coverage of the ENVISAT ASAR Global Mode data was adequate to capture the dynamics of the propagation of the 2010 flood across the entire Indus Basin (Figs. 4.5 and 4.6). However, a lower acquisition frequency from 17 August 2010 onwards only allows for partial coverage of the recession of the flood.

A major limitation of the ASAR GM mapping technique used here is that inundated area with emergent vegetation can be confused with land area of high soil moisture. Wind-induced waves can also generate a roughening of the water surface which increases the scattering of the radar signal due to Bragg resonance; a phenomenon that is more pronounced in C than L-band over inland water bodies (Alpers, 1985; Alsdorf et al., 2007). Finally, partial flooding inside a pixel can be a common feature at this scale (spatial resolution of ∼1 km) especially along braided channels and will result in mixed pixels composed of land, water and flooded vegetation, which can return a wide range of signals.

The original mapping presented here was carried out in response to a request for a timely indication of the extent of flooding by a UN emergency response team when the event was in full throw, when expediency and simplicity competed with sophistication of technique as priorities. The methods lend themselves to seek greater accuracy by further analysis, building on thresholding and region-growing techniques by, for example, Galland et al. (2009); Matgen et al. (2011); Yu & Clausi (2007) and Silveira & Heleno (2009). Other successful segmentation methods for SAR images involving texture and shape (van der Werff & van der Meer, 2007), active contours (Ben Ayed et al., 2005; Chakraborty et al., 2009;
4. Pakistan floods

Fu et al., 2008) and multi-objective algorithms (Collins & Kopp, 2008) may be suitable. However, it is felt that the basic premise of same-track image differencing (to mitigate incidence-angle effects and ambiguous low-backscatter response due to absorption), coupled with a robust region-growing segmentation technique (e.g. Matgen et al. (2011)) to account for the small inter-modal range in the probability density functions of flooded and non-flooded areas, is well suited to map flooding in arid regions using SAR data.

### 4.6.3 A complement to other mapping techniques

Optical sensors, such as Landsat TM and MODIS, can easily detect open water using the strong absorption of solar energy by water in the near and middle infra-red. Shallow depths and turbid waters, are better detected at greater wavelengths (> 1µm; short-wave infra-red) where the illumination of the suspended materials or of the shallow bottom of a water column is considerably reduced (Bukata, 2005; Li et al., 2003). However during storm events the use of optical data can be severely limited by cloud cover. Radar imaging is less affected by cloud cover and can penetrate vegetation at a depth which depends on the wavelength used and the structure (density and height) of the vegetation (Alsdorf et al., 2007; Hess et al., 2003; Martinez & Le Toan, 2007; Rosenqvist et al., 2007). The use of L band data from the JERS and ALOS PALSAR sensors for flood monitoring is mainly restricted by acquisition times and limited archives, rather than by weather or vegetation condition. ALOS PALSAR data in the wide swath mode are particularly attractive to cover large regions, but unfortunately the system failed in April 2011 and this resource is therefore no longer available for data beyond that time\(^1\). Passive microwave data (e.g. SSM/I and ISCCP) are helpful for delineating inundated areas (e.g., Hamilton et al. (2002); Sippel et al. (1998)), in particular when used in conjunction with other sensors to limit confounding factors such as atmospheric condition and vegetation (Prigent et al., 2007), but their use in natural disaster response is limited by their low spatial resolution (tens of km). The geosynchronous weather satellites (e.g. Meteosat II, GOES, GMS) may often be able to bypass clouds with their high temporal resolution allowing

---

\(^1\)http://www.palsar.ersdac.or.jp/e/
4. Pakistan floods

for the mapping of open water at ~1km spatial resolution. In the specific case of
flooding events occurring in semi-arid and arid regions, such as the Pakistan flood
studied here, water under flooded vegetation can also be mapped using composite
data from the thermal bands of these weather satellites (Leblanc et al., 2003,
2011).

Amongst the most promising potential developments in remote sensing of
surface water is the future Surface Water and Ocean Topography (SWOT) mis-
sion. It is currently planned to be launched in 2020 and will provide significant
improvements in our abilities to map inundated areas from space 1. Using wide-
swath altimetry technology, SWOT will provide temporal and spatial variations
in water volumes stored in rivers, lakes, and wetlands at unprecedented resolution
(Biancamaria et al., 2010). SWOT will generate a global 3D mapping of all ter-
restrial water bodies whose surface area exceeds 250 m² and rivers whose width
exceeds 100 m (Biancamaria et al., 2010). The principal instrument of SWOT
will be a Kα-band Radar Interferometer (KaRIN), which will provide heights and
co-registered all-weather imagery of water over 2 swaths, each 60 km wide, with
an expected precision of 1 cm/km for water slopes, and absolute height level pre-
cision of 10 cm/km². ESA will also be extending and improving C-band SAR
capabilities with the launch of the Sentinel-1 system, expected in 2013. This pair
of satellites is planned to provide data with a spatial resolution of 20 m, with a
revisit time of between 1 and 3 days for Europe and Canada 2.

4.6.4 Other applications of GM data

Space borne technologies are increasingly found to be a key source of information
for wetlands conservation and management, as many of the World’s wetlands have
insufficient on-ground data in part due to their size, number and limited accessi-
bility (Jones et al., 2009). Even at such a time when technological advancement
in data processing, storage and communication enables ever higher spatial reso-
lutions from airborne and satellite sensors, the use of coarser resolution data still
has a very firm place in the remote sensing field where broad-scale monitoring

1http://swot.jpl.nasa.gov/mission/
2http://www.esa.int/esaLP/SEMBRS4KXMF_LPgmes_0.html
4. Pakistan floods

is required, such as the monitoring of algal blooms (Ahn & Shanmugam, 2006), assessing risks of fire (A. & R., 2008; Chéret & Denux, 2007) or drought (Rojas et al., 2011), the development of land surface models (Jarlan et al., 2005), the assessment of animal stocking rates (Hunt & Miyake, 2006) and the mapping of shorelines (M.-Muslim et al., 2007). The perceived role of GM data was primarily in the monitoring of sea ice (ESA, 2007b), to which field it has indeed contributed (e.g. Quincey & Luckman (2009)). However, its coverage and availability have already been identified as useful advantages in other areas, and have been put to good use, particularly in the areas of relative soil moisture (Bartsch et al., 2008, 2009), surface soil wetness (Pathe et al., 2009; Scipal et al., 2005) and in wetness dynamics (Scipal et al., 2005). The implication of GM data’s sensitivity to surface wetness has recently led to the interesting extension to its potential use in the monitoring of freeze-thaw cycles in permafrost regions (Park et al., 2011).

The need for coarse scale inundation mapping has been identified and acted upon, resulting in, for example, the 25 km resolution surface water product described by Schroeder et al. (2010) which tests favourably against finer resolution products, and the 0.25° global inundation map produced by Prigent et al. (2007).

This study aims to contribute to this latter group at a resolution that falls between those global-oriented scales and the finer river-channel-scale resolutions that can be achieved with other sensor-mode configurations such as JERS, ALOS PALSAR or ASAR in its finer modes. It is felt that ENVISAT ASAR GM data could be used to monitor inundation patterns over large wetlands in complement to estimates from other sensors (e.g. Sakamoto et al. (2007)).

4.7 Conclusions

It is clear that during periods of cloud cover, in which optical satellite data with which to map a flood event are not available, GM data may be available to varying degrees that cover the region of interest. Much of the rain (and the cloud) that affected the flooding in Pakistan occurred in the mountains away from the flood plain, and so it is fair to say that the relative availability of GM data with optical data may be greater in flood events closer to the precipitation source, making the use of GM data of greater value.
Ambiguity resulting from low backscatter values from non-flooded areas have been shown to have been reduced greatly by the image differencing process, as these low backscatter values are reasonably consistent between orbit cycles. A greater challenge is presented by the ambiguities in origin of data values, where effects such as dihedral scattering, Bragg resonance and speckle can raise and lower pixel values and prohibit the accurate classification of water. Where the objective is to ascertain the extent of flooding in near real time, there is little that can be done about Bragg resonance unless precise wind conditions are known, other than to hope that temporal frequency of data is sufficient to allow us to understand where this effect may have occurred and to rectify it in an updated image. The effects of speckle can be reduced with filtering. Dihedral scattering can, to some extent, be managed by acceptance that only open water is being mapped, or by further analysis using textural measures. Where, for example, dihedral scattering is dominant in a pixel, we expect the $\Delta \sigma^0$ value to be high, and would expect such areas of partial inundation to surround areas of total inundation. Further analysis could therefore encompass such regions into the flooded class and improve the overall accuracy. The coarse spatial resolution of GM data compounds all of the above problems, where adjacent regions of low and high backscatter values return an averaged mid-range value.

It has been shown, however, that a reasonable level of overall accuracy can be achieved using GM data which allows an understanding of the dynamics and broad-scale extents of a large flood during periods when there are no other means by which to judge these parameters. In the interests of flood mitigation planning, where thousands of lives are at stake, we feel that the potential use of GM data for this purpose is significant.
Chapter 5

Effects of vegetation on mapping of floods using satellite radar data

Chapter context
Abstract

The viability of mapping the extents of flooding in two rivers in northern Queensland, Australia, using Envisat ASAR Global Monitoring Mode (GM) data, is examined, through observations of a major flood in each. In the Flinders River event, the flooding, which covered over 9,000 square kilometres, can clearly be distinguished, from low backscatter values in the case of total inundation, to high backscatter where partial inundation occurs. The Staaten flood event demonstrates neither of these characteristics. The resulting classification of the Flinders flood achieves a kappa statistic of 0.7, with the region of lowest accuracy being topographically distinct, and more similar to that of the Staaten region. Water height loggers were set up in the lower Mitchell flood plain, neighbouring the Staaten, and the GM data response monitored throughout a year, during a quarter of which time the vegetation was partially inundated. It was found that extensive flooding under local conditions resulted in a barely perceptible drop in backscatter values, due to the mixing of high and low values, confining the use of GM data for the monitoring of floods to regions with particular surface conditions.

5.1 Introduction

The use of radar as a consistent and reliable tool for the mapping of floods has been shown to be complicated, and requires an understanding of when to expect, under particular environmental conditions, a low backscatter response due to specular reflection on the surface of water, and a high backscatter response due to dihedral scattering (double-bounce) (e.g. Bartsch et al. (2008); Costa & Telmer (2006); Grings et al. (2009); Kasischke et al. (2003)). To further complicate the process, open water can exhibit Bragg Resonance, where coherent additive effects return very bright signals from targets with regular textural qualities, such as waves on the surface of the water (Alsdorf et al., 2007; Liebe et al., 2009; O’Grady et al., 2011; Schaber et al., 1997), or furrows in a ploughed field in a particular orientation (Bonn & Dixon, 2005; Sanyal & Lu, 2004). While such limiting phenomena can occur in any environment under study, there are
some ground-cover characteristics which render a region particularly prone to ambiguous radar signals, and therefore preclude the use of radar of particular specifications for the mapping of flood extents. It is useful to understand what some of these conditions may be in the context of a particular data product, to avoid abortive work and erroneous inundation maps. Here we look at two large floods in near-neighbouring catchments in the Gulf Plains of north-west Queensland, observing a major difference in our ability to map their extents using GM radar data, and examine the likely reasons. From this, it may be possible to determine those areas that are off-limits to the use of GM data for flood mapping, and also help in our understanding of the particular radar characteristics of grasslands and tropical savanna.

5.2 Study Area and Flood Events

![Figure 5.1: Catchments under study in Northern Queensland, Australia](image)
5. Vegetation effects

5.2.1 Flinders

The Flinders Basin occupies an area of nearly 110,000 square kilometres of the northern Queensland Gulf Plains bioregion, draining into the Gulf of Carpentaria. The lower floodplain is dominated by grasslands, in particular Dichanthium, or Bluegrass. Tussock grasses such as these occur in clay soils through which water does not readily penetrate, resulting in either water logging or an absence of moisture, depending on rainfall (Spessa et al., 2005). Such conditions are not favourable to trees, and this region is therefore dominated by open grassland plains. In February 2009, a period of sustained rainfall resulted in the flooding of the lower basin, the inundated area covering over 9,000 square kilometres. Figure 5.2 shows the rainfall average for the whole basin, leading up to the flood event. The largest extent of the flood occurred following the final large rainfall peak in mid-February. At this time the gauge height at Walkers Bend, situated downstream of the largest flooded area, was on the decline from its peak at 15m (Fig. 5.2). Following 15 February, the gauge height increased as the upstream flood evacuated.

![Figure 5.2: Flinders basin rainfall (blue) and river height (red) at Walker’s Bend gauging station on the Flinders river in early 2009](image)

Figure 5.3 shows a MODIS Band 6 reflectance image (left) of the flooded
5. Vegetation effects

region on 15 February 2009, in which the low values representing absorption of the short-wave infra-red radiation by water appear as dark. The image on the right shows GM data acquired on the same day, with radar backscatter values ranging from close to zero dB (bright) to -25 dB (dark), the latter representing the specular reflection of the radar signal on open water. It is clear that the main body of flood water in the centre of the images has been captured in the GM data, along with the coastal wetlands of the Southern Gulf Aggregation to the north-west of the images. However, the fan of laden tributaries running into the Norman River to the east of the MODIS image is less clearly defined in the GM data. The bright patch in the GM image to the west of the gulf coast is attributed to Bragg Resonance.

Figure 5.3: ASAR GM image of flooding in the Flinders catchment on 15 Feb 2009 (left), compared with an image derived from MODIS Band 6 reflectance values from the same date (right)

5.2.2 Staaten

Covering an area of 25,572 square kilometres, the Staaten Basin ranges from the Bulleringa National Park in central Cape York to the west coast of the cape, narrowing as it reaches the Gulf of Carpentaria. The basin comprises ephemeral
5. Vegetation effects

Riverine channels in an alluvial plain dissected with many streams, oxbow lakes and billabongs, or drainage depressions, in widely spaced valleys that form a uniform pattern shared with the neighbouring Mitchell and Gilbert basins.

The tropical savanna woodland of the Staaten region comprises tall dense grasses, together with various densities of mainly Eucalypt trees, predominantly Melaleuca.

A continually wet January in 2007 preceded a significant rainfall event on 7 February (see Figure 5.4). The gauge at Dorunda peaked at around 7.2m on 10 February, at which time a substantial part of the lower reaches of the Staaten basin were flooded.

![Graph showing rainfall and river height at Dorunda gauging station on the Staaten river in early 2007.](image)

Figure 5.4: Staaten basin rainfall (blue) and river height (red) at Dorunda gauging station on the Staaten river in early 2007

The extent of the flood at this time in the Staaten Basin can be seen in the MODIS Band 6 image in the left of Figure 5.5. From the image on the right, it can be seen that the only open water visible from the GM data is along the coast, which represents the semi-permanent wetlands of the coastal South-east Karumba Plain Aggregation. The inability to distinguish the presence of flooding is representative of all the GM radar data seen during this event.

The relative separability of water and land between the two instances is high-
5. Vegetation effects

Figure 5.5: (left) Flooding (in black) of the Staaten river on 10 Feb 2007, derived from MODIS Band 6 reflectance values, and (right) the near-simultaneously acquired GM backscatter image lighted in the density plots shown in Figure 5.6. Ground-projected values, normalised to an incidence angle of $30^\circ$ ($\gamma_{30}$) for flooded regions are represented by the blue lines, non-flooded by red. The Flinders plot on the left shows that $\gamma_{30}$ values for the flooded region appear to have at least two means, one at around -15 dB and another at around -10 dB, the latter made up largely of those pixels in the vicinity of the Norman River alluvial fan already discussed, that is susceptible to misclassification. The Staaten plot on the right shows that, beyond the tail of low values returned from the coastal wetlands, the flooded region shows a distribution around a mean that is perhaps a decibel higher than that of the non-flooded land.

The Staaten and the Flinders flooded regions are both very flat, and GM data was acquired at similar angles of incidence, and so it must be concluded that the vastly different radar responses represent differences in the nature of the vegetation, in the nature and pattern of the floods themselves, or a combination of the two. The positive displacement of the flooded density plot for the Staaten Basin indicates a higher set of backscatter values for the flooded region than
5. Vegetation effects

Figure 5.6: Backscatter value density plots of the Flinders (left) and Staaten (right) regions, during their respective flood events. Values from pixels representing flood waters, as identified in MODIS SWIR data, give the blue curve, non-flooded pixel values the red.

the surrounding land. We know that the C-Band signal used to produce GM data is liable to interact with structures at the physical scale of leaves and small branches of trees, and that the Staaten and the neighbouring basins do experience floods which reach the lower tree crown (see Figure 5.7). Flooding of this extent, which completely covers the grasses, tends to return a low value due to specular reflection and to attenuation and volume scattering from the upper branches.

Figure 5.7: Part of the Mitchell River during a flood (Source: Kowanyama Aboriginal Land and Natural Resources Management Office)

Where flood levels are much lower, the radar signal is liable to interact with grasses, and a flood level rising through the height of the grass and then covering it should be characterised by an increase in backscatter due to dihedral scattering, followed by a sharp drop in value at total inundation. Grasses form the tallest
stratum growth form of the Flinders flooded region.

5.3 **Theoretical Basis**

As has been seen, the two regions under study have different land cover regimes. Both are extensively covered by grasses—in the case of the Flinders flood region, these are dominated by Dichanthium Sericeum, or Bluegrass (see Figure 5.8), the stems and leaves of which stand erect and rarely reach 30 cm in height (FAO, 2012). The major contrast in the Staaten region is the woody over-storey of ironwood / eucalypt species, with various grass species (Themeda Triandra, Heteropogon Contortus, Heteropogon Triticeus, Sorghum Plumosum, Chrysopogon Fallax, Alloteropsis Semialata, Eriachne Obtusa) which can grow up to anywhere between 70 and 300 cm tall, depending on phenology and fire history (see Figure 5.9).

![Figure 5.8: Dichanthium Sericeum, < 30 cm (Image: D. Greig, Source: ANH (2012)](Image: D. Greig, Source: ANH (2012)

![Figure 5.9: Sorghum Plumosum, 70–300 cm (Image: L. Wallis, Source: GBIF (2012)](Image: L. Wallis, Source: GBIF (2012)

The characteristics of radar backscatter reflected from a surface comprising water with emergent vegetation has been extensively studied (Costa & Telmer, 2006; Grings et al., 2009; Henderson & Lewis, 2008; Hess et al., 2003; Noernberg
5. Vegetation effects

et al., 1999; Parmuchi et al., 2002; Pope et al., 1997; Rosenqvist et al., 2007; Silva et al., 2008; Töyrä & Pietroniro, 2005; Töyrä et al., 2002; Wang et al., 1995). Some work presents evidence as to the merits, or otherwise, of the use of C-band horizontally co-polarised radar (C-HH, in which both the sent and received signals are filtered in the horizontal plane). Töyrä & Pietroniro (2005) note that SAR has difficulty separating inundated shrub or forest from other vegetation which is dry. They tackle this problem with the use of auxiliary visible/infrared (VIR) data. Pope et al. (1997) identify an apparent threshold in the height and density of grasses that determines the response (whether rising or falling) of backscatter to inundation, and suggest that the use of C-HH data to detect partial flooding in marshes with sparse cover “will probably not be possible”. This conclusion is shared by Pope et al. (1992), who note the lack of distinction between flooded emergent grasses and non-flooded grasses using C-HH. Another potential problem with C-HH data is that, due to the polarisation of incident radiation on the surface of water, co-polarised signals (HH and VV) are more susceptible to Bragg resonance than cross-polarised, although HH appears less so than VV (Liebe et al., 2009).

