UNIVERSITY OF SOUTHAMPTON

An Investigation into Ocean Wave Sources of Ambient Seismic Noise.

by

Jennifer F. Ward Neale

A thesis submitted in partial fulfillment for the degree of Doctor of Philosophy

in the
Faculty of Natural and Environmental Sciences
Ocean and Earth Science

May 2017
Periodic pressure fluctuations beneath ocean waves deform the seabed beneath them and can generate seismic waves that propagate around the globe. These pressure fluctuations are recorded on seafloor pressure gauges and the associated ground displacements on seismometers, where they contribute to ambient seismic noise. The signals offer an opportunity to study or monitor ocean waves that are traditionally difficult to measure because of their low wave heights (deep water infragravity waves) or their remote locations (deep water swell waves). However, the link between ocean waves and the oscillations of the pressure and seismic wavefields has remained unclear.

The aim of this thesis was to increase our understanding of ocean sources of ambient seismic noise, including their location, how well they can be located, and their relationship with ocean wave parameters.

Using cross-correlated pairs of seafloor pressure records, infragravity waves offshore California/Oregon were found to originate mostly from local coastlines during northern-hemisphere winter and from the south during summer. A first attempt to estimate the coastal reflection coefficient of remote arrivals was made and found to be 0.49-0.74, which has implications for infragravity energy in the deep ocean and may be important for models in which infragravity waves are propagated across ocean basins.

P-wave sources in the North Pacific during winter were located using seismometers in California and found to be concentrated around 40-50°N 160-180°E. Observed source locations were within 10° of the modelled source locations. Significant wave height in the deep ocean was estimated from P-waves and correlated with modelled wave height with a correlation coefficient of 0.63. Previous work only attempted to estimate coastal wave heights. Combining additional records from Japan and Europe improved source location, including imaging of multiple sources. Accuracy in source location and amplitude estimation are essential if microseisms are to be used to monitor wave activity in the deep ocean.
Contents

List of Figures ix
List of Tables xiii
Declaration of Authorship xv
Acknowledgements xvii

1 Introduction 1
  1.1 Background . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . 1
    1.1.1 Historical context and motivations for studying ambient noise . . 3
    1.1.2 Double frequency P-wave microseisms . . . . . . . . . . . . . . . . 5
    1.1.3 Infragravity waves and seismic hum . . . . . . . . . . . . . . . . . 7
    1.1.4 Contemporary motivations for study . . . . . . . . . . . . . . . . . 8
  1.2 Data and methods . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . 12
    1.2.1 Ocean bottom pressure . . . . . . . . . . . . . . . . . . . . . . . 12
      1.2.1.1 Cross-correlation . . . . . . . . . . . . . . . . . . . . . . 12
      1.2.1.2 Backprojection . . . . . . . . . . . . . . . . . . . . . . . 14
      1.2.1.3 Incoherent beamforming . . . . . . . . . . . . . . . . . . 20
    1.2.2 Wave buoys . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . 20
    1.2.3 Seismic records . . . . . . . . . . . . . . . . . . . . . . . . . . . . 21
      1.2.3.1 Beamforming . . . . . . . . . . . . . . . . . . . . . . . . 22
    1.2.4 WAVEWATCH III® . . . . . . . . . . . . . . . . . . . . . . . . . . 26
      1.2.4.1 Description of model . . . . . . . . . . . . . . . . . . . . 26
      1.2.4.2 Model setup . . . . . . . . . . . . . . . . . . . . . . . . . 29
      1.2.4.3 Modelling of P-wave microseisms . . . . . . . . . . . . . 30

2 Source regions and reflection of infragravity waves offshore of the USA’s Pacific Northwest. 33
  2.1 Abstract . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . 33
  2.2 Introduction . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . 34
  2.3 Data and methods . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . 36
    2.3.1 Cross-correlation . . . . . . . . . . . . . . . . . . . . . . . . . . . 37
    2.3.2 Backprojection . . . . . . . . . . . . . . . . . . . . . . . . . . . . 38
      2.3.2.1 Backprojection for isotropic source distribution . . . . . . 40
      2.3.2.2 Backprojection weighting . . . . . . . . . . . . . . . . . . 41
    2.3.3 Beamforming . . . . . . . . . . . . . . . . . . . . . . . . . . . . . 42
    2.3.4 Effects of normalisation on directionality estimates . . . . . . . . 43
<table>
<thead>
<tr>
<th>Section</th>
<th>Title</th>
<th>Page</th>
</tr>
</thead>
<tbody>
<tr>
<td>2.3.5</td>
<td>Calculating reflection coefficients</td>
<td>44</td>
</tr>
<tr>
<td>2.3.5.1</td>
<td>Seaward:shoreward wave propagation</td>
<td>44</td>
</tr>
<tr>
<td>2.3.5.2</td>
<td>Assumption of no local infragravity wave generation</td>
<td>44</td>
</tr>
<tr>
<td>2.3.6</td>
<td>Cases of reflection</td>
<td>46</td>
</tr>
<tr>
<td>2.4</td>
<td>Results</td>
<td>46</td>
</tr>
<tr>
<td>2.4.1</td>
<td>Infragravity wave energy and directionality</td>
<td>46</td>
</tr>
<tr>
<td>2.4.2</td>
<td>Reflection of infragravity waves</td>
<td>47</td>
</tr>
<tr>
<td>2.5</td>
<td>Discussion</td>
<td>54</td>
</tr>
<tr>
<td>2.6</td>
<td>Conclusions</td>
<td>56</td>
</tr>
<tr>
<td>3</td>
<td>Monitoring remote ocean waves using P-wave microseisms</td>
<td>59</td>
</tr>
<tr>
<td>3.1</td>
<td>Abstract</td>
<td>59</td>
</tr>
<tr>
<td>3.2</td>
<td>Introduction</td>
<td>60</td>
</tr>
<tr>
<td>3.3</td>
<td>Data and processing</td>
<td>61</td>
</tr>
<tr>
<td>3.4</td>
<td>Results</td>
<td>66</td>
</tr>
<tr>
<td>3.4.1</td>
<td>Comparison between observed and modelled source</td>
<td>66</td>
</tr>
<tr>
<td>3.4.2</td>
<td>Relating microseism source to significant wave height</td>
<td>70</td>
</tr>
<tr>
<td>3.5</td>
<td>Discussion</td>
<td>75</td>
</tr>
<tr>
<td>3.6</td>
<td>Conclusions</td>
<td>79</td>
</tr>
<tr>
<td>4</td>
<td>Improving microseismic P-wave source location with multiple seismic</td>
<td>81</td>
</tr>
<tr>
<td>4.1</td>
<td>Abstract</td>
<td>81</td>
</tr>
<tr>
<td>4.2</td>
<td>Introduction</td>
<td>81</td>
</tr>
<tr>
<td>4.3</td>
<td>Data and methods</td>
<td>83</td>
</tr>
<tr>
<td>4