For both C- and L-band data, Wang et al. (1995) found that HH data displayed a greater flooded/non-flooded backscatter ratio than VV in their study of response to flooding in the Amazon forests, as was the case for Grings et al. (2009) when exploring the use of Envisat ASAR data to estimate water storage in the marshes of the Paraná Delta in Argentina.

In studies of aquatic or flooded vegetation, the presence of water is generally manifest in a higher backscatter signal (e.g. Bartsch et al. (2008); Costa & Telmer (2006)), but this is not always the case. While total submersion of vegetation, in the absence of Bragg scattering, results in a low backscatter response due to specular reflection away from the sensor, the behaviour of the signal response as the water rises to that point depends on several factors.

Grings et al. (2009) found that verticality of the vegetation determined the behaviour of backscatter from flooded marshes: they observed that the response went up when the inundated marsh was mainly of vertical orientation, and down when of random orientation, a finding shared by Silva et al. (2008), Töyrä et al. (2001) and Pope et al. (1997). The latter describe a similar trend reversal in terms
5. Vegetation effects

of vegetation density, observing an increasing backscatter value with flooding where the density of emergent Cattail and Saw Grass was greater than 60% cover, and a decrease in densities less than 50%. Flooding was also associated with a decrease in backscatter where small-stemmed rushes occurred at 50%–80% cover (Pope et al., 1997).

Costa & Telmer (2006) used RADARSAT and JERS-1 data to examine the lakes of the Brazilian Pantanal. They found that taller denser aquatic species give the highest C-band backscatter values due to dihedral scattering. However, Ramsey (1998) found that in taller marshes, C-band and X-band returns were dominated by volume scattering, with dihedral scattering being predominant in shorter marsh grasses.

Kasischke et al. (2003) studied ERS SAR data at 13 sites in the South Florida wetlands, and found that in all but one of the study sites, radar backscatter decreased with increasing water level. The one exception, i.e. where backscatter increased with increasing water level, was the site with the highest above-ground biomass.

It is clear that certain combinations of structural orientation / homogeneity, vegetation density and spatial coverage preclude the ability to distinguish flooded and non-flooded vegetation, as was found by Parmuchi et al. (2002) in the case of marshes in the Paraná Delta, despite the knowledge that there existed a large difference in water levels between data acquisitions. The vegetation species in this case was *Scirpus giganteus* with 100% coverage, which comprised vertical stems rising to 1.5–2m. This inability to distinguish between flooded and non-flooded states would seem to contradict the findings of Grings et al. (2009), mentioned above, who apparently observed a marked separability between ASAR readings in flooded and non-flooded states of the same species (identified in this case as *Cortaderia*), in the same region. Marsh density and biomass conditions during winter are suggested as possible reasons for the phenomenon by Parmuchi et al. (2002).

Considering the varied results, it seems that the flooding of grasslands may be observed either as a drop in backscatter due to specular reflection, or a rise due to dihedral scattering, both of which may conceivably be experienced within the bounds of a single pixel. Pope et al. (1997) refers to the trade off between
increasing backscatter from dihedral interactions and decreasing backscatter from forward scattering. The determination as to which of these is dominant, is governed by some vegetation cover threshold, which may vary between species and growth stage. It seems unavoidable that there must be a range of cover within which the presence of flooding is undetectable with radar data. In either case, the detectability of water depends on the amount of attenuation present through volume scattering above the water level. Töyrä & Pietroniro (2005) note this fact, finding difficulty separating inundated shrub or forest from non-flooded vegetation or bare soil without the help of auxiliary data in the form of optical imagery. Pope et al. (1997) acknowledges that there “appears to be an important height-density threshold in C-HH”. Töyrä et al. (2001) found that their ability to detect flooded sedges and grasses was confined to the summer when the newly-grown stalks were upright, and that in the spring, the dihedral backscatter became attenuated by brown vegetation and thatching.

The mechanisms which conspire to render the Staaten flood invisible using GM data can only be understood by measuring backscatter responses under known vegetation and flood conditions.

### 5.4 Method

GM data covering the two flood events was preprocessed according to methods described in Chapter 2. A list of data used may be found in Appendix B.

The ability to use GM data to map the Flinders flood was tested by classifying flooded and non-flooded pixels using the methods described in Chapter 4 O’Grady et al. (2011), which involve a region-growing algorithm and track-for-track baseline image deduction. This was done on three dates for which near-simultaneous cloud-free MODIS data were available, in order that the accuracy of the map could be tested. Kappa tests (Cohen, 1960; Foody, 2006; Hunt et al., 2010; Tolpekinn & Stein, 2009) were carried out to this end.

In order to analyse the radar response to varying water levels in the savanna of the western Cape York river basins, a series of capacitance-type water level loggers (Fig. 5.11) was set up across a transect running south from the Mitchell River, over a distance of 22km, approximately 2km apart (see Fig. 5.10). These
Figure 5.10: Location of water height loggers in the Mitchell floodplain
recorded data over the period of one year, including dry periods and periods of flooding at various depths.

**Figure 5.11:** Capacitance loggers set up in the Mitchell floodplain (as numbered). The main channel of the Mitchell near Shello Station is shown in its dry season state in the lower right-hand image

---

### 5.5 Results and Discussion

#### 5.5.1 Flinders

Figure 5.12 shows the backscatter response at arbitrary points in the Flinders floodplain from the middle of the 2008/9 wet season through a complete wet-
dry-wet season cycle. From this plot, it can be seen that the tussocky grasses dominating the floodplain produce high backscatter values, between -5 and -9 dB at the height of the wet season, and low values of -11 to -13 decibels during the dry. Total submersion of the grasses in February 2009 produced a backscatter value of around -17 dB, but the same sample point a few weeks later returned a value at the high end of the observed range. Where no flooding occurred, the range of values at any time throughout the year across the sample points was fairly consistent at -3 to -4 dB.

![Figure 5.12: Variation of average GM backscatter values through the year from the same orbit track, at various points in the Flinders floodplain](image)

The seasonal variation of backscatter values is consistent with findings by others (e.g. Bartsch et al. (2008); Kasischke et al. (2003); Wang et al. (1995)) as to the domination of backscatter signals by soil moisture. In fact, high soil-moisture response is all but indistinguishable from the response of semi-submerged grasses causing dihedral backscatter.

The build-up of water from dry soil through to increased wetness, to inundation, to reduction in water level below the top of vegetation and increased dihedral scattering may be seen in Figure 5.13.

On 31 January, the dry conditions across the floodplain are evident from the general low backscatter values below around -10 dB. By 15 February the flooding can clearly be seen, surrounded by much higher (-5 – 0 dB) backscatter corresponding to wet soil and dihedral scattering. In the next two images showing
19 February and 3 March, the area of total inundation represented by the homogeneous region of values at or below -16 dB reduces, and the radar-bright areas, tending towards zero loss, increase, as the flood water drops below the height of the vegetation, and interactions between the water surface and emergent vegetation result in dihedral scattering. Between 4 March and 9 March these areas are replaced by lower pixel values as the water level drops.

The effect of incidence angle on dihedral scattering is evident when comparing the images in Fig. 5.14, taken 11 hours apart on 9 and 10 March 2009. The former was taken was taken on an ascending orbit, where the data corresponding to the flooded region was acquired at a high incidence angle ($\approx 40^\circ$). The latter was acquired on a descending orbit, with the flooded region corresponding to lower incidence angles ($\approx 15^\circ$). The normalisation process with respect to incidence angle, using parameters derived from the methods described in Chapter 3, and the classification described in Chapter 4 and in O’Grady et al. (2011), mitigate this effect to a large extent.
5. Vegetation effects

Figure 5.14: Two GM images of the Flinders floodplain, acquired 11 hours apart on 9–10 March 2009. On the left, the region of interest (ROI) is on the outside of the swath of an ascending orbit, where incidence angles are around 40°. On the right, the ROI is on the inside of a descending orbit, with incidence angles around 15°.

Progression of flooding

In order to observe the various states of vegetation–water interaction through the Flinders flood, an unsupervised classification was performed on the time series of images that were acquired through the flood event. This resulted in the grouping together of those areas with a similar behaviour of backscatter response over time.

Figure 5.15 follows the progress of the mean backscatter returned by four value groups, from the resultant classification, through the flood event. Prior to the build-up of flood water in late January 2009, pixels in the red, blue and orange classes showed a mean backscatter values of -11 dB consistent with dry soil. The green class, dominated by the highly-weathered bedrock to the southwest of the region, and the dry sand of the coastal areas to the north, returned a lower mean value of -12 dB, consistent with absorption and specular reflection. The next period, up to 15 February, sees a peaking of the gauge height at Walkers Bend of around 15 m. The blue-red class values dive to -15 dB, where total submersion is reached. It is unfortunate that no GM data were available earlier in February,
5. Vegetation effects

Figure 5.15: (Below) Mean $\gamma$ values for classes determined using an unsupervised classification of five GM images straddling the Flinders flood event, shown against gauge height at Walker’s Bend. The geographical location of the classes is shown at the top.
when we would expect to see, for these two classes, the backscatter values rise due to dihedral scattering, prior to total submersion. However, we do see this occur as the waters recede. In early March we see the blue class values approach those of the orange class at around -6 dB, representing those areas where the flood has drained, exposing wet soil. The red class values peak at an average of -2 dB, and represent those areas where the grasses remain emergent through the remnant flood waters. The orange class shows those areas which did not flood, and the progression from dry soils to wet soils can be seen between late January and mid February, where they remain fairly constant at around -6 dB until the gauge height at Walkers Bend is seen to drop off in early March. At this time, the average backscatter values for all of the classes begin their descent to their normal values for the time of year (driven by soil moisture), averaging around -9 dB.

Class map of Flinders flood

The full extend of the Flinders flood was mapped using the preprocessing, image differencing, thresholding and classification techniques described in Section 4.4. The results of the subsequent classification of flooded and non-flooded pixels are shown in Figure 5.16. Correctly classified water and land are shown in blue and yellow respectively, and pixels incorrectly classified as water and land are shown in cyan and red. The $\kappa$-statistic resulting from the error analysis on the 15-February classification is 0.70 (see Table 5.1). No hard and fast rule exists to determine a definitive interpretation of the $\kappa$ value, though some attempts have been offered (e.g. Fleiss (1981); Landis & Koch (1977)). In all assessments, a $\kappa$-statistic value 0.7 ranks highly, indicating a good classification result. It is significant that errors of omission, where water is classified incorrectly as land, are concentrated to the greater part on the north-east area, which actually represents the lower reaches of the Norman River. This error is demonstrated best in the classification for 3 March, seen in the lower left image in Figure 5.16.
Figure 5.16: Accuracy of flood classification using GM data against a MODIS Band 6 threshold classification, at various stages of the flood.
Figure 5.17: Water level measured at five logger locations (in blue) in the Mitchell floodplain over a nine month period, together with ground-range projected backscatter ($\gamma_30$, in red) derived from GM data. The bottom right-hand plot shows daily rainfall at Kowanyama Airport for the same period (Courtesy Australian Bureau of Meteorology http://www.bom.gov.au) Logger locations were shown in Fig. 5.10
5. Vegetation effects

<table>
<thead>
<tr>
<th>Class</th>
<th>MODIS B6</th>
<th>Row Sum</th>
</tr>
</thead>
<tbody>
<tr>
<td>G</td>
<td>Land</td>
<td>102136</td>
</tr>
<tr>
<td></td>
<td>Water</td>
<td>10758</td>
</tr>
<tr>
<td></td>
<td></td>
<td>112894</td>
</tr>
<tr>
<td>M</td>
<td>Land</td>
<td>10941</td>
</tr>
<tr>
<td></td>
<td>Water</td>
<td>42209</td>
</tr>
<tr>
<td></td>
<td></td>
<td>53150</td>
</tr>
<tr>
<td>Col Sum</td>
<td>113077</td>
<td>52967</td>
</tr>
<tr>
<td></td>
<td></td>
<td>166044</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Class</th>
<th>% Commission</th>
<th>% Ommission</th>
<th>Estimated κ</th>
</tr>
</thead>
<tbody>
<tr>
<td>Land</td>
<td>9.529293</td>
<td>9.675708</td>
<td>0.701270</td>
</tr>
<tr>
<td>Water</td>
<td>20.585136</td>
<td>20.310760</td>
<td>0.697725</td>
</tr>
</tbody>
</table>

\[ \kappa = 0.699493, \quad \kappa \text{ Variance} = 0.000003 \]

Table 5.1: Kappa (κ) results from the accuracy test performed on the classification of flooding in the Flinders basin on 15 February 2009

5.5.2 Staaten

Figure 5.17 shows the backscatter values being returned under certain flood levels in very similar environmental conditions to the Staaten flood, in the neighbouring Mitchell floodplain. Five data loggers were set up and left for one year, recording water height above ground level hourly. The results are plotted in Figure 5.17. Although the loggers were a few kilometres apart, the region is extremely flat, at around 11m above sea level. The three loggers furthest away from the Mitchell channel (33901, 34762, 34828) recorded two peaks, in early and late February, taking the water height up to between 200 and 1100 mm. It is only the first peak in early February, however, common to all of the logger data, which causes a discernible drop in the backscatter (γ) value which we would expect to see due to specular reflection on open water. In the case of loggers 33901 and 34827, it is certainly true that a water height of at least 300mm would be required to submerge the grasses, and this height was only achieved during this peak.

We can determine from the plots that values returned during the flooded period are of a broad range whose mean tends to follow the range of values expected as a result of soil moisture alone, at this time of year. We have seen the seasonal trend of backscatter values in the Flinders region in Figure 5.12, in which the low backscatter values during the dry season reflect the extremely low
soil moisture content during that period.

![Figure 5.18: Tallest stratum growth-form in flooded areas](image)

In the tussock grass-dominated Flinders floodplain, we see, in late January 2009, when the soil was dry, a backscatter value across the floodplain of around -11 dB. From this point, we see the flooded region drop to -15 dB, and then rise sharply as the flood recedes. This is consistent with the findings by others that short, more vertically oriented grasses are more inclined to return a response resulting from dihedral scattering. Another difference between the two regions is the presence of trees in the Mitchell and Staaten floodplains, in contrast to the Flinders flood region, the tallest stratum growth form of which is grass (see Figure 5.18). We know that C-Band radar will interact with structures of the order of scale of the small branches and foliage within the canopy of the eucalypts which dominate the Mitchell and Staaten, which can only serve to dilute the signature resulting from the grass and water interactions, with volume scatter.
5. Vegetation effects

5.6 Conclusion

From the classification and accuracy test performed using the Flinders data, GM data was shown to be a useful tool in the mapping of floods, given favourable surface conditions. It was mentioned in section 5.5 that the errors of omission in this classification were concentrated in the Norman River section of the study region. This is the section of the greater Flinders floodplain that shares a similar vegetation growth form class to the Staaten region, as can be seen in Figure 5.18.

Findings by others regarding the influence of vertical orientation and above-surface density in grasses on C-HH data is mixed. We found that in the short, vertically-oriented bluegrass of the Flinders floodplain, during partial inundation, a lower water level produced a higher backscatter value as dihedral scattering increased. In the denser, thatched grasses of the Mitchell and Staaten floodplains, the presence of water was barely distinguishable. Our findings straddle the demarcation of the expected radar response in flooded grasslands and wooded savanna, depending on the nature of vegetation. To model precisely the synergy between vegetation structure, density, orientation and phenology may not be possible, given the diversity of species and distribution. It is clear that certain environmental conditions preclude the use of GM data for the monitoring of floods. When radar is used to detect the presence of flooding, we rely on a low return in the case of total inundation, due to specular reflection, and on a high return where emergent vegetation is present, due to dihedral scattering.

The benefits of the use C-HH radar data to monitor shallow but extensive flooding may therefore come with the cost of a similar study on a floodplain-by-floodplain basis as needed, to calibrate thresholds, or to discount the use of C-band radar entirely.
Chapter 6

Conclusion

6.0.1 Research questions

The research questions posed in Section 1.6.1 were addressed as follows:

How does radar backscatter vary with incidence angle for different surface conditions? How does this affect the segmentation of open water? The variation of backscatter with incidence angle showed a close fit ($R^2 = 0.8$) to a sinusoidal model over open water, but a far less close fit ($R^2 = 0.2$) over land, due, most likely, to environmental variations of the target areas throughout the time series over which the regression was carried out. Despite this, the slope of a linear approximation to the model showed clear regional clustering. This clustering, when compared with various land cover classifiers, was found to correlate best with regolith.

The slope of the linear approximation to the variation of backscatter with incidence angle proved less sensitive to soil moisture and to Bragg resonance than the absolute backscatter values themselves, suggesting a potential avenue of research into more accurate mapping of water, in the event of future C-Band missions providing sufficiently frequent coverage.

The calculation of slope and offset of the linear model for particular regions of interest provided a reasonable means to normalise individual GM images with respect to incidence angle through the rest of the research done for this thesis.
6. Conclusion

Under what vegetation conditions (vegetation type, size, density, orientation) does multihedral backscatter distort the radar signal so as to make flood water indistinguishable from its surroundings? The short, vertically-oriented bluegrass of the Flinders floodplain demonstrated a predictable backscatter signature throughout the process of soil saturation, partial flooding and total inundation, allowing a flood to be mapped to a high level of accuracy ($\kappa$-statistic = 0.7).

The inundation of the taller, denser thatched grasses of the Mitchell/Staaten floodplains returned backscatter signals that were indistinguishable from the surrounding wet soil and vegetation through periods of partial inundation, with the only predictable signal being the fall in backscatter values at total inundation.

Prior knowledge of vegetation characteristics of the Flinders region would further increase the accuracy of the flood classification, by segregating the image into regions of greater and lesser certainty.

**How can we separate dry soil/sand from floods through arid regions?**

**How significant is this problem?** Track-for-track image differencing techniques can be used to eliminate ambiguity between low backscatter response due to specular reflection on flood waters and low backscatter response due to absorption in dry sand.

The problem was found to be potentially significant, with, in one example, the area of absorption surrounding flood waters in one GM image having the same order of scale as that of the specular reflection on the flood itself.

The work done to resolve this problem also demonstrated the ability of GM data to fill-in the flood mapping function during critical early stages of a major flood, were VIR data were unavailable due to cloud cover, and also demonstrated the stability of the backscatter threshold for flood demarcation, in comparison to that for SWIR reflectance from MODIS images.

The combination of techniques and data sources allowed a large flood to be mapped in its entirety over the course of several months.

**How can the processing of such a large dataset be managed?** How can we automate the download, registration and orthorectification of
6. Conclusion

a high volume of GM data files to allow efficient analysis? The open source geospatial community, led by the Open Source Geospatial Foundation (OSGeo\(^\text{1}\)) play a vital role in the provision of fundamental tools to read, write and process most file formats used by the GIS field, and these are relied upon by both commercial and non-commercial software providers.

By developing the scripts used in this thesis, I have demonstrated that open source tools can be used throughout the whole data download, registration, pre-processing and analysis process, allowing control, transparency, flexibility and efficiency, at very little cost.

The primary reason for the decision to take the time to carry out this stage of the work, rather than relying on bespoke software such as ESA’s NEST tools, was the fact that the whole process could be carried out using the university’s high-performance computing system, in a parallel processing environment. It is easy to underestimate the resultant advantage of being able to run and re-run scripts many times, tweaking parameters in each case, making and correcting mistakes without having to wait a considerable time to find out that a mistake may have been made in the first place. A fast and efficient processing system allows for an environment of exploration, in which trial-and-error can lead to new paths in the research, as it has done for many people in all fields of research.

6.1 In summary

I have investigated the potential role that Envisat ASAR Global Monitoring Mode data plays in the mapping of broad scale flooding. Now that Envisat is no longer operational, although we have alternative C-band satellite sensor data available to us, we await a contender for GM data which matches its coverage and ready availability. We can be sure that, as technology in processing circuitry, data storage capacity and, in particular, battery power storage and solar recharge capabilities advance, we may not have to rely on such coarse spatial resolution as that of GM data. Any increase in resolution will pose further challenges to the processing involved in the type of research that makes use of large time series,

\(^{1}\text{http://www.osgeo.org/}\)
such as that with which I have dealt here. Depending on the scales being looked at, the type of pre-processing described in Chapter 2 may involve generalisation computations where they were absent here, in order that resultant data sets were spatially comparable to other data with which they were to be compared. We can see this process happening even with GM data at its 1km resolution, where for example, soil moisture products have been derived and compared to similar products that have 25km or even 50km spatial resolutions. The foreseen finer spatial resolutions to be dealt with in the next generation of research will present plenty of new opportunity to gain a better understanding of the complexities of backscatter response. To take advantage of these resolutions, orthorectification must be carried out in compliance with comprehensive and proven techniques as set out by, for example van Zyl et al. (1993) and Ulander (1996), using a similarly fine-scaled DEM. More careful consideration will need to be given to georeferencing, though it is expected that much of this preprocessing will be offered by the data providers, as was the case, for example, with L-band ALOS PALSAR. Given such anticipated increased requirements in terms of processing, it is felt that the principles described in this thesis, in the use of universally available open source data interoperability tools, open platform scripting and GIS software and multi-threading techniques that can easily be ported to super-computers such as JCU’s HPC, will continue to offer the best solutions.

Whilst the use, on the same broad scale, of data at a much finer resolution, will lead to less ambiguity caused by the aggregation of sub-pixel backscatter responses from a surface of varied structural, textural and dielectric properties, many factors which contribute to the complexities faced with the use of GM data will remain. The availability of finer resolution data at a similar temporal frequency as GM will open the path for the extension of work described here. More precise matching of radar response with topographical parameters will lead to a better understanding of the variance of backscatter with incidence angle. Flooding extents will be able to be mapped with greater precision. Reduced pixel-mixing will lead to better segregation of soil moisture, for example, from dihedral scatter at the boundaries of floods. Partially flooded vegetation may be more precisely matched with radar response categorically in terms of vegetation species, which, we have learnt, is vital in order to interpret backscatter from such
environments to derive the presence of water, or even its relative height.

At the outset of the work described here, the decision was made to concentrate on GM as a primary source of data with which to investigate the mapping of floods. This decision was made for a number of reasons. The time constraints governing most Ph.D. research (which are really the manifestation of financial constraints), impose the necessary trade-off between depth of research and scope. The temporal frequency of GM data provided a huge time series with which to investigate the behaviour of backscatter against various topographical parameters, as well as giving me the opportunity to test the availability of data in flood events against VIR alternatives. Also, from research by others it was found that, of all possible wave bands and polarisation configurations, the C-HH configuration offered by GM data might well be the best one for the job. It is worth noting here that one of the most confounding problems found in all research in the detection of water using both C- and L-band data was the phenomenon of Bragg Resonance. As I have discussed, this is dependent on the orientation, as well as relative scale, of regular waves on the surface of the water. As such the effect is far reduced using a cross-polarised, C-HV or C-VH configuration, although this may come at the cost of a lower radiometric resolution on surfaces where no polarity shift has occurred. Sentinel-1 will provide both cross-polarised configurations (ESA, 2012b).

From my initial review of literature on the subject of water detection using radar remote sensing, fundamental questions emerged: What were the effects of incidence angle on backscatter? How could specular reflection be distinguished from absorption in flooded arid regions? How did backscatter response vary with vegetation conditions under partial inundation? To answer these questions, I devised a means to process and manage the large dataset, necessary for such analysis at a pixel-by-pixel level. I used regression to gain values, for each pixel and under seasonal conditions, of slope and intercept in a linear approximation of the relationship between incidence angle and backscatter, by which to normalise the data. I managed to further mitigate incidence angle effects through image differencing, and with this method I also separated flood from desert, successfully tackling a problem that had not been dealt with before. In this way I captured the progress of an entire flood event, stretching across a thousand kilometres over
a hundred days. I investigated radar response to total and partial submersion of vegetation in two regions in the tropical savanna of northern Queensland, and gained an understanding of the interaction of radar with different vegetative conditions, adding to a field of research that was, and is, in great need of further investigation. The study reinforced my conviction in the ability of GM data to map a large rapid flood event covering nine thousand kilometres with open water, while at the same time highlighting the failure of GM to capture a similarly sized event in a neighbouring region, with a population close to zero; where floods, even on this scale, can conceivably pass by unnoticed.

In attempting to answer these questions, other questions inevitably arose, and some were answered. I observed the correlation between the variation of backscatter with incidence angle and regolith. I found that the unique and highly separable value of $\Delta \gamma / \Delta \alpha$ on open water appeared to be relatively independent of Bragg Resonance, paving the way to the possibility of taking advantage of this fact to map flooding using C-HH data under unfavourable wind conditions, with impunity. Further, it was found that this opportunity was unlikely to be met using GM data, due to margins of error in consecutive images where the difference in incidence angle was insufficient. The question of how to optimise radiometric thresholding was answered with a unique incremental threshold convergence method using complementary MODIS data, in a process which highlighted both the stability of the resultant GM-derived threshold, and the instability of the MODIS SWIR reflectance threshold for open water.

The understanding gained from this research, and the questions raised, coincide well with the opportunity offered by JCU for further research, to use radar and optical/thermal satellite data to contribute to knowledge being gathered by a team under the Hydrology Department, as well as to investigate further the interaction of radar with vegetation and water in a tropical floodplain. I look forward to this work, and to the discoveries that await the field of remote sensing as new satellite missions come online.
6. Conclusion

6.2 Beyond the thesis

6.2.1 C-band synthetic aperture radar

On 8 April 2012, while the conclusions to this thesis were being written, the European Space Agency lost contact with Envisat. After spending a month trying to restore contact to the satellite, ESA declared the mission at an end on 9 May. As with many successful earth observation satellite missions, Envisat was in operation for well over its planned lifetime, having been in orbit since 2002.

With the corresponding loss of the ASAR sensor, there is a gap in availability of data data similar to ASAR’s GM mode, with its temporal frequency and availability. In the short term, this will have an immediate effect on research such as this, which seeks to take advantage of this particular data mode that has been peculiar to Envisat ASAR, and at least one systematically generated product derived from GM data, that being ESA’s Tiger Innovator project Soil moisture for Hydrometeorological Applications over SADC (SHARE) (Bartsch et al., 2008).

Thankfully, the end of Envisat does not mark the end of satellite C-band radar availability. The Canadian Space Agency’s (CSA) remarkable Radarsat-1 has been earmarked to fill part of the gap left by Envisat’s demise (Boucher, 2012). Radarsat-1, launched on a 5 year mission in November 1995, is still going strong in May 2012. Though having no GM mode as such, Radarsat-1 does have a wide ScanSAR mode with a nominal resolution of 100m, as well as fine-beam modes operating up to an 8m nominal resolution. The later Radarsat-2 satellite, launched in December 2007 and being capable of dual-sided imaging to a nominal resolution of 3m, also adds to the available list of C-band data. Further to this, CSA are planning the continuation and improvement of Radarsat-1 and -2 in the planned launches in 2016 and 2017 of the RADARSAT Constellation, which is intended to provide daily C-band SAR coverage of 95% of the world (CSA, 2012).

ESA has been planning to launch its next C-band SAR sensor aboard the Sentinel-1 mission in 2013 (ESA, 2012b), and the cessation of operation of Envisat gives this mission renewed significance. Sentinel-1 will be one of a string of satellites launched under a programme entitled Global Monitoring for Environ-
6. Conclusion

\textbf{GMES}, headed by the European Commission in partnership with the ESA and the European Environment Agency (ESA, 2012a). The subsequent Sentinels 2-5 will include high resolution VIR sensors for the monitoring of land, ocean and atmosphere.

The Indian Space Research Organisation (ISRO) launched its \textit{Radar Imaging Satellite-1} (RISAT-1) on 26 April 2012 (ISRO, 2012), though the future availability of data from the C-band SAR sensor on board to the global scientific community is unclear.

\subsection{Mary River, Northern Territory}

\textbf{The need to monitor greenhouse gas emissions from floodplains in Australia}

One vital component in gauging the extent of the greenhouse-gas-induced climate change problem is the monitoring of emissions. The United Nations Convention on Climate Change (UNCCC) identified the systematic observation of surface greenhouse gases as an essential component of climate change policy (Onoda, 2008). Currently there are large uncertainties in surface fluxes of CO$_2$, CH$_4$ and N$_2$O, but an increasing feedback between climate change and greenhouse gasses is expected (Bréon & Ciais, 2010). The Global Warming Potential (GWP) of CH$_4$ is 25 times that of CO$_2$ (Dalal \textit{et al.}, 2008), which amounts to 20\% of the total global warming effect. To put the GWP of CH$_4$ in context, Dalal \textit{et al.} (2008) tell us that, since 1750, CO$_2$ concentrations have gone up by 33\%, while CH$_4$ concentrations have gone up by 75\%.

Somewhere between a quarter and a third of global CH$_4$ emissions come from wetlands and lake sediments (Chen \textit{et al.}, 2011; Dalal \textit{et al.}, 2008). The feedback between climate change and methanogenesis in such environments in Australia is poorly understood, due to a lack of research data, but we do know that there is a positive feedback between atmospheric CO$_2$ concentrations and CH$_4$ emissions (Dalal \textit{et al.}, 2008). Law & Garnett (2011) assess the use of the National Carbon Accounting Toolbox (NCAT) for estimating and mapping carbons stocks in Australia’s Northern Territory, and conclude that further work is needed on soils,
fire, grasslands, wetlands and woody debris in order to improve the validity of NCAT for carbon estimates.

The role of C-band radar

Soil and vegetation characteristics are major factors governing the production and emission of CH$_4$ in wetlands. In addition, the hydrodynamics of wetland systems are crucially important to understand, for a number of reasons. Flooded wetlands can produce forty times the amount of CH$_4$ as wet soils (Arnell, 2002). In dry ecosystems, methane fluxes are linked largely to soil porosity, whereas, once flooded, methane emissions are a function of vegetation characteristics and the actual level of the water (Chen et al., 2011; Dalal et al., 2008). Also, the very dynamic density and composition of aquatic vegetation in a floodplain are governed by its hydrology and morphology (Coops et al., 1999).

Vegetation types determine relative CH$_4$ emissions. There is a positive correlation between the distribution of vegetation with well-developed aerenchyma and CH$_4$ production and emission. Such emissions may also be influenced by the provision of carbon substrates through the roots. In addition, vegetation can be an indicator of soil types, which in turn effect CH$_4$ production through relative levels of trace elements, electron acceptors, pH and salinity (Dalal et al., 2008).

Research is currently underway by a team here at JCU, monitoring greenhouse gas emissions in a section of the Mary River floodplain in the Northern Territory. The intention is to upscale results to the broader floodplain, and beyond, based on common indicators that can be correlated to the results. In light of what has been discussed here, data must be collected in terms of vegetation type, density and phenology, as well as the hydrodynamics of the flood plain - water height, flood duration and water body connectivity. Cost and practical accessibility limit the provision of large scale data of this sort to remote sensing as a method. Research described in this thesis has outlined the limitations of VIR data alone as a continuous monitoring tool, due to cloud cover. We have seen that microwave radar data are largely free from this problem, and that C-band data, in particular, have been the radar data of choice for many studies of aquatic vegetation. We have also seen that C-band radar allows us to determine the presence of water
and, in certain vegetation environments, gives us an indication of the level of submersion of aquatic vegetation.

We have also learnt the important limitations of C-band data. The disadvantage of coarseness of scale suffered by GM data is somewhat irrelevant, given that there remains no similar contender with a 500m pixel spacing. However, as mentioned, Radarsat-1 and -2 do provide an invaluable stepping-stone until the advent of ESA’s Sentinel mission, at which time we can hope for a 1–3 day repeat coverage once again (ESA, 2012b). Other limitations, such as we found in the ambiguity of signal in the Staaten grasslands under flood, are peculiar to a particular vegetation regime. In this regard, the particular conditions of the Mary River ecosystem can be evaluated on their own merits. In addition to radar, VIR remote sensing will necessarily be used to distinguish, as needed, the vegetation types and stages. Such requirements can be met by a number of available satellite sensors, depending on the resolution required, ranging from MODIS and Landsat TM at 500, 250 or 30m resolutions, all the way to the 0.3m panchromatic, 2m VIR and the almost daily revisit time offered by DigitalGlobe’s WorldView-2 (DigitalGlobe, 2012).

Participation in the research at Mary River constitutes not only a contribution as a tool to upscale the observed data, but also provides the opportunity to answer questions arising from the work described in Chapter 5 and by others, surrounding the complex relationship of C-band backscatter, with water level and structure, density, orientation and height of emergent vegetation.

### 6.2.3 Combination with other data types

The satellite Gravity Recovery and Climate Experiment (GRACE) is a twin satellite mission which provides data in the form of a set of Stokes coefficients to a truncated spherical harmonic expansion of the geoid. The data are used to compute gravity anomalies, which are deviations from the large-scale gravity field, caused by density variations in the subsurface. As ground and surface water constitute a substantial fluid mass with significant variation over time, GRACE’s potential contribution to the field of hydrology is clear. A significant amount of literature directly related to GRACE data are devoted to the establishment of the
accuracy of the reduction process used to derive the anomaly figures, as this is a complex procedure. The further derivation of hydrological parameters, and the apportioning of values to various components of the hydrological cycle are even more complex. GM data allows us the opportunity to test the relationship of observed rainfall extent and relative soil moisture with GRACE anomaly derivations for the Australian region, forming a component of the ground-truthing work necessary to increase the accuracy of GRACE derived products. I am currently involved in discussions towards the design of a research model to take advantage of this opportunity.

Back in Section 4.6.3 I commented on the potential of the future Surface Water and Ocean Topography (SWOT) mission, to be launched in 2020. Using wide-swath altimetry technology, SWOT will provide temporal and spatial variations in water volumes stored in rivers, lakes, and wetlands at unprecedented resolution (Biancamaria et al., 2010). SWOT will generate a global 3D mapping of all terrestrial water bodies whose surface area exceeds 250 m$^2$ and rivers whose width exceeds 100 m (Biancamaria et al., 2010). The principal instrument of SWOT will be a K_a-band Radar Interferometer (KaRIN), which will provide heights and co-registered all-weather imagery of water over 2 swaths, each 60 km wide, with an expected precision of 1 cm/km for water slopes, and absolute height level precision of 10 cm/km$^2$. The high spatial and radiometric resolution of SWOT data, with repeat coverage occurring at least twice every 21 days, will enable cross-validation of C-band derived flood extents, facilitate volume calculations and add precision to the derivation of GRACE products. The contribution of SWOT to flood mapping will be enormous.

### 6.2.4 Characteristics of fire occurrence and spread under a changing Australian environment

Tropical savannas (TS) cover more than one third of the continent of Australia (Landsberg et al., 2011). Globally, they contain almost 15% of the world’s carbon stock, and contribute 38%, 19% and 59% of emissions of CO$_2$, CH$_4$ and NO$_2$ respectively (Spessa et al., 2005). This makes TS a significant component of
6. Conclusion

the global greenhouse gas budget. The TS landscapes of northern Australia are
dynamic, continually changing due to anthropogenic and natural effects. One
important feature straddling both of these categories is fire. In their study of
fire patterns in two regions within the TS in northern Australia, Felderhof &
Gillieson (2006) found that up to 74% of the Cape York study region burnt
in a single year, and determined that an understanding of fire patterns in the
TS was a necessary step towards understanding the vegetation dynamics in the
region.

Fire patterns are determined by ignition and propagation characteristics (Chu-
vieco et al., 2004). Propagation depends on fuel and environmental conditions.
Ignition depends on man or nature, and the latter takes the form of lightning
(Felderhof & Gillieson, 2006). Neilson (1995) models fire likelihood on lightning
conditions and intensity of rainfall in summer, as do M.B. & Pivello (2000), who
note that most fires occur in the transition from the dry to the wet season. In
the context of greenhouse gas emissions, it is expected that lightning occurrences
will increase, with an increased concentration of CO$_2$ in the atmosphere (Cardoso
et al., 2008).

Many researchers have used remote sensing in an attempt to gauge fire risk.
Chuvieco et al. (2004) cite three disadvantages in their endeavours to this end,
being insufficient temporal frequency, obscuration by cloud cover and calibration
difficulties. Certainly the GM data used in this thesis would overcome the first
two of these - to the third problem it may offer its own brand of challenges. The
authors also highlight the disadvantage with their VIR data in the case of dead
fuel, in that wetness is not followed by greenness, and therefore the moisture
content is less discernible. Spessa et al. (2005), in their investigation into the re-
lationship of fire frequency, rainfall and vegetation patterns in northern Australia,
used a Normalised Difference Vegetation Index (NDVI) to distinguish vegetation
patterns, but note potential errors due to differences in “background soil colour”. The
importance of dead fuel moisture to the spread of fire in grasslands is em-
phasised by Sullivan (2010), who also reiterate the difficulty in measuring this
parameter. They also list the grass curing rate as a factor influencing fire spread,
and note that whilst rainfall does not slow down the curing rate of annual grasses
(which continue to cure once they have started), it does slow down the curing
rate in perennial grasses through regeneration.

The scale of TS regions under study confine any broad-reaching real-time fire occurrence or likelihood mapping to remote sensing. Apart from the use of thermal infrared data to detect fire occurrences directly, VIR reflectance data is limited in its response purely to surface reflection. Thermal emissions have been used to some effect by, for example, Vidal et al. (1994), who relate fire occurrences to the ratio between actual and potential evapotranspiration (AET/PET) derived from daily NOAA-AVHRR surface temperatures and synoptic air temperature (Leblon, 2005).

Little work has been done to explore the use of radar to add a measure of biomass density to fire likelihood detection. Leblon (2005) explore the use of Radarsat-1 to help determine the fuel moisture component of fire probability, but find that the 35 day revisit time is prohibitive, and see the contribution to the signal by soil moisture as noise. As we have seen in Chapter 5, it is important to know the characteristics of the vegetation species that cover the region of interest, as, depending on the density, structure and orientation of grasses, C-band backscatter may represent the wetness/dryness of the grass almost independently of the surface beneath. As discussed, the context of vegetation species is also crucial where rainfall may or may not effect fuel curing.

Bartsch et al. (2008) showed how GM data could provide an effective and, at 1km resolution, spatially fine (compared to alternatives) soil moisture product. We have shown that in some grasslands, C-band backscatter is representative of the wetness of grasses, rather than the surface below. The temporal frequency of GM data allows us to see, during the dry season, where in the savanna rain has recently fallen.

Figure 6.1 shows four GM images over the Flinders region in Queensland during the dry season. The top left image shows the low backscatter values where no rain has fallen recently, whereas the other images show where rain has fallen in the previous 24 hours or so. It may be, that analysis of the time taken for the radar response to drop, following a rain event, could give us an indication as to whether the backscatter is representative of the soil moisture, or the wetness of the vegetation. Where soil moisture is represented, the use of X-band data could be explored, the shorter wavelength of which ensures interaction of the radar sig-
nal with the smaller structured grass components. At the very least, with GM data we can conceivably generate maps showing rainless periods throughout the dry season. Work by Felderhof & Gillieson (2006) shows us how to effectively map fires in Australia’s tropical savannas over long periods using remote sensing. With such data sets, the path has been laid for us to analyse any correlation between fire distribution and spread, with products derived from radar data, together with VIR and thermal imagery. Although Envisat is no longer operational, its wealth of data remains for much potential future research, especially in the analysis of the relationship of backscatter with environmental conditions. In light of the importance of tropical savannas globally, and of fire patterns within those savannas, it is surely an avenue of research worth exploring.

Figure 6.1: Four images of the Flinders region in Queensland, in which the high backscatter response following rainfall is manifest as bright patches against the otherwise low response due to dry surface conditions.
Appendix A - Scripts

GrassAscii.pm

A perl module building a class with the ability to read and write GRASS GIS native ASCII raster images.

Listing 1: GrassAscii.pm

package GrassAscii;
use strict;
use warnings;

sub new {
    my $class = shift;
    my $fullpath = shift or die "You must argue a full path\n";
    my $self = {
        FULLPATH  => $fullpath,
        NORTH    => undef,
        SOUTH    => undef,
        EAST     => undef,
        WEST     => undef,
        ROWS     => undef,
        COLS     => undef,
        NULL     => undef,
        TYPE     => 'float',
        MULTIPLIER => 1,
        DECIMAL_PLACES => 2,
        DATA    => []
    };
    bless($self, $class);
    return $self;
}

sub write_ascii {
    my $self = shift;
    if (@_) {
        ($self->{TYPE}, $self->{DECIMAL_PLACES}) = @_;
    }
    my $dps = $self->{DECIMAL_PLACES};
    open my $output, ">$self->{FULLPATH}" or die "Could not edit/create ".
    $self->{FULLPATH}".\n";
ORTHO.pm

A perl module building a child class of GrassAscii, which reads raster value data generated by the GM module and performs orthorectification.

Listing 2: ORTHO.pm

```
package ORTHO;
use Math::Trig;
```
use GrassAscii;

@ISA = "GrassAscii";

use strict;
use warnings;

my $P = 500; # pixel size

sub my
{
  my $class = shift;
  my ($root,$R) = 0 or die "You need to argue a root and a satellite radius
"
  my $self = $class->GrassAscii::new($root.'.ORTHO');
  $self->[ROOT] =$root;
  $self->[R] =$R;
  $self->[HEIGHT] = undef;
  $self->[WIDTH] = undef;
  $self->[MARGIN] = 2; # How much to trim the edges by
  bless($self,$class);
  return $self;
}

sub fix_header
{
  my $self = shift;
  $self->[NORTH] = sprintf("%.0f",$self->[HEIGHT] - $self->[MARGIN]);
  $self->[SOUTH] = $self->[MARGIN];
  $self->[EAST] = sprintf("%.0f",$self->[WIDTH] - $self->[MARGIN]);
  $self->[WEST] = $self->[MARGIN];
  $self->[ROWS] = sprintf("%.0f",$self->[HEIGHT]);
  $self->[COLS] = sprintf("%.0f",$self->[WIDTH]);
  $self->[NULL] = 0;
  $self->[TYPE] = "int";
  $self->[DECIMAL_PLACES] = 0;
}

sub get_r
{
  # radius at geographic latitude
  my $flat = shift;
  my $flat = 1 / 298.257223560; # flattening
  my $a = 6378137; # Equatorial radius of WGS84 ellipsoid
  my $theta = atan((1 - $flat)**2 * tan(deg2rad($flat)));
  my $radius = $a * (1 - ($flat * (sin($theta)**2)));
  return $radius;
}

# Note check_dimensions() must initialize $WIDTH and $HEIGHT
sub check_dimensions
{
  my $self = shift;
  my $colsrows = (cols => undef,rows => undef);
  foreach my $thing ("cols","rows")
    {
    $colsrows{$$thing} = get_dim($$thing,$self->{ROOT}).".DATA" or die "Data appears to have no dimensions\n"
    }
  foreach my $field ("DEM","LAT","SRT")
    {
    my $dim = get_dim($$thing,$self->{ROOT}).".$$field";
    unless ($dim == $colsrows{$$thing}) { return 0 }
    }
}

$self->[WIDTH] = $colsrows{cols};
$self->[HEIGHT] = $colsrows{rows};
return 1;
sub get_dim {
    my ($thing, $fullpath) = @_;  # First pass to establish contributions (see notes 2 April 2011)
    my $response = `grep $thing $fullpath`;  
    chomp $response;
    if ($response =~ /:\t\s*(\d+)/) {
        $output = 1 + $1;
    }  
    return $output;
}
sub orthorectify {
    my $self = shift;
    # First to get past the headers
    open DEM, "$self->[ROOT].".DEM" or die "Couldn't open ".$self->[ROOT].".DEM";
    open DATA, "$self->[ROOT].".DATA" or die "Couldn't open ".$self->[ROOT].".DATA";
    open LAT, "$self->[ROOT].".LAT" or die "Couldn't open ".$self->[ROOT].".LAT";
    open SRT, "$self->[ROOT].".SRT" or die "Couldn't open ".$self->[ROOT].".SRT";
    open THETA, "$self->[ROOT].".THETA" or die "Couldn't open ".$self->[ROOT].".THETA";
    my %type = (DEM => [], DATA => [], LAT => [], SRT => [], THETA => []);
    my %handles = (DEM => $DEM, DATA => $DATA, LAT => $LAT, SRT => $SRT, THETA => $THETA);
    foreach my $key (keys %handles) {
        my $line = 'x';
        while ($line =~ /[^a-z]/) {
            local $FH = $handles{$key};
            $line = <$FH>;
        }
        my $row = split(/	|\s/, $line);
        $type{$key} = $row;
    }
    foreach (my $row = 0; $row <$self->[HEIGHT]; $row++){
        my $phi_max = 0;
        my @newrasterrow = ();
        my @contributions = ();
        my @displacements = ();
        my @shadows = ();
        for (my $i = 0; $i <$self->[WIDTH]; $i++) {
            push @newrasterrow, 0;
            push @contributions, 0;
            push @displacements, 0;
            push @shadows, 0;
        }
        # First pass to establish contributions (see notes 2 April 2011)
        for (my $col = 0; $col <$self->[WIDTH]; $col++) {
            my $h = $type{DEM}->[col];
unless ($h \text{ eq } 's' ) \{ \ # because these represent null values in DEM
    my $radians = $type{THETA} - [ ] [ col ] * 3.1415927 / 180;
    my $s = 299792458 * $type{SRT} - [ ] [ col ] * 10**s( -9 ) / 2; \ # Slant \rightarrow 
    range distance in metres to Ellipsoid! Corrected!
    my $phi = \text{asin}( $s * \text{sin}( $radians ) / \text{sqrt}( $s**2 + $h**2 - (2 * $s + $h) * \text{cos}( $radians ) ) ) ; \ # to calc shadow see notes 7
    April 2011
    if ($phi_max > $phi) \{ \ # This must be shadow - update phi_max, \rightarrow 
        shadow and leave
        $shadows$[ col ] = 1;
    } else \{ \ # carry on with contributions
        $phi_max = $phi;
        my $r = \text{get_r}( $type{LAT} - [ ] [ col ] ) ;
        my $big_omega = \text{acos}( ( ( $r + $h )**2 + ( $self - [ ] [ R ] )**2 - $s**2 ) / ( 2 * ( $r + $h ) * $self - [ ] [ R ] ) ) ) ;
        my $small_omega = \text{acos}( ( ( $r**2 + ( $self - [ ] [ R ] )**2 - $s**2 ) / ( 2 * $r * $self - [ ] [ R ] ) ) ) ;
        my $delta = \text{r*} ( $big_omega - $small_omega ) ;
        my $displacements = $delta / $P ; \ # Number of cols as decimal \rightarrow 
        fraction
        $displacements$[ col ] = $displacement ; \ # Store for second \rightarrow 
        pass
        my $int = \text{int}( $displacement ) ;
        if ($col - $int >= 0 \&\& ($col - $int) < $self - [ ] [ WIDTH ] ) \{ 
            $contributions$[ col - $int ] += ( 1 - $displacement + \rightarrow 
                $int) ; 
        
        } if ($col - $int - 1 >= 0 \&\& ($col - $int - 1) < $self - [ ] [ WIDTH ] ) \{ 
            $contributions$[ col - $int - 1 ] += ( $displacement - \rightarrow 
                $int) ;
        } 

# OK now for second pass to place values
for ( my $col = 0 ; $col < $self - [ ] [ WIDTH ] ; $col++ ) \{ 
    my $h = $type{DEM} - [ ] [ col ] ;
    unless ($h \text{ eq } 's' || $shadows$[ col ] == 1 ) \{ \ # because these \rightarrow 
        represent null values in DEM or shadow, respectively
        my $displacement = $displacements$[ col ] ;
        my $int = \text{int}( $displacement ) ;
        if ($col - $int >= 0 \&\& ($col - $int) < $self - [ ] [ WIDTH ] \&\& ($col - $int - 1) >= 0 \&\& ($col - $int - 1) < $self - [ ] [ WIDTH ] ) \{ 
            $contributions$[ col - $int - 1 ] += ( $displacement - \rightarrow 
                $int - 1) ;
        } if ($contributions$[ col - $int - 1 ] ) \{ \ # Just avoiding \rightarrow 
            divisions by zero
            $neurasterov$[ col ] = ( $displacement - $int ) * $type[ \rightarrow 
                DATA] - [ ] [ col ] - $int - 1) / $contributions$[ col - \rightarrow 
                $int - 1 ] ;
        } if ($contributions$[ col - $int ] ) \{ \ # again, avoiding \rightarrow 
            divisions by zero
        

    }
\[ \newrasterrow[\col] := (1 - \text{displacement} + \text{int}) \times \left( \text{type}(\text{DATA}) \right) / \text{contributions}[\col - \text{int}]; \]

```perl
$\text{int}_\text{row} = \text{map}(\text{int}, \newrasterrow); 
foreach my $\text{key} (\text{keys} \%\text{handles}) { 
    local $\text{FH} = \text{handles}[$\text{key}]; 
    my @thisrow = split(/\t|\s/,$\text{FH}); 
    $\text{type}[$\text{key}] = \@thisrow; 
}
\text{fix\_header}($\text{self});
\text{close} \text{DEM};
\text{close} \text{DATA};
\text{close} \text{LAT};
\text{close} \text{SRT};
```

**ALPHA.pm**

A perl child class of GrassAscii, which calculates local incidence angles \(\alpha\) from nominal incidence angles \(\theta\).

**Listing 3: ALPHA.pm**

```
package ALPHA;
use GrassAscii;
use Math::Trig;
```
use strict;
use warnings;

my $P = 500; # Pixel spacing

sub new {
  my $class = shift;
  my $root = shift or die "You must argue a fullpath root\n";
  my $self = $class->GrassAscii::new($root, '.ALPHA');
  $self->{ROOT} = $root;
  bless ($self, $class);
  return $self;
}

sub create_alpha {
  my $self = shift;
  my %ascii;
  print "Creating Alpha...\n";
  foreach my $type qw(DEM THETA) {
    $ascii{$type} = new GrassAscii($self->{ROOT}, '.ALPHA');
    $ascii{$type}->read_header();
  }
  die "Cols and Rows are not equal\n" unless (check_dims(%ascii));
  foreach my $type qw(DEM THETA) {
    $ascii{$type}->read_data();
  }
  my $null = 's';
  my $tempfile = $self->{ROOT}, '.TEMP';
  open TEMPDATA, ">$tempfile";
  for (my $rows = 1; $rows < ($ascii{THETA})->{ROWS} - 1; $rows++) {
    my $output_row = 0;
    for (my $cols = 1; $cols < ($ascii{THETA})->{COLS} - 1; $cols++) {
      my ($a, $b, $c, $d, $f, $g, $h, $i) = (
        $ascii{DEM}->[DATA][$rows - 1][$cols + 1],
        $ascii{DEM}->[DATA][$rows][$cols + 1],
        $ascii{DEM}->[DATA][$rows + 1][$cols + 1],
        $ascii{DEM}->[DATA][$rows - 1][$cols],
        $ascii{DEM}->[DATA][$rows + 1][$cols],
        $ascii{DEM}->[DATA][$rows - 1][$cols - 1],
        $ascii{DEM}->[DATA][$rows + 1][$cols - 1],
        $ascii{DEM}->[DATA][$rows][$cols - 1],
        $ascii{DEM}->[DATA][$rows + 1][$cols - 1]
      )
    }
    my $alpha = 0;
    unless (($a eq $null) || ($b eq $null) || ($c eq $null) || ($d eq $null) || ($f eq $null) || ($g eq $null) || ($h eq $null)) {
      my $theta = deg2rad($ascii{THETA})->{DATA}[$rows][$cols];
      my $s = $i - $a + $c - $g + $f - $d;
      my $t = $a - $i - $g + $c + $b - $h;
      my $L = sqrt($s**2 + ($P**2) + $t**2);
      my $A = $s/$L;
      my $C = (6 * $P)/$L;
      $alpha = 180 + acos(((C * cos($theta)) - ($A * sin($theta))) / 3.1415927);
    }
  }
  close TEMPDATA;
  print "\n";
}

@ISA = ("GrassAscii");

Appendix A - Scripts

192
GM.pm

A perl module that is responsible for reading a GM data file, extracting geometric parameters, and interpolating tie point data to create raster surfaces for the incidence angle at the WGS84 ellipsoid ($\theta$), the slant range time (SRT), and latitude (LAT), all of which are necessary for preprocessing calculations:
Listing 4: GM.pm

1 package GM;
2 use MY_CONFIG;
3 use Math::Trig;
4 use strict;
5 use warnings;
6
7 my $config = new MY_CONFIG('config');
8 my $PDS = $config->{PDS};
9 my $GDALINFO = $config->{GDALINFO};
10
11 sub new {
12  my $class = shift;
13  my $self = {};
14  $self->{FULLPATH} = shift;
15  $self->{WIDTH} = undef;
16  $self->{HEIGHT} = undef;
17  $self->{TIEPOINTS} = [];
18  $self->{SAMPLE_NUMBERS} = [];
19  bless ($self, $class);
20  return $self;
21 }
22
23 sub get_R {
24  my $self = shift;
25  my $fullpath = $self->{FULLPATH};
26  my @response = @{$PDS->ds 2 -field orbit_state_vectors $fullpath};
27  chomp @response;
28  my @keys = split(/\t/,$response[3]);
29  my @values = split(/\t/,$response[4]);
30  my %hash;
31  for (my $i = 0; $i <= $#keys; $i++) {
32    $hash{$keys[$i]} = $values[$i];
33  }
34  my $R = sqrt(($hash{'orbit_state_vectors[2].x_pos_1'})**2 + ($hash{'orbit_state_vectors[2].y_pos_1'})**2 + ($hash{'orbit_state_vectors[2].z_pos_1'})**2) / 100; # as they are in cm.
35  return (1 + sprintf("%.0f",$R));
36 }
37
38 sub interpolate {
39  my $self = shift;
40  my $root = $self->{FULLPATH};
41  $root =~ s/\N1//;
42  my @theta; #Nominal Incidence Angles
43  my @srt; #Slant Range Times
44  my @lat; # Latitudes
45  my @all = (theta => \@theta, srt => \@srt, lat => \@lat);
46  print "Interpolating...\n";
47  #Create empty raster arrays
48  foreach my $key (keys %all) {
49    for (my $i = 0; $i < $self->{HEIGHT}; $i++) {
50      my @temp;
51      for (my $j = 0; $j < $self->{WIDTH}; $j++) {
52        push @temp, 0;
53      }
54      push @all{$key}, \@temp;
55    }
56  }
57  # Interpolate across rows
58  # Invert across columns
foreach my $tp_row (@{$self->{TIEPOINTS}}) {
  my $this_raster_row = $self->{HEIGHT} - ($tp_row->[0]["yx"][0] + 0.5);  

  for (my $tp_col = 0; $tp_col < 10; $tp_col++) {
    my $last_sample_value;
    my $next_sample_value;
    my $last_sample_col = $self->{SAMPLE_NUMBERS}[$tp_col] - 1;
    my $next_sample_col = $self->{SAMPLE_NUMBERS}[$tp_col + 1] - 1;

    foreach my $field (keys %all) {
      $last_sample_value{$field} = $tp_row->[10]->{$field}[$field];
      $next_sample_value{$field} = $tp_row->[11]->{$field}[$field];
    }

    for (my $raster_col = $self->{SAMPLE_NUMBERS}[$tp_col] - 1; $raster_col < $self->{SAMPLE_NUMBERS}[$tp_col + 1] - 1; $raster_col++) {
      my $this_value;

      foreach my $fld (keys %all) {
        $this_value{$fld} = (($next_sample_value{$field}) - $last_sample_value{$field})
        * ($raster_col - $last_sample_col) / ($next_sample_col - $last_sample_col);
        $all{$fld}[$this_raster_row][$raster_col] = $this_value{$field};
      }
    }
  }

  foreach my $fd (keys %all) {
    $all{$fd}[$this_raster_row->[10]->{SAMPLE_NUMBERS}[10] - 1] = $tp_row->[10]->{SAMPLE_NUMBERS}[$tp_col];
  }
}

# Now interpolate down columns

for (my $rast_col = 0; $rast_col < $self->{WIDTH}; $rast_col++) {
  for (my $tiepoint_row = 0; $tiepoint_row < scalar @{$self->{TIEPOINTS}}) - 1; $tiepoint_row++) {
    my $last_sample_raster_row = $self->{HEIGHT} - @{$self->{TIEPOINTS}}[$tiepoint_row][0]["yx"][0] + 0.5;
    my $next_sample_raster_row = $self->{HEIGHT} - @{$self->{TIEPOINTS}}[$tiepoint_row][0]["yx"][0] + 0.5;

    my $proportion = ($raster_row - $last_sample_raster_row) / ($next_sample_raster_row - $last_sample_raster_row);

    foreach my $field (keys %all) {
      $this_value = $proportion
      * ($all{$field}[$next_sample_raster_row][0]["rast_col"] - $all{$field}[$last_sample_raster_row][0]["rast_col"])
      + $all{$field}[$last_sample_raster_row][0]["rast_col"];
      $all{$field}[$raster_row][0]["rast_col"] = $this_value;
    }
  }
}
. Appendix A - Scripts

```perl
109 }  
110  
111  
112  
113  
114 sub output_rasters {  
115  my ($all, $root, $self) = @_;  
116  foreach my $file (keys %$all) {  
117    output_raster($all->{$file}, uc($file), $root, $self);  
118  }  
119  
120  
121 sub output_raster {  
122  my ($data, $suffix, $root, $self) = @_;  
123  open my $output, ">$fullpath" or die "Couldn't create/edit $fullpath\n";  
124  my $east = sprintf("%.0f",$self->[HEIGHT]);  
125  my $least = sprintf("%.0f",$self->[WIDTH]);  
126  print $output <<="EOF";  
127  north: $north  
128  south: 0  
129  east: $east  
130  west: 0  
131  rows: $north  
132  cols: $east  
133  type: float  
134  null: 0  
135  EOF  
136  foreach my $datarow (@$data) {  
137    my @row = map { sprintf("%.2f",$_) } @$datarow;  
138    print $output join("\t", @row);\n";  
139  }  
140  
141  
142  
143 sub create_points_file {  
144  my $self = shift;  
145  my ($points_path, $reverse) = @_ or die "Points file path not specified. ← Aborting\n";  
146  open POINTS, ">$points_path" or die "Cannot create/edit POINTS file ← $points_path\n";  
147  print POINTS "# Nearest\tNear\tFar\tF\tF\tInclude\n";  
148  my @tiepoints = $self->[TIEPOINTS];  
149  for (my $i = 0; $i < @$tiepoints; $i++) {  
150    for (my $j = 0; $j < 11: $j++) {
151      my $x = $tiepoints->[$i][$j]{'x'}{1};  
152      my $y = $tiepoints->[$i][$j]{'y'}{0};  
153      my $long = $tiepoints->[$i][$j]{'long'};  
154      my $lat = $tiepoints->[$i][$j]{'lat'};  
155      my $local = "%x\%y";  
156      my $destination = "$long\$lat";  
157      my $output = ($reverse ? "$destination\t$local\t$1\n" : "$local←
$t$destination\t$1\n") ;  
158      print POINTS $output;  
159    }  
160 }  
161 }  
162 
163 }  
164 
165 sub open {  
166  my $self = shift;  
167  ($self->[HEIGHT], $self->[WIDTH]) = @get_dimensions($self->[FULLPATH]);
```
sub get_root { } # Appendix A - Scripts

sub get_tiepoints { } # Appendix A - Scripts

196 } # Appendix A - Scripts
process_gm.pl

The perl script which uses the methods in modules above to implement the pre-
processing of a batch of GM data files. The script runs within its own GRASS
GIS environment, creating a unique MAPSET which is destroyed once complete.
In this way the routine can be run multiple times simultaneously within one
GRASS instance, without falling foul of any locks.

Listing 5: process_gm.pl

```perl
#!/usr/bin/perl
use lib '/home/damien/Modules';
use GM;
use ORTHO;
use ALPHA;
use MY_CONFIG;
use strict;
use warnings;
my $config = new MY_CONFIG('config');
my $GISBASE = $config->{GISBASE};
my $DEM_LOCATION = $config->{DEM_LOCATION}; #'North_Aus';
my $DEM_NAME = $config->{DEM_NAME};
my $LOCATION = $config->{LOCATION};
my $time = time();
# We'll just search the current directory for GM N1 files
# Note: This must be done from within a Grass shell.
my ($BATCH) = @ARGV[0] or die "You must argue a batch number\n";
my ($search_dir) = @ARGV[1] or die "You must argue a search directory (full ←
path)\n";
```
my @fullpaths = `~/usr/bin/find $search_dir -type l -name 'ASA*.N1'`;  
chomp @fullpaths;

die "No GM files found
" unless (scalar(@fullpaths));

# OK, what follows is the main procedure — a sequence of commands, most of which are defined below in sub-procedures. I know, I know, it’s ugly, but all of the necessary (and more beautiful) algorithm structures have been set up in the modules called at the top. This is just the final implementation.

foreach my $fullpath (@fullpaths) {  
    next unless check_latitude($fullpath);
    mapset($DEM_LOCATION, "$BATCH");
    delete_old_location();
    recreate_dem_group();
    bring_in_raw($fullpath);
    target_dem();
    my $gm = new GM($fullpath);
    $gm->open();
    $gm->create_points_file("$GISDBASE/$BATCH/PERMANENT/group/temp/POINTS"); # replace dodgy imported one
    $gm->create_points_file("$GISDBASE/$DEM_LOCATION/$BATCH/group/$BATCH/POINTS".
    rectify_dem();
    rectify_output_group();
    set_target();
    copy_points_file();
    rectify_output_group();
    mapset($LOCATION, "$BATCH");
    output_tif($root);
    clean_up($root);
}

rid_gis_folders();

print "That took ",(time() - $time)." seconds\n";

sub check_latitude {  
    my $fullpath = shift;
    my @lats;
    my @result = `gdalinfo $fullpath`;
    foreach my $line (@result) {#(595.5,480.5) -> (-70.374488,-10.937347,0)
        if ($line =~ /\((-?[\d.]+,-?[\d.]+)\)/) {
            push @lats, 1;
        }
    }
Appendix A - Scripts

```perl
# @articles = sort {&b <=> &a} @files;
my @sorted = sort {&a <=> &b} @lats;
return (@sorted[0] < 60);

sub rid_gis_folders {
    unlink "$GISDBASE/L\&BATCH";
    unlink "$GISDBASE/LOCATION/M\&BATCH";
    unlink "$GISDBASE/DEM\&LOCATION/M\&BATCH";
}

sub clean_up {
    my $root = shift;
    my @extensions = qw(ALPHA ORTHO THETA LAT SRT DATA DEM TEMP);
    my @files = map("$root.$_", @extensions);
    unlink @files;
}

sub output_tif {
    my $root = shift;
    my $command = "<<EOF"
    g.region zoom=ORTHO.r
    r.mapcalc 'ALPHA = round(100 * ALPHA.r)'
    r.mapcalc 'ORTHO = round(ORTHO.r)'
    g.remove group=output
    i.group group=output input=ORTHO.ALPHA
    r.out.gdal -c input=output format=GTiff type=UInt16 output="$root.tif"
    EOF
    system $command;
}

sub rectify_output_group {
    system "i.rectify -a group=output extension=r order=3 ";
}

sub copy_points_file {
    system "/bin/cp "$GISDBASE/L\&BATCH/PERMANENT/group/temp/POINTS $GISDBASE/\&L\&BATCH/PERMANENT/group/output/POINTS ";
}

sub create_output_group {
    system "i.group group=output input=ORTHO,ALPHA ";
}

sub import_asciis {
    my $root = shift;
    my $command = "<<EOF"
    r.in.ascii input=$root.ORTHO output=ORTHO
    r.in.ascii input=$root.ALPHA output=ALPHA
    EOF
    system $command;
}

sub set_target {
    system "i.target group=output location=\&LOCATION mapset=M\&BATCH ";
}

sub ascify_data_and_dem {
    my ($root) = @_;  
    system "r.out.ascii -i input=$DEM\&NAME.r output="$root.DEM null=\" ";  
    system "r.out.ascii -i input=temp.1 output="$root.DATA null=0 ";
    
}```
This perl script looks for all GM data files under the current directory and separates them into batches, creating a directory for each batch and symbolic links within each to those folders. Each batch will then be processed concurrently. Two batch scripts are then created by the main script for each batch - one being the instructions for Grass GIS, the other being the PBS batch script. The latter is submitted to the HPC server. The start of each PBS script sets the GRASS_BATCH_JOB variable to point to the individual Grass batch script, and
starts a new instance of GRASS GIS, creating a unique temporary MAPSET within which to work.

Listing 6: pbs.pl

```perl
#!/usr/bin/perl
use strict;
use warnings;

my @GMs = `find \`pwd\` -type f -name 'ASA_GMs.N1'`;
chomp @GMs;

my ($prefix, $number_of_batches) = @ARGV;
my $lastbatch = ($#GMs > $number_of_batches ? $number_of_batches : $#GMs);
my $files_per_folder = int((#GMs + 1) / ($lastbatch + 1));
my $index = 0;
my $localdir = 'pwd';
my @folders;

for (my $batch = 0; $batch <= $lastbatch; $batch++) {
    mkdir "$prefix$batch" or die "Couldn't create folder $prefix$batch\n";
    push @folders, $prefix$batch;
    system "ln -s $GMs[$index] $prefix$batch/";
    $index++;
}

# Now to get remainders, if there are any
for (my $i = $index; $i <= $#GMs; $i++) {
    my $put_it_in = $#GMs - $i;
    system "ln -s $GMs[$i] $prefix$put_it_in/";
}

foreach my $folder (@folders) {
    open my $grassbatch, ">grassbatch$folder" or die "Couldn't create ← grassbatch$folder\n";
    print $grassbatch "#!/bin/bash";
    print $grassbatch "process_gm.pl $folder \`pwd\`/$folder\n";
    close $grassbatch;
    chmod 0755, "grassbatch$folder";
    open my $pbsbatch, ">pbsbatch$folder" or die "Couldn't create ← pbsbatch$folder\n";
    print $pbsbatch "#!/bin/bash";
    print $pbsbatch "#!/bin/bash\n";
    #PBS -c
    #PBS -j oe
    #PBS -m ae
    #PBS -N Preprocess_GM_Files
    #PBS@damien.ogrady@my.jcu.edu.au
    #PBS -l walltime=9999:00:00
    #PBS -l pmem=8gb
    echo "  ";
    echo "  This job is allocated 1 cpu on "
    cat $PBS_NODEFILE
    echo "  ";
```

regression.pl

This perl script uses GDAL to carry out regression as described in section 3.4.1. GRASS GIS is only used at the very end to create visual interpretations.

Listing 7: regression.pl

```perl
#!/usr/bin/perl
use strict;
use warnings;
use lib '/home/damien/Modules';
use GRASSSESSION;
use Geo::GDAL;
use MY_CONFIG;

my $SPECIFIC_DIRECTORY = undef; # speaks for itself. Comes from command line ←
                                 #NB!! NOT full path, just a name
my $raw_files = get_arguments();
my $config = new MY_CONFIG('config');

my $RESOLUTION = $config->{RESOLUTION}; #Full will be 0.05
my $TARGET_DIRECTORY = $config->{TARGET_DIRECTORY}; # Root location for ←
                                                      #regression, world and timestamp maps
my $MASK_DIRECTORY = $config->{MASK_DIRECTORY}; # location of the MASK.tif file
my $GISDBASE = $config->{GISDBASE};
my $LOCATION = $config->{LOCATION};
my ($N, $S, $E, $W) = ($config->{N}, $config->{S}, $config->{E}, $config->{W}); # ←
         # Defining the north, south, east and west limits
my $WKT = $config->{WKT};
$config = undef;

### This is the main procedure ###

foreach my $fullpath (@$raw_files) {
my $root = get_root($fullpath) or die "Something wrong with the full path $fullpath
";
if ($root) {
my $dn_dataset = Geo::GDAL::Open($fullpath);
if (within_bounds($dn_dataset)) {
  #
  eval {
    my $clipped_dn_dataset = clip_dn($root, $dn_dataset); #contains dn and alpha
    my $sigma_alpha_dataset = create_sigmaalpha($root, $clipped_dn_dataset);
    update_regression($sigma_alpha_dataset, $clipped_dn_dataset);
  }
  #
  }
  #
  print LOG "Error for $fullpath: $@
" if $@;
} else {
  print "File with root $root not within bounds
"
}
delete_temporary_files($root);
}
grass_calc_and_output();
### End of main procedure ###
sub grass_calc_and_output {
my $grass = new GRASSSESSION($GISDBASE);
$grass->setLocation($LOCATION);
foreach my $prefix qw(N SumX SumX2 SumY SumY2 SumXY SumGamma) {
  $grass->addCommand("r.in.gdal input=TARGET_DIRECTORY/$SPECIFIC_DIRECTORY/$prefix.tif output=$prefix overwrite");
  #
  $grass->addCommand("rm TARGET_DIRECTORY/$SPECIFIC_DIRECTORY/$prefix.tif");
}
my $command = <<"GCAO"
  r.mapcalc 'B = eval (numerator = (N * SumXY) - (SumX * SumY), denominator = (N * SumX2) - pow(SumX, 2), numerator / denominator)'
  r.mapcalc 'A = (SumY - (B * SumX)) / N'
  r.mapcalc 'R2 = eval (numerator = (N * SumXY) - (SumX * SumY), denominator = sqrt((N * SumX2) - pow(SumX, 2)) * sqrt((N * SumY2) - pow(SumY, 2)), R = numerator / denominator, pow(R, 2))'
  r.mapcalc 'SD = sqrt((SumY2 / N) - pow(SumY / N, 2))'
  r.colors.stddev SD
  r.colors.stddev B
  r.mapcalc 'MEANGAMMA = eval(mg = SunGamma / N, \n    if(mg > 255, 255, mg))'
  echo '0% red
  0 red
  70 green
  140 0 128 0
  160 blue
  255 0 0 128' | r.colors MEANGAMMA col=rules
  r.out.png input=MEANGAMMA output=TARGET_DIRECTORY/
    MEANGAMMA_SPECIFIC_DIRECTORY.png
  r.out.png input=SD output=TARGET_DIRECTORY/SD_MEANGAMMA_SPECIFIC_DIRECTORY.png
GCAO
$grass->addCommand($command);
foreach my $suffix qw(B A R2 MEANGAMMA SD) {
  my $type = ($suffix eq 'OUTPUT') ? 'Byte' : 'Float32';
Appendix A - Scripts

```perl
$grass->addCommand("r.out.gdal input=$suffix output=$TARGET_DIRECTORY/\$SPECIFIC_DIRECTORY/$suffix.tiff type=file format=GTiff nodata=0");
%
$grass->run();
%
sub get_root {
    my $fullpath = shift;
    if ($fullpath =~ /(ASA_GM.+)/) {
        return $1;
    }
%
sub get_mask_data {
    my ($clipped_dn_dataset) = @_;;
    my $mask_dataset = Geo::GDAL::Open("$MASK_DIRECTORY/MASK.tif") or die "←
        Couldn't open mask $MASK_DIRECTORY/MASK.tif\n";
    my $mask_band = $mask_dataset->GetRasterBand(1);
    my ($minx, $maxx, $miny, $maxy, $dx, $dy, $width, $height, $xoff, $yoff) = ←
        @get_gridding_data($clipped_dn_dataset);
    my $data = $mask_band->ReadTile($xoff, $yoff, $width, $height);
    return $data;
%
sub delete_temporary_files {
    my ($root) = @_;;
    foreach ('ALPHA', 'SIGMAALPHA', 'DN', 'CLIP') {
        unlink("$_.$root.tif");
    }
%
sub clip_dn {
    my ($root, $dn_dataset) = @_;;
    my ($minx, $maxx, $miny, $maxy, $dx, $dy, $width, $height, undef, undef) = ←
        @get_gridding_data($dn_dataset);
    my $west = $minx < $W ? $W : $minx;
    my $east = $maxx > $E ? $E : $maxx;
    my $north = $maxy > $N ? $N : $maxy;
    my $south = $miny < $S ? $S : $miny;
    my $xoff = sprintf("%.0f",( $west - $minx )/$dx);
    my $yoff = sprintf("%.0f",abs(( $north - $maxy )/$dy));
    my $xsize = sprintf("%.0f",($east - $west)/$dx);
    my $ysize = abs( sprintf("%.0f",($north - $south)/$dy));
    my $dn_band = $dn_dataset->GetRasterBand(1);
    my $alpha_band = $dn_dataset->GetRasterBand(2);
    my $dn_data = $dn_band->ReadTile($xoff, $yoff, $xsize, $ysize);
    my $alpha_data = $alpha_band->ReadTile($xoff, $yoff, $xsize, $ysize);
    my $dataset = Geo::GDAL::Driver('GTiff')->Create("CLIP$root.tiff", $xsize, ←
        $ysize, 2, 'UInt16');
    $dataset->GeoTransform($west, $RESOLUTION, 0, $north, 0, -$RESOLUTION);
    $dataset->SetProjection($WKT);
    my $dn_band_out = $dataset->Band(1);
    $dn_band_out->NoDataValue(65535);
    $dn_band_out->WriteTile($dn_data);
    my $alpha_band_out = $dataset->Band(2);
    $alpha_band_out->NoDataValue(65535);
```
1.

```plaintext
for (my $Row=0; $Row<$height; $Row++) {
  for (my $Col=0; $Col<$width; $Col++) {
    if ($Sigma_alpha_data->[[$Row][[$Col]]] &
        $alpha_data->[[$Row][[$Col]]] != 65535) {
      my $Y = $Sigma_alpha_data->[[$Row][[$Col]]] &
        $alpha_data->[[$Row][[$Col]]] / 100;
      $alpha = 360 if $alpha < 0;
    }
  }
}
```

2.

```plaintext
my ($dataset) = 0;
my ($minx, $dx, undef, $maxx, undef, $dy) = $dataset->GeoTransform();
my ($width, $height) = $dataset->Size;
my ($maxx, $miny) = (($minx + ($dx * $width)), ($maxx - (abs($dy) * $height← )));
my $xoff = sprintf("%.0f",($minx - $W) / $dx);
my $yoff = sprintf("%.0f",($maxx - $H) / $dy);
my $output = [[$minx, $maxx, $miny, $maxy, $dx, $dy, $width, $height, $xoff, ← $yoff];
return $output;
```

3.

```plaintext
my ($dn, $dataset) = 0;
my ($minx, $maxx, $miny, $maxy) = @{$get_gridding_data($dn, $dataset)};
return ($miny < $N) & (maxy < $N) & (minx < $W) & (maxx > $W);
```

4.

```plaintext
my ($N_band, $SumY_band, $SumX_band, $SumXY_band, $SumX2_band, $SumY2_band, $SumGamma_band) = @{$get_regression_bands($alpha_band)};
```

5.

```plaintext
my ($dx, $dy, $width, $height, $xoff, $yoff) = @{$get_gridding_data($sigma_alpha_dataset)};
my $xoff = $xoff * $width, $yoff = $yoff * $height;
```

6.

```plaintext
my $N_data = $N_band->ReadTile($xoff, $yoff, $width, $height);  
my $SumY_data = $SumY_band->ReadTile($xoff, $yoff, $width, $height);  
my $SumX_data = $SumX_band->ReadTile($xoff, $yoff, $width, $height);  
my $SumXY_data = $SumXY_band->ReadTile($xoff, $yoff, $width, $height);  
my $SumX2_data = $SumX2_band->ReadTile($xoff, $yoff, $width, $height);  
my $SumY2_data = $SumY2_band->ReadTile($xoff, $yoff, $width, $height);  
my $SumGamma_data = $SumGamma_band->ReadTile($xoff, $yoff, $width, $height)
```

7.

```plaintext
return $dataset;
```

8.

```plaintext
$alpha_band = $clipped_dn_dataset->GetRasterBand(1);
my $alpha_nodata = $alpha_band->GetNoDataValue;
my $Sigma_alpha_data = $Sigma_alpha_band->ReadTile;
my $alpha_band = $clipped_dn_dataset->GetRasterBand(2);
my $alpha_nodata = $alpha_band->GetNoDataValue;
my $alpha_data = $alpha_band->ReadTile;
print "Updating regression...\n";
```

9.

```plaintext
get_gridding_data {
  my ($dataset) = 0;
  my ($minx, $dx, undef, $maxx, undef, $dy) = $dataset->GeoTransform();
  my ($width, $height) = $dataset->Size;
  my ($maxx, $miny) = (($minx + ($dx * $width)), ($maxx - (abs($dy) * $height← )));
  my $xoff = sprintf("%.0f",($minx - $W) / $dx);
  my $yoff = sprintf("%.0f",($maxx - $H) / $dy);
  my $output = [[$minx, $maxx, $miny, $maxy, $dx, $dy, $width, $height, $xoff, ← $yoff];
  return $output;
}
```

10.

```plaintext
update_regression {
  # In this case using psi – a power of cosine alpha.
  See notes 1 Dec 09.
  my ($sigma_alpha_dataset, $clipped_dn_dataset) = 0;
  my $regression_dataset = get_regression_dataset();
  my ($N_band, $SumY_band, $SumX_band, $SumXY_band, $SumX2_band, $SumY2_band, $SumGamma_band) = @{$get_regression_bands($alpha_band)};
  my ($minx, $maxx, $miny, $maxy, $dx, $dy, $width, $height, $xoff, $yoff) = @{$get_gridding_data($sigma_alpha_dataset)};
  my $N_data = $N_band->ReadTile($xoff, $yoff, $width, $height);
  my $SumY_data = $SumY_band->ReadTile($xoff, $yoff, $width, $height);
  my $SumX_data = $SumX_band->ReadTile($xoff, $yoff, $width, $height);
  my $SumXY_data = $SumXY_band->ReadTile($xoff, $yoff, $width, $height);
  my $SumX2_data = $SumX2_band->ReadTile($xoff, $yoff, $width, $height);
  my $SumY2_data = $SumY2_band->ReadTile($xoff, $yoff, $width, $height);
  my $SumGamma_data = $SumGamma_band->ReadTile($xoff, $yoff, $width, $height);
  print "Updating regression...\n";
  for (my $Row=0; $Row<$height; $Row++) {
    for (my $Col=0; $Col<$width; $Col++) {
      if ($Sigma_alpha_data->[[$Row][[$Col]]] &
          $alpha_data->[[$Row][[$Col]]] != 65535) {
        my $Y = $Sigma_alpha_data->[[$Row][[$Col]]] &
          $alpha_data->[[$Row][[$Col]]] / 100;
        $alpha = 360 if $alpha < 0;
      }
    }
  }
}
```
. Appendix A - Scripts

191 $X = -100 \ast \log_{10}(\cos(3.1415927 \ast \alpha / 180));
192 $\gamma = Y - X;
193 \$N_data\rightarrow\$Row[\$Col] += 1;
194 \$SumY_data\rightarrow\$Row[\$Col] += \$Y;
195 \$SumX_data\rightarrow\$Row[\$Col] += \$X;
196 \$SumXY_data\rightarrow\$Row[\$Col] += (\$X \ast \$Y);
197 \$SumX2_data\rightarrow\$Row[\$Col] += (\$X ** 2);
198 \$SumY2_data\rightarrow\$Row[\$Col] += (\$Y ** 2);
199 \$SumGamma_data\rightarrow\$Row[\$Col] += $\gamma$;
200 
201 
202 
203 \$N_band\rightarrow\WriteTile(\$N_data, \$xoff, \$yoff);
204 \$SumY_band\rightarrow\WriteTile(\$SumY_data, \$xoff, \$yoff);
205 \$SumX_band\rightarrow\WriteTile(\$SumX_data, \$xoff, \$yoff);
206 \$SumXY_band\rightarrow\WriteTile(\$SumXY_data, \$xoff, \$yoff);
207 \$SumX2_band\rightarrow\WriteTile(\$SumX2_data, \$xoff, \$yoff);
208 \$SumY2_band\rightarrow\WriteTile(\$SumY2_data, \$xoff, \$yoff);
209 \$SumGamma_band\rightarrow\WriteTile(\$SumGamma_data, \$xoff, \$yoff);
210 
211 
212 
213 # $N_band, $SumY_band, $SumX_band, $SumXY_band, $SumX2_band, $psi_band
214 
215 \# Byte/Int16/UInt16/Int32/UInt32/Float32/Float64/
216 \# CInt16/CInt32/CFloat32/CFloat64
217 sub get_regression_bands 
218 my $bands = [ 
219 [N, 'UInt16'],
220 [SumY, 'UInt16'],
221 [SumX, 'Float32'],
222 [SumXY, 'Float32'],
223 [SumX2, 'Float32'],
224 [SumY2, 'Float32'],
225 [SumGamma, 'UInt16'],
226 ];
227 my @output;
228 foreach (@$bands) {
229 my ($name, $datatype) = @$_;
230 my $path = "$TARGET_DIRECTORY/SPECIFIC_DIRECTORY";
231 if (-e "$path/$name.tif") {
232 my $dataset = Geo::GDAL::Open("$path/$name.tif", 'Update');
233 my $band = $dataset->GetRasterBand(1);
234 push (@output, $band);
235 } else {
236 my $width = sprintf("%.1f",($W - $W) / $RESOLUTION);
237 my $height = sprintf("%.1f",($H - $S) / $RESOLUTION);
238 print "Creating $path/$name.tif...\n"
239 my $dataset;
240 # eval
241 mkdir $path unless -d $path;
242 my $dataset = Geo::GDAL::Driver ('GTiff')\Create("$path/$name.tif", $width, $height, 1, $datatype);
243 #
244 # print STDERR "Error: @\n"
245 # $dataset->GeoTransform($W, $RESOLUTION, 0, $H, 0, -$RESOLUTION);
246 $dataset->SetProjection($WKT);
247 my $band = $dataset->Band(1);
248 $band->Fill(0);
249 $band->NoDataValue(0);
250 push (@output, $band);
sub create_sigmaalpha {
    my ($root, $clipped_dn_dataset) = @_;
    my ($width, $height) = $clipped_dn_dataset->Size;
    my ($minx, $dx, undef, $maxy, undef, $dy) = $clipped_dn_dataset->GeoTransform();
    my $alpha_band = $clipped_dn_dataset->GetRasterBand(2);
    my $alphas = $alpha_band->ReadTile;
    my $dn_band = $clipped_dn_dataset->GetRasterBand(1);
    my $dn_nodata = $dn_band->GetNoDataValue;
    my $dns = $dn_band->ReadTile;
    my $mask_data = get_mask_data($clipped_dn_dataset);

    my $name = "SIGMAALPHA$r.root.tif";
    my $datatype = 'UInt16';
    my $nodata = 0;
    my $data = ();

    my $dataset = Geo::GDAL::Driver('GTiff')->Create($name, $width, $height, 1, $datatype);
    $dataset->GeoTransform($minx, $dx, 0, $maxy, 0, $dy);
    $dataset->SetProjection(WKT);
    $band = $dataset->Band(1);
    $band->NoDataValue($nodata);

    print "Creating SigmaAlpha for $root...
";

    # die "DN and ALPHA maps are different sizes\n" unless ($alpha_dataset->Size == ($width, $height));

    for (my $row=0; $row<$height; $row++) {
        my @rowvalue;
        for (my $col=0; $col<$width; $col++) {
            my $value = $nodata;
            if (($dns->[0][$row][$col] != 65535) &
                ($alphas->[0][$row][$col] != 65535) &
                ($mask_data->[0][$row][$col] != 65535) &
                ($mask_data->[0][$row][$col] != 65535)) {
                my $alpha = $alphas->[0][$row][$col] / 100;
                $value = sprintf("%.0f", (-100 * log10(((($dns->[0][$row][$col] > 2)/21900000) * sin((3.1415927/180) * $alpha)))));
            } else if ($value < 0) {
                $value = 0;
            } else if ($value > 65534) {
                $value = 65534;
            }
            push (@rowvalue, $value);
        }
        push (@rowvalue, $value);
    }

    $band->WriteTile(@data, @rowvalue);
    return $dataset;
}


309  sub log10 {
310      my $n = shift;
311      return log($n)/log(10) unless $n < 0;
312  }
313  
314  sub get_arguments {
315      die "You must argue a directory and some raw data filename(s)\n"
316         if (@ARGV) == 0;
317      $SPECIFIC_DIRECTORY = shift @ARGV;
318      my @output;
319      foreach my $raw (@ARGV) {
320          die "Sorry, but the argument(s) must contain the FULL, not RELATIVE, path\n"
321              unless $raw =~ /\//;
322          push(@output, $raw);
323      }
324      return @output;
325  }
Appendix B - Data

All GM data was acquired systematically via download from ESA’s Kiruna and ESRIN ground stations, made available in a two-week moving window through the Category 1 Fast Track Registration agreement, under the project number C1P.5908 (ESA, 2007a).

All MODIS data was obtained from http://lpdaac.usgs.gov, maintained by the NASA Land Processes Distributed Active Archive Center (LP DAAC) at the USGS/Earth Resources Observation and Science (EROS) Center, Sioux Falls, South Dakota. 2003. Data for the images were provided by NASA.

Chapter 3

Radar data


Dominant vegetation species, vegetation growth form

Title: Vegetation - Post-European Settlement (1988)

Custodian: Geoscience Australia

Geology
Title: Surface geology of Australia 1:1,000,000 scale, Queensland - 2nd edition
Authors: Whitaker, A.J., Champion, D.C., Sweet, I.P., Kilgour, P. and Connolly, D.P.
Custodian: Geoscience Australia
Further information: http://www.ga.gov.au

Soil type
Title: Digital Atlas of Australian Soils - Soil Landscapes Map
Custodian: Department of Agriculture Fisheries and Forestry; Australian Bureau of Agricultural and Resource Economics and Sciences

Regolith
Title: Regolith landforms polygon dataset AGSOCAT (AGSO catalogue) Number: 21805
Custodian: Geoscience Australia

MODIS

<table>
<thead>
<tr>
<th>Satellite</th>
<th>Year</th>
<th>Day</th>
<th>H</th>
<th>V</th>
</tr>
</thead>
<tbody>
<tr>
<td>Aqua</td>
<td>2011</td>
<td>130</td>
<td>30</td>
<td>12</td>
</tr>
<tr>
<td>Aqua</td>
<td>2011</td>
<td>130</td>
<td>31</td>
<td>10</td>
</tr>
<tr>
<td>Aqua</td>
<td>2011</td>
<td>130</td>
<td>31</td>
<td>11</td>
</tr>
</tbody>
</table>
Chapter 5

Radar data

The following lists, each entry comprises 5 numbers representing the following:

1. Date of acquisition (yyyymmdd)
2. Time of acquisition (hhmmss)
3. Orbit cycle
4. Orbit track
5. Absolute orbit
<table>
<thead>
<tr>
<th>Date</th>
<th>Time</th>
<th>Year</th>
<th>Month</th>
<th>Day</th>
<th>Hour</th>
<th>Minute</th>
<th>Second</th>
<th>Hour</th>
<th>Minute</th>
<th>Second</th>
<th>Hour</th>
<th>Minute</th>
<th>Second</th>
</tr>
</thead>
<tbody>
<tr>
<td>20060403</td>
<td>00450</td>
<td>046</td>
<td>00288</td>
<td>21384</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>20070212</td>
<td>00458</td>
<td>055</td>
<td>00288</td>
<td>25893</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>20090203</td>
<td>123426</td>
<td>076</td>
<td>00109</td>
<td>36235</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>20090213</td>
<td>140138</td>
<td>076</td>
<td>00253</td>
<td>36579</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>20090217</td>
<td>134576</td>
<td>076</td>
<td>00310</td>
<td>36436</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>20090219</td>
<td>123234</td>
<td>076</td>
<td>00336</td>
<td>36464</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>20090221</td>
<td>004813</td>
<td>076</td>
<td>00360</td>
<td>36486</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>20090222</td>
<td>141833</td>
<td>076</td>
<td>00382</td>
<td>36508</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>20090223</td>
<td>012659</td>
<td>076</td>
<td>00390</td>
<td>36515</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>20090225</td>
<td>124234</td>
<td>076</td>
<td>00425</td>
<td>36550</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>20090228</td>
<td>124952</td>
<td>076</td>
<td>00468</td>
<td>36594</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>20090301</td>
<td>013916</td>
<td>076</td>
<td>00475</td>
<td>36601</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>20090302</td>
<td>132510</td>
<td>076</td>
<td>00496</td>
<td>36622</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>20090304</td>
<td>0001157</td>
<td>076</td>
<td>00614</td>
<td>36643</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>20090307</td>
<td>122910</td>
<td>076</td>
<td>00666</td>
<td>36693</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>20090309</td>
<td>004520</td>
<td>076</td>
<td>00884</td>
<td>36715</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>20090310</td>
<td>124328</td>
<td>076</td>
<td>00909</td>
<td>36736</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>20090312</td>
<td>005106</td>
<td>076</td>
<td>01131</td>
<td>36758</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>20090313</td>
<td>124103</td>
<td>076</td>
<td>01152</td>
<td>36779</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>20090315</td>
<td>005747</td>
<td>076</td>
<td>01747</td>
<td>36801</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>20090316</td>
<td>124532</td>
<td>076</td>
<td>01905</td>
<td>36822</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>20090317</td>
<td>135559</td>
<td>076</td>
<td>02010</td>
<td>36837</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>20090319</td>
<td>0031152</td>
<td>076</td>
<td>02311</td>
<td>36858</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>20090320</td>
<td>140137</td>
<td>076</td>
<td>02521</td>
<td>36880</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>20090323</td>
<td>000446</td>
<td>076</td>
<td>02848</td>
<td>36915</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>20090324</td>
<td>133455</td>
<td>076</td>
<td>03100</td>
<td>36937</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>20090326</td>
<td>001144</td>
<td>076</td>
<td>03311</td>
<td>36958</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>20090326</td>
<td>234112</td>
<td>076</td>
<td>03455</td>
<td>36972</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>20090329</td>
<td>001751</td>
<td>076</td>
<td>03740</td>
<td>37001</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>20090330</td>
<td>134627</td>
<td>076</td>
<td>03960</td>
<td>37023</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>20090401</td>
<td>002323</td>
<td>076</td>
<td>04177</td>
<td>37044</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>20090401</td>
<td>231332</td>
<td>076</td>
<td>04311</td>
<td>37058</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>20090403</td>
<td>005943</td>
<td>076</td>
<td>04464</td>
<td>37073</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>20090403</td>
<td>234564</td>
<td>076</td>
<td>04747</td>
<td>37101</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>20090406</td>
<td>010603</td>
<td>076</td>
<td>04891</td>
<td>37116</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>20090407</td>
<td>000154</td>
<td>076</td>
<td>05010</td>
<td>37137</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>20090409</td>
<td>140428</td>
<td>076</td>
<td>05024</td>
<td>37152</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>20090410</td>
<td>133909</td>
<td>076</td>
<td>05908</td>
<td>37159</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>20090411</td>
<td>141104</td>
<td>076</td>
<td>06076</td>
<td>37180</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>20090412</td>
<td>137375</td>
<td>076</td>
<td>06811</td>
<td>37209</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>20090414</td>
<td>001313</td>
<td>076</td>
<td>07002</td>
<td>37230</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>20090414</td>
<td>234339</td>
<td>076</td>
<td>07116</td>
<td>37244</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>20090416</td>
<td>131044</td>
<td>076</td>
<td>07386</td>
<td>37266</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>20090417</td>
<td>234846</td>
<td>076</td>
<td>07587</td>
<td>37287</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>20090420</td>
<td>124522</td>
<td>076</td>
<td>08195</td>
<td>37323</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>20090421</td>
<td>013618</td>
<td>076</td>
<td>08203</td>
<td>37331</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>20090422</td>
<td>132216</td>
<td>076</td>
<td>08224</td>
<td>37352</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>20090423</td>
<td>235919</td>
<td>076</td>
<td>08245</td>
<td>37373</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>20090425</td>
<td>132803</td>
<td>076</td>
<td>08267</td>
<td>37395</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

Continued on next page
<table>
<thead>
<tr>
<th>Date</th>
<th>Time</th>
<th>Day</th>
<th>Month</th>
<th>Year</th>
<th>Value1</th>
<th>Value2</th>
<th>Value3</th>
<th>Value4</th>
<th>Value5</th>
</tr>
</thead>
<tbody>
<tr>
<td>20090427</td>
<td>122630</td>
<td>078</td>
<td>04</td>
<td>2009</td>
<td>078295</td>
<td>37423</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>20090428</td>
<td>011434</td>
<td>078</td>
<td>04</td>
<td>2009</td>
<td>03003</td>
<td>37431</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>20090430</td>
<td>122321</td>
<td>078</td>
<td>04</td>
<td>2009</td>
<td>03388</td>
<td>37466</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>20090502</td>
<td>130775</td>
<td>078</td>
<td>05</td>
<td>2009</td>
<td>03677</td>
<td>37495</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>20090503</td>
<td>141833</td>
<td>079</td>
<td>05</td>
<td>2009</td>
<td>03882</td>
<td>37510</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>20090505</td>
<td>005359</td>
<td>079</td>
<td>05</td>
<td>2009</td>
<td>04041</td>
<td>37531</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>20090506</td>
<td>124224</td>
<td>079</td>
<td>05</td>
<td>2009</td>
<td>04224</td>
<td>37552</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>20090507</td>
<td>013315</td>
<td>079</td>
<td>05</td>
<td>2009</td>
<td>04332</td>
<td>37560</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>20090509</td>
<td>124932</td>
<td>079</td>
<td>05</td>
<td>2009</td>
<td>04677</td>
<td>37595</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>20090510</td>
<td>135845</td>
<td>080</td>
<td>05</td>
<td>2009</td>
<td>04825</td>
<td>37610</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>20090512</td>
<td>003354</td>
<td>080</td>
<td>05</td>
<td>2009</td>
<td>05002</td>
<td>37631</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>20090513</td>
<td>233301</td>
<td>079</td>
<td>05</td>
<td>2009</td>
<td>06300</td>
<td>37659</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>20090515</td>
<td>125928</td>
<td>079</td>
<td>05</td>
<td>2009</td>
<td>06524</td>
<td>37681</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>20090517</td>
<td>011724</td>
<td>080</td>
<td>05</td>
<td>2009</td>
<td>06774</td>
<td>37703</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>20090518</td>
<td>135057</td>
<td>079</td>
<td>05</td>
<td>2009</td>
<td>06959</td>
<td>37724</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>20090519</td>
<td>141546</td>
<td>079</td>
<td>05</td>
<td>2009</td>
<td>07110</td>
<td>37739</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>20090520</td>
<td>134339</td>
<td>079</td>
<td>05</td>
<td>2009</td>
<td>07424</td>
<td>37753</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>20090522</td>
<td>002204</td>
<td>079</td>
<td>05</td>
<td>2009</td>
<td>07774</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>20090523</td>
<td>134934</td>
<td>079</td>
<td>05</td>
<td>2009</td>
<td>08017</td>
<td>37796</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>20090525</td>
<td>002612</td>
<td>079</td>
<td>05</td>
<td>2009</td>
<td>08188</td>
<td>37817</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>20090527</td>
<td>235412</td>
<td>079</td>
<td>05</td>
<td>2009</td>
<td>08202</td>
<td>37831</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>20090530</td>
<td>010317</td>
<td>079</td>
<td>05</td>
<td>2009</td>
<td>08217</td>
<td>37846</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>20090601</td>
<td>140726</td>
<td>079</td>
<td>06</td>
<td>2009</td>
<td>08326</td>
<td>37899</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>20090602</td>
<td>134545</td>
<td>079</td>
<td>06</td>
<td>2009</td>
<td>08310</td>
<td>37939</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>20090604</td>
<td>001028</td>
<td>079</td>
<td>06</td>
<td>2009</td>
<td>08331</td>
<td>37960</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>20090605</td>
<td>134051</td>
<td>079</td>
<td>06</td>
<td>2009</td>
<td>08353</td>
<td>37982</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>20090607</td>
<td>001775</td>
<td>079</td>
<td>06</td>
<td>2009</td>
<td>08374</td>
<td>38003</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>20090607</td>
<td>234612</td>
<td>079</td>
<td>06</td>
<td>2009</td>
<td>08388</td>
<td>38017</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>20090618</td>
<td>131100</td>
<td>080</td>
<td>06</td>
<td>2009</td>
<td>08388</td>
<td>38168</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>20090620</td>
<td>141019</td>
<td>080</td>
<td>06</td>
<td>2009</td>
<td>08607</td>
<td>38197</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>20090622</td>
<td>004522</td>
<td>080</td>
<td>06</td>
<td>2009</td>
<td>08842</td>
<td>38218</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>20090623</td>
<td>123430</td>
<td>080</td>
<td>06</td>
<td>2009</td>
<td>10019</td>
<td>38239</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>20090624</td>
<td>012349</td>
<td>080</td>
<td>06</td>
<td>2009</td>
<td>10177</td>
<td>38247</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>20090625</td>
<td>130408</td>
<td>080</td>
<td>06</td>
<td>2009</td>
<td>10138</td>
<td>38268</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>20090704</td>
<td>132805</td>
<td>080</td>
<td>07</td>
<td>2009</td>
<td>08267</td>
<td>38397</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>20090706</td>
<td>122630</td>
<td>080</td>
<td>07</td>
<td>2009</td>
<td>08295</td>
<td>38425</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>20090707</td>
<td>011436</td>
<td>080</td>
<td>07</td>
<td>2009</td>
<td>08303</td>
<td>38433</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>20090708</td>
<td>130210</td>
<td>080</td>
<td>07</td>
<td>2009</td>
<td>08324</td>
<td>38454</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>20090711</td>
<td>004814</td>
<td>080</td>
<td>07</td>
<td>2009</td>
<td>08360</td>
<td>38490</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>20090712</td>
<td>128314</td>
<td>080</td>
<td>07</td>
<td>2009</td>
<td>08381</td>
<td>38511</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>20090713</td>
<td>134630</td>
<td>080</td>
<td>07</td>
<td>2009</td>
<td>08396</td>
<td>38526</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>20090715</td>
<td>002326</td>
<td>080</td>
<td>07</td>
<td>2009</td>
<td>08417</td>
<td>38547</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>20090717</td>
<td>123173</td>
<td>080</td>
<td>07</td>
<td>2009</td>
<td>08431</td>
<td>38561</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>20090717</td>
<td>131928</td>
<td>080</td>
<td>07</td>
<td>2009</td>
<td>08453</td>
<td>38583</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>20090718</td>
<td>235646</td>
<td>080</td>
<td>07</td>
<td>2009</td>
<td>08474</td>
<td>38604</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>20090720</td>
<td>010607</td>
<td>080</td>
<td>07</td>
<td>2009</td>
<td>08489</td>
<td>38619</td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

continued on next page
### Appendix B - Data

GM Data used in Chapter 5

<table>
<thead>
<tr>
<th>Date</th>
<th>Time</th>
<th>Day</th>
<th>Hour</th>
<th>Minute</th>
<th>Second</th>
<th>Hourly Value</th>
</tr>
</thead>
<tbody>
<tr>
<td>20090722</td>
<td>02:00</td>
<td>01</td>
<td>01</td>
<td>00:00</td>
<td>00:00</td>
<td>000158</td>
</tr>
<tr>
<td>20090723</td>
<td>02:00</td>
<td>01</td>
<td>02</td>
<td>00:00</td>
<td>00:00</td>
<td>000168</td>
</tr>
<tr>
<td>20090725</td>
<td>02:00</td>
<td>01</td>
<td>03</td>
<td>00:00</td>
<td>00:00</td>
<td>000178</td>
</tr>
<tr>
<td>20090726</td>
<td>02:00</td>
<td>01</td>
<td>04</td>
<td>00:00</td>
<td>00:00</td>
<td>000188</td>
</tr>
<tr>
<td>20090727</td>
<td>02:00</td>
<td>01</td>
<td>05</td>
<td>00:00</td>
<td>00:00</td>
<td>000198</td>
</tr>
<tr>
<td>20090728</td>
<td>02:00</td>
<td>01</td>
<td>06</td>
<td>00:00</td>
<td>00:00</td>
<td>000208</td>
</tr>
<tr>
<td>20090729</td>
<td>02:00</td>
<td>01</td>
<td>07</td>
<td>00:00</td>
<td>00:00</td>
<td>000218</td>
</tr>
<tr>
<td>20090730</td>
<td>02:00</td>
<td>01</td>
<td>08</td>
<td>00:00</td>
<td>00:00</td>
<td>000228</td>
</tr>
<tr>
<td>20090731</td>
<td>02:00</td>
<td>01</td>
<td>09</td>
<td>00:00</td>
<td>00:00</td>
<td>000238</td>
</tr>
<tr>
<td>20090732</td>
<td>02:00</td>
<td>01</td>
<td>10</td>
<td>00:00</td>
<td>00:00</td>
<td>000248</td>
</tr>
<tr>
<td>20090733</td>
<td>02:00</td>
<td>01</td>
<td>11</td>
<td>00:00</td>
<td>00:00</td>
<td>000258</td>
</tr>
<tr>
<td>20090734</td>
<td>02:00</td>
<td>01</td>
<td>12</td>
<td>00:00</td>
<td>00:00</td>
<td>000268</td>
</tr>
<tr>
<td>20090735</td>
<td>02:00</td>
<td>01</td>
<td>13</td>
<td>00:00</td>
<td>00:00</td>
<td>000278</td>
</tr>
<tr>
<td>20090736</td>
<td>02:00</td>
<td>01</td>
<td>14</td>
<td>00:00</td>
<td>00:00</td>
<td>000288</td>
</tr>
<tr>
<td>20090737</td>
<td>02:00</td>
<td>01</td>
<td>15</td>
<td>00:00</td>
<td>00:00</td>
<td>000298</td>
</tr>
<tr>
<td>20090738</td>
<td>02:00</td>
<td>01</td>
<td>16</td>
<td>00:00</td>
<td>00:00</td>
<td>000308</td>
</tr>
<tr>
<td>20090739</td>
<td>02:00</td>
<td>01</td>
<td>17</td>
<td>00:00</td>
<td>00:00</td>
<td>000318</td>
</tr>
<tr>
<td>20090740</td>
<td>02:00</td>
<td>01</td>
<td>18</td>
<td>00:00</td>
<td>00:00</td>
<td>000328</td>
</tr>
<tr>
<td>20090741</td>
<td>02:00</td>
<td>01</td>
<td>19</td>
<td>00:00</td>
<td>00:00</td>
<td>000338</td>
</tr>
<tr>
<td>20090742</td>
<td>02:00</td>
<td>01</td>
<td>20</td>
<td>00:00</td>
<td>00:00</td>
<td>000348</td>
</tr>
<tr>
<td>20090743</td>
<td>02:00</td>
<td>01</td>
<td>21</td>
<td>00:00</td>
<td>00:00</td>
<td>000358</td>
</tr>
<tr>
<td>20090744</td>
<td>02:00</td>
<td>01</td>
<td>22</td>
<td>00:00</td>
<td>00:00</td>
<td>000368</td>
</tr>
<tr>
<td>20090745</td>
<td>02:00</td>
<td>01</td>
<td>23</td>
<td>00:00</td>
<td>00:00</td>
<td>000378</td>
</tr>
</tbody>
</table>

Continued on next page
Appendix B - Data

Continued on next page

Continued on next page
Appendix B - Data

Continued on next page
GM Data used in Chapter 5

Continued on next page
<table>
<thead>
<tr>
<th>Date</th>
<th>Time</th>
<th>Year</th>
<th>Model</th>
<th>Code</th>
<th>Age</th>
</tr>
</thead>
<tbody>
<tr>
<td>20100328</td>
<td>01:12:11</td>
<td>088</td>
<td>00:07:42</td>
<td>42212</td>
<td></td>
</tr>
<tr>
<td>20100329</td>
<td>00:46:09</td>
<td>088</td>
<td>00:08:44</td>
<td>42226</td>
<td></td>
</tr>
<tr>
<td>20100330</td>
<td>12:37:26</td>
<td>088</td>
<td>00:10:09</td>
<td>4247</td>
<td></td>
</tr>
<tr>
<td>20100330</td>
<td>23:43:38</td>
<td>088</td>
<td>00:14:16</td>
<td>42524</td>
<td></td>
</tr>
<tr>
<td>20100402</td>
<td>06:15:56</td>
<td>088</td>
<td>00:14:45</td>
<td>42823</td>
<td></td>
</tr>
<tr>
<td>20100402</td>
<td>14:21:14</td>
<td>088</td>
<td>00:15:34</td>
<td>42991</td>
<td></td>
</tr>
<tr>
<td>20100403</td>
<td>01:21:51</td>
<td>088</td>
<td>00:16:09</td>
<td>42984</td>
<td></td>
</tr>
<tr>
<td>20100403</td>
<td>13:04:36</td>
<td>089</td>
<td>00:16:73</td>
<td>42905</td>
<td></td>
</tr>
<tr>
<td>20100404</td>
<td>05:53:28</td>
<td>088</td>
<td>00:17:43</td>
<td>42312</td>
<td></td>
</tr>
<tr>
<td>20100405</td>
<td>14:26:59</td>
<td>088</td>
<td>00:19:46</td>
<td>42334</td>
<td></td>
</tr>
<tr>
<td>20100406</td>
<td>01:36:31</td>
<td>088</td>
<td>00:20:31</td>
<td>42341</td>
<td></td>
</tr>
<tr>
<td>20100407</td>
<td>00:59:04</td>
<td>088</td>
<td>00:21:17</td>
<td>42355</td>
<td></td>
</tr>
<tr>
<td>20100408</td>
<td>00:28:11</td>
<td>088</td>
<td>00:21:34</td>
<td>42369</td>
<td></td>
</tr>
<tr>
<td>20100408</td>
<td>23:54:44</td>
<td>088</td>
<td>00:24:35</td>
<td>42383</td>
<td></td>
</tr>
<tr>
<td>20100410</td>
<td>01:04:39</td>
<td>088</td>
<td>00:26:20</td>
<td>42398</td>
<td></td>
</tr>
<tr>
<td>20100410</td>
<td>06:33:38</td>
<td>088</td>
<td>00:27:44</td>
<td>42412</td>
<td></td>
</tr>
<tr>
<td>20100410</td>
<td>13:04:41</td>
<td>089</td>
<td>00:28:14</td>
<td>42419</td>
<td></td>
</tr>
<tr>
<td>20100412</td>
<td>14:07:16</td>
<td>089</td>
<td>00:29:56</td>
<td>42434</td>
<td></td>
</tr>
<tr>
<td>20100413</td>
<td>01:09:35</td>
<td>088</td>
<td>00:30:47</td>
<td>42441</td>
<td></td>
</tr>
<tr>
<td>20100414</td>
<td>13:02:10</td>
<td>088</td>
<td>00:32:46</td>
<td>42462</td>
<td></td>
</tr>
<tr>
<td>20100414</td>
<td>14:12:57</td>
<td>088</td>
<td>00:33:39</td>
<td>42477</td>
<td></td>
</tr>
<tr>
<td>20100415</td>
<td>23:41:08</td>
<td>088</td>
<td>00:34:43</td>
<td>42483</td>
<td></td>
</tr>
<tr>
<td>20100416</td>
<td>00:48:05</td>
<td>088</td>
<td>00:36:40</td>
<td>42498</td>
<td></td>
</tr>
<tr>
<td>20100418</td>
<td>12:38:04</td>
<td>089</td>
<td>00:38:11</td>
<td>42515</td>
<td></td>
</tr>
<tr>
<td>20100419</td>
<td>01:26:51</td>
<td>088</td>
<td>00:38:59</td>
<td>42527</td>
<td></td>
</tr>
<tr>
<td>20100420</td>
<td>05:03:38</td>
<td>088</td>
<td>00:40:31</td>
<td>42541</td>
<td></td>
</tr>
<tr>
<td>20100421</td>
<td>12:42:28</td>
<td>088</td>
<td>00:42:45</td>
<td>42562</td>
<td></td>
</tr>
<tr>
<td>20100421</td>
<td>13:21:67</td>
<td>088</td>
<td>00:43:41</td>
<td>42589</td>
<td></td>
</tr>
<tr>
<td>20100422</td>
<td>13:20:49</td>
<td>088</td>
<td>00:43:59</td>
<td>42577</td>
<td></td>
</tr>
<tr>
<td>20100423</td>
<td>03:56:16</td>
<td>088</td>
<td>00:44:46</td>
<td>42584</td>
<td></td>
</tr>
<tr>
<td>20100424</td>
<td>12:48:01</td>
<td>088</td>
<td>00:46:27</td>
<td>42605</td>
<td></td>
</tr>
<tr>
<td>20100425</td>
<td>01:29:31</td>
<td>088</td>
<td>00:47:25</td>
<td>42613</td>
<td></td>
</tr>
<tr>
<td>20100426</td>
<td>01:05:10</td>
<td>088</td>
<td>00:48:26</td>
<td>42627</td>
<td></td>
</tr>
<tr>
<td>20100427</td>
<td>23:53:08</td>
<td>089</td>
<td>00:16:05</td>
<td>42655</td>
<td></td>
</tr>
<tr>
<td>20100428</td>
<td>14:42:53</td>
<td>089</td>
<td>00:24:52</td>
<td>42663</td>
<td></td>
</tr>
<tr>
<td>20100430</td>
<td>03:06:36</td>
<td>089</td>
<td>00:45:42</td>
<td>42684</td>
<td></td>
</tr>
<tr>
<td>20100501</td>
<td>00:05:38</td>
<td>089</td>
<td>00:50:59</td>
<td>42698</td>
<td></td>
</tr>
<tr>
<td>20100501</td>
<td>14:10:06</td>
<td>089</td>
<td>00:56:70</td>
<td>42706</td>
<td></td>
</tr>
<tr>
<td>20100502</td>
<td>13:37:48</td>
<td>089</td>
<td>00:58:42</td>
<td>42720</td>
<td></td>
</tr>
<tr>
<td>20100503</td>
<td>13:04:59</td>
<td>089</td>
<td>00:59:25</td>
<td>42734</td>
<td></td>
</tr>
<tr>
<td>20100504</td>
<td>14:15:38</td>
<td>089</td>
<td>01:00:10</td>
<td>42749</td>
<td></td>
</tr>
<tr>
<td>20100504</td>
<td>23:43:31</td>
<td>089</td>
<td>01:10:56</td>
<td>42755</td>
<td></td>
</tr>
<tr>
<td>20100506</td>
<td>00:47:50</td>
<td>089</td>
<td>01:13:31</td>
<td>42770</td>
<td></td>
</tr>
<tr>
<td>20100507</td>
<td>02:03:01</td>
<td>089</td>
<td>01:14:52</td>
<td>42784</td>
<td></td>
</tr>
<tr>
<td>20100508</td>
<td>02:15:18</td>
<td>089</td>
<td>01:16:09</td>
<td>42799</td>
<td></td>
</tr>
<tr>
<td>20100508</td>
<td>13:56:22</td>
<td>089</td>
<td>01:17:46</td>
<td>42806</td>
<td></td>
</tr>
<tr>
<td>20100510</td>
<td>09:02:32</td>
<td>089</td>
<td>01:18:48</td>
<td>42827</td>
<td></td>
</tr>
<tr>
<td>20100510</td>
<td>14:28:19</td>
<td>089</td>
<td>01:19:75</td>
<td>42835</td>
<td></td>
</tr>
<tr>
<td>20100510</td>
<td>23:54:01</td>
<td>089</td>
<td>02:02:49</td>
<td>42841</td>
<td></td>
</tr>
</tbody>
</table>

GM Data used in Chapter 5

Continued on next page
GM Data used in Chapter 5

20100512 020902 089 00217 42856
20100513 020810 089 00231 42870
20100513 143359 090 00239 42878
20100514 122648 090 00252 42891
20100515 132956 090 00267 42906
20100517 092328 090 00289 42928
20100517 130272 090 00289 42928
20100520 140715 089 00296 42935
20100520 010939 090 00303 42942
20100519 040222 089 00317 42956
20100520 141908 090 00339 42978
20100521 011439 090 00346 42985
20100522 044803 089 00360 42999
20100523 123803 089 00381 43020
20100524 011928 090 00389 43028
20100525 050399 090 00403 43042
20100526 020316 090 00417 43056
20100526 235127 089 00431 43070
20100527 135910 089 00439 43078
20100528 131921 089 00453 43092
20100529 124848 090 00467 43106
20100529 235636 089 00474 43113
20100531 032069 089 00489 43128
20100601 033479 090 00502 43142
20100602 122953 090 00523 43163
20100603 010708 090 00531 43171
20100604 125918 090 00552 43192
20100605 141007 090 00567 43207
20100605 012090 090 00574 43214
20100607 045159 090 00588 43228
20100608 123417 090 00599 43249
20100608 243331 090 00616 43256
20100610 131039 090 00638 43278
20100611 123940 090 00652 43292
20100611 234840 090 00659 43299
20100612 135621 090 00670 43307
20100613 131629 090 00681 43321
20100614 124515 090 00695 43335
20100615 021080 090 00709 43343
20100616 010310 090 00727 43357
20100617 080579 090 00731 43371
20100617 235639 090 00745 43385
20100618 140130 090 00753 43393
20100620 030637 090 00774 43414
20100620 130491 090 00781 43421
20100621 020235 090 00788 43428
20100621 140714 090 00796 43436
20100622 010938 090 00803 43443
20100623 130209 090 00824 43464
20100624 141256 090 00839 43479

Continued on next page
### Appendix B - Data

GM Data used in Chapter 5

<table>
<thead>
<tr>
<th>Date</th>
<th>Time</th>
<th>Value 1</th>
<th>Value 2</th>
<th>Value 3</th>
<th>Value 4</th>
<th>Value 5</th>
</tr>
</thead>
<tbody>
<tr>
<td>20100624</td>
<td>23:36:26</td>
<td>090</td>
<td>00345</td>
<td>43485</td>
<td></td>
<td></td>
</tr>
<tr>
<td>20100625</td>
<td>13:40:37</td>
<td>090</td>
<td>00357</td>
<td>43493</td>
<td></td>
<td></td>
</tr>
<tr>
<td>20100626</td>
<td>13:07:56</td>
<td>090</td>
<td>00367</td>
<td>43507</td>
<td></td>
<td></td>
</tr>
<tr>
<td>20100627</td>
<td>14:18:22</td>
<td>090</td>
<td>00382</td>
<td>43522</td>
<td></td>
<td></td>
</tr>
<tr>
<td>20100628</td>
<td>01:19:27</td>
<td>090</td>
<td>00389</td>
<td>43529</td>
<td></td>
<td></td>
</tr>
<tr>
<td>20100628</td>
<td>13:53:31</td>
<td>090</td>
<td>00396</td>
<td>43536</td>
<td></td>
<td></td>
</tr>
<tr>
<td>20100630</td>
<td>00:19:42</td>
<td>090</td>
<td>00417</td>
<td>43557</td>
<td></td>
<td></td>
</tr>
<tr>
<td>20100630</td>
<td>14:25:53</td>
<td>090</td>
<td>00425</td>
<td>43565</td>
<td></td>
<td></td>
</tr>
<tr>
<td>20100701</td>
<td>01:26:52</td>
<td>090</td>
<td>00432</td>
<td>43572</td>
<td></td>
<td></td>
</tr>
<tr>
<td>20100702</td>
<td>00:56:15</td>
<td>090</td>
<td>00446</td>
<td>43586</td>
<td></td>
<td></td>
</tr>
<tr>
<td>20100703</td>
<td>00:28:07</td>
<td>090</td>
<td>00460</td>
<td>43600</td>
<td></td>
<td></td>
</tr>
<tr>
<td>20100703</td>
<td>23:53:46</td>
<td>090</td>
<td>00474</td>
<td>43614</td>
<td></td>
<td></td>
</tr>
<tr>
<td>20100704</td>
<td>12:23:57</td>
<td>090</td>
<td>00481</td>
<td>43621</td>
<td></td>
<td></td>
</tr>
<tr>
<td>20100705</td>
<td>01:05:59</td>
<td>090</td>
<td>00489</td>
<td>43629</td>
<td></td>
<td></td>
</tr>
<tr>
<td>20100706</td>
<td>23:59:38</td>
<td>091</td>
<td>00016</td>
<td>43657</td>
<td></td>
<td></td>
</tr>
<tr>
<td>20100707</td>
<td>14:04:23</td>
<td>091</td>
<td>00024</td>
<td>43665</td>
<td></td>
<td></td>
</tr>
<tr>
<td>20100709</td>
<td>00:36:36</td>
<td>091</td>
<td>00045</td>
<td>43686</td>
<td></td>
<td></td>
</tr>
<tr>
<td>20100710</td>
<td>00:05:33</td>
<td>091</td>
<td>00059</td>
<td>43700</td>
<td></td>
<td></td>
</tr>
<tr>
<td>20100710</td>
<td>14:10:06</td>
<td>091</td>
<td>00067</td>
<td>43708</td>
<td></td>
<td></td>
</tr>
<tr>
<td>20100711</td>
<td>01:12:08</td>
<td>091</td>
<td>00074</td>
<td>43715</td>
<td></td>
<td></td>
</tr>
<tr>
<td>20100712</td>
<td>00:45:19</td>
<td>091</td>
<td>00088</td>
<td>43729</td>
<td></td>
<td></td>
</tr>
<tr>
<td>20100713</td>
<td>14:15:38</td>
<td>091</td>
<td>00110</td>
<td>43751</td>
<td></td>
<td></td>
</tr>
<tr>
<td>20100714</td>
<td>01:17:05</td>
<td>091</td>
<td>00117</td>
<td>43758</td>
<td></td>
<td></td>
</tr>
<tr>
<td>20100715</td>
<td>13:10:09</td>
<td>091</td>
<td>00138</td>
<td>43779</td>
<td></td>
<td></td>
</tr>
<tr>
<td>20100716</td>
<td>12:40:54</td>
<td>091</td>
<td>00152</td>
<td>43793</td>
<td></td>
<td></td>
</tr>
<tr>
<td>20100717</td>
<td>01:30:04</td>
<td>091</td>
<td>00160</td>
<td>43801</td>
<td></td>
<td></td>
</tr>
<tr>
<td>20100719</td>
<td>00:18:32</td>
<td>091</td>
<td>00183</td>
<td>43829</td>
<td></td>
<td></td>
</tr>
<tr>
<td>20100719</td>
<td>14:26:57</td>
<td>091</td>
<td>00196</td>
<td>43837</td>
<td></td>
<td></td>
</tr>
<tr>
<td>20100720</td>
<td>12:21:09</td>
<td>091</td>
<td>00209</td>
<td>43850</td>
<td></td>
<td></td>
</tr>
<tr>
<td>20100720</td>
<td>00:59:03</td>
<td>091</td>
<td>00217</td>
<td>43858</td>
<td></td>
<td></td>
</tr>
<tr>
<td>20100722</td>
<td>00:28:10</td>
<td>091</td>
<td>00231</td>
<td>43872</td>
<td></td>
<td></td>
</tr>
<tr>
<td>20100722</td>
<td>23:38:34</td>
<td>091</td>
<td>00245</td>
<td>43886</td>
<td></td>
<td></td>
</tr>
<tr>
<td>20100723</td>
<td>14:01:31</td>
<td>091</td>
<td>00254</td>
<td>43894</td>
<td></td>
<td></td>
</tr>
<tr>
<td>20100724</td>
<td>13:27:58</td>
<td>091</td>
<td>00267</td>
<td>43908</td>
<td></td>
<td></td>
</tr>
<tr>
<td>20100725</td>
<td>13:04:41</td>
<td>091</td>
<td>00281</td>
<td>43922</td>
<td></td>
<td></td>
</tr>
<tr>
<td>20100726</td>
<td>12:26:52</td>
<td>091</td>
<td>00295</td>
<td>43936</td>
<td></td>
<td></td>
</tr>
<tr>
<td>20100727</td>
<td>23:35:42</td>
<td>091</td>
<td>00302</td>
<td>43943</td>
<td></td>
<td></td>
</tr>
<tr>
<td>20100728</td>
<td>00:92:62</td>
<td>091</td>
<td>00317</td>
<td>43958</td>
<td></td>
<td></td>
</tr>
<tr>
<td>20100728</td>
<td>00:08:24</td>
<td>091</td>
<td>00331</td>
<td>43972</td>
<td></td>
<td></td>
</tr>
<tr>
<td>20100729</td>
<td>14:19:12</td>
<td>091</td>
<td>00339</td>
<td>43980</td>
<td></td>
<td></td>
</tr>
<tr>
<td>20100730</td>
<td>13:40:39</td>
<td>091</td>
<td>00353</td>
<td>43994</td>
<td></td>
<td></td>
</tr>
<tr>
<td>20100801</td>
<td>00:14:02</td>
<td>091</td>
<td>00374</td>
<td>44015</td>
<td></td>
<td></td>
</tr>
<tr>
<td>20100801</td>
<td>14:25:06</td>
<td>091</td>
<td>00382</td>
<td>44023</td>
<td></td>
<td></td>
</tr>
<tr>
<td>20100802</td>
<td>01:26:52</td>
<td>091</td>
<td>00389</td>
<td>44030</td>
<td></td>
<td></td>
</tr>
<tr>
<td>20100803</td>
<td>13:13:40</td>
<td>091</td>
<td>00410</td>
<td>44051</td>
<td></td>
<td></td>
</tr>
<tr>
<td>20100804</td>
<td>12:43:20</td>
<td>091</td>
<td>00424</td>
<td>44065</td>
<td></td>
<td></td>
</tr>
<tr>
<td>20100804</td>
<td>23:51:29</td>
<td>091</td>
<td>00431</td>
<td>44072</td>
<td></td>
<td></td>
</tr>
<tr>
<td>20100806</td>
<td>00:56:18</td>
<td>091</td>
<td>00446</td>
<td>44087</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Date</td>
<td>Time</td>
<td>Code</td>
<td>GM Value</td>
<td>Data Value</td>
<td></td>
<td></td>
</tr>
<tr>
<td>-------------</td>
<td>--------</td>
<td>------</td>
<td>----------</td>
<td>------------</td>
<td></td>
<td></td>
</tr>
<tr>
<td>2010-08-07</td>
<td>00:28:09</td>
<td>09:1</td>
<td>441001</td>
<td>910467 441108</td>
<td></td>
<td></td>
</tr>
<tr>
<td>2010-08-07</td>
<td>23:53:40</td>
<td>09:1</td>
<td>441115</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>2010-08-08</td>
<td>12:44:09</td>
<td>09:1</td>
<td>441222</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>2010-08-09</td>
<td>13:24:00</td>
<td>09:1</td>
<td>441337</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>2010-08-08</td>
<td>01:37:49</td>
<td>09:2</td>
<td>441459</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>2010-08-08</td>
<td>23:23:52</td>
<td>09:2</td>
<td>441472</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>2010-08-07</td>
<td>01:06:39</td>
<td>09:2</td>
<td>441487</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>2010-08-08</td>
<td>12:29:01</td>
<td>09:2</td>
<td>441520</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>2010-08-07</td>
<td>23:30:38</td>
<td>09:2</td>
<td>441521</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>2010-08-06</td>
<td>04:21:09</td>
<td>09:2</td>
<td>441620</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>2010-08-07</td>
<td>00:11:14</td>
<td>09:2</td>
<td>441644</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>2010-08-17</td>
<td>14:22:06</td>
<td>09:2</td>
<td>441652</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>2010-08-08</td>
<td>13:43:35</td>
<td>09:2</td>
<td>442066</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>2010-08-08</td>
<td>02:02:32</td>
<td>09:2</td>
<td>442087</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>2010-08-08</td>
<td>23:44:42</td>
<td>09:2</td>
<td>443001</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>2010-08-08</td>
<td>13:49:29</td>
<td>09:2</td>
<td>443090</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>2010-08-08</td>
<td>13:16:32</td>
<td>09:2</td>
<td>443232</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>2010-08-23</td>
<td>14:46:58</td>
<td>09:2</td>
<td>443338</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>2010-08-08</td>
<td>01:36:01</td>
<td>09:2</td>
<td>443445</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>2010-08-25</td>
<td>01:00:31</td>
<td>09:2</td>
<td>443559</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>2010-08-26</td>
<td>12:50:52</td>
<td>09:2</td>
<td>443800</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>2010-08-08</td>
<td>01:55:17</td>
<td>09:2</td>
<td>443817</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>2010-08-08</td>
<td>23:54:02</td>
<td>09:2</td>
<td>444098</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>2010-08-08</td>
<td>03:00:49</td>
<td>09:2</td>
<td>444090</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>2010-08-08</td>
<td>03:03:48</td>
<td>09:2</td>
<td>444116</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>2010-08-08</td>
<td>02:02:37</td>
<td>09:2</td>
<td>444303</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>2010-08-08</td>
<td>14:07:16</td>
<td>09:2</td>
<td>444348</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>2010-08-08</td>
<td>01:09:39</td>
<td>09:2</td>
<td>444445</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>2010-08-08</td>
<td>13:02:10</td>
<td>09:2</td>
<td>444666</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>2010-08-08</td>
<td>14:19:12</td>
<td>09:2</td>
<td>444841</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>2010-08-09</td>
<td>01:43:49</td>
<td>09:2</td>
<td>444848</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>2010-08-09</td>
<td>13:07:50</td>
<td>09:2</td>
<td>445095</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>2010-08-09</td>
<td>14:18:23</td>
<td>09:2</td>
<td>445244</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>2010-08-05</td>
<td>23:46:02</td>
<td>09:2</td>
<td>445350</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>2010-08-06</td>
<td>13:46:34</td>
<td>09:2</td>
<td>445358</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>2010-08-08</td>
<td>12:42:29</td>
<td>09:2</td>
<td>445566</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>2010-08-09</td>
<td>23:15:28</td>
<td>09:2</td>
<td>445753</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>2010-08-05</td>
<td>13:25:00</td>
<td>09:2</td>
<td>445851</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>2010-08-09</td>
<td>13:59:14</td>
<td>09:2</td>
<td>445851</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>2010-08-09</td>
<td>00:25:21</td>
<td>09:2</td>
<td>446062</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>2010-08-09</td>
<td>23:53:47</td>
<td>09:2</td>
<td>446116</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>2010-08-11</td>
<td>12:25:38</td>
<td>09:2</td>
<td>446232</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>2010-08-13</td>
<td>01:05:15</td>
<td>09:2</td>
<td>446361</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>2010-08-13</td>
<td>13:25:03</td>
<td>09:2</td>
<td>446368</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>2010-08-14</td>
<td>13:56:38</td>
<td>09:3</td>
<td>446562</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>2010-08-15</td>
<td>12:29:53</td>
<td>09:3</td>
<td>446666</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>2010-08-16</td>
<td>01:07:08</td>
<td>09:3</td>
<td>446674</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>2010-08-17</td>
<td>12:59:17</td>
<td>09:3</td>
<td>446950</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>2010-08-18</td>
<td>12:55:52</td>
<td>09:3</td>
<td>447097</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>2010-08-18</td>
<td>14:00:16</td>
<td>09:3</td>
<td>447098</td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

Continued on next page
## GM Data used in Chapter 5

<table>
<thead>
<tr>
<th>Date</th>
<th>Time</th>
<th>GMID</th>
<th>Event ID</th>
<th>Lat</th>
<th>Lon</th>
</tr>
</thead>
<tbody>
<tr>
<td>2010-09-18</td>
<td>23:38:28</td>
<td>093</td>
<td>00073</td>
<td>44716</td>
<td></td>
</tr>
<tr>
<td>2010-09-19</td>
<td>01:12:08</td>
<td>093</td>
<td>00074</td>
<td>44717</td>
<td></td>
</tr>
<tr>
<td>2010-09-20</td>
<td>00:45:14</td>
<td>093</td>
<td>00088</td>
<td>44731</td>
<td></td>
</tr>
<tr>
<td>2010-09-21</td>
<td>14:15:37</td>
<td>093</td>
<td>00110</td>
<td>44753</td>
<td></td>
</tr>
<tr>
<td>2010-09-22</td>
<td>23:43:30</td>
<td>093</td>
<td>00145</td>
<td>44759</td>
<td></td>
</tr>
<tr>
<td>2010-09-23</td>
<td>00:04:47</td>
<td>093</td>
<td>00131</td>
<td>44774</td>
<td></td>
</tr>
<tr>
<td>2010-09-24</td>
<td>00:16:52</td>
<td>093</td>
<td>00145</td>
<td>44788</td>
<td></td>
</tr>
<tr>
<td>2010-09-25</td>
<td>14:21:10</td>
<td>093</td>
<td>00153</td>
<td>44796</td>
<td></td>
</tr>
<tr>
<td>2010-09-26</td>
<td>05:52:25</td>
<td>093</td>
<td>00174</td>
<td>44810</td>
<td></td>
</tr>
<tr>
<td>2010-09-27</td>
<td>02:22:29</td>
<td>093</td>
<td>00188</td>
<td>44831</td>
<td></td>
</tr>
<tr>
<td>2010-09-28</td>
<td>14:26:55</td>
<td>093</td>
<td>00196</td>
<td>44839</td>
<td></td>
</tr>
<tr>
<td>2010-09-29</td>
<td>01:36:07</td>
<td>093</td>
<td>00203</td>
<td>44846</td>
<td></td>
</tr>
<tr>
<td>2010-09-30</td>
<td>13:55:40</td>
<td>093</td>
<td>00210</td>
<td>44853</td>
<td></td>
</tr>
<tr>
<td>2010-10-01</td>
<td>12:21:08</td>
<td>093</td>
<td>00209</td>
<td>44852</td>
<td></td>
</tr>
<tr>
<td>2010-10-02</td>
<td>01:03:08</td>
<td>093</td>
<td>00217</td>
<td>44860</td>
<td></td>
</tr>
<tr>
<td>2010-10-03</td>
<td>12:50:48</td>
<td>093</td>
<td>00238</td>
<td>44881</td>
<td></td>
</tr>
<tr>
<td>2010-10-04</td>
<td>01:32:05</td>
<td>093</td>
<td>00246</td>
<td>44889</td>
<td></td>
</tr>
<tr>
<td>2010-10-05</td>
<td>01:04:35</td>
<td>093</td>
<td>00260</td>
<td>44903</td>
<td></td>
</tr>
<tr>
<td>2010-10-06</td>
<td>03:03:43</td>
<td>093</td>
<td>00274</td>
<td>44917</td>
<td></td>
</tr>
<tr>
<td>2010-10-07</td>
<td>02:26:29</td>
<td>093</td>
<td>00283</td>
<td>44931</td>
<td></td>
</tr>
<tr>
<td>2010-10-08</td>
<td>14:07:11</td>
<td>093</td>
<td>00295</td>
<td>44938</td>
<td></td>
</tr>
<tr>
<td>2010-10-09</td>
<td>01:37:19</td>
<td>093</td>
<td>00289</td>
<td>44932</td>
<td></td>
</tr>
<tr>
<td>2010-10-10</td>
<td>14:04:21</td>
<td>093</td>
<td>00317</td>
<td>44960</td>
<td></td>
</tr>
<tr>
<td>2010-10-11</td>
<td>01:23:44</td>
<td>093</td>
<td>00328</td>
<td>44974</td>
<td></td>
</tr>
<tr>
<td>2010-10-12</td>
<td>01:14:09</td>
<td>093</td>
<td>00348</td>
<td>44982</td>
<td></td>
</tr>
<tr>
<td>2010-10-13</td>
<td>03:04:34</td>
<td>093</td>
<td>00360</td>
<td>45003</td>
<td></td>
</tr>
</tbody>
</table>

223
References


REFERENCES


REFERENCES


226
REFERENCES


ESA (2004). Absolute calibration of ASAR level 1 products generated with PF-ASAR. European Space Agency, ESRIN, Via Galileo Galilei, Casella Postale 64, 00044 Frascati (Rome), Italy, 1st edn. 14, 40, 42, 43, 88, 113


227
REFERENCES


ESA (2010). Enviview. In Earthnet Online, European Space Agency, Earth Observation Ground Segment Department User Services and Mission Planning Office Via Galileo Galilei 00044 Frascati (Rome) Italy. 79


REFERENCES


REFERENCES


REFERENCES


232
REFERENCES


REFERENCES


REFERENCES


REFERENCES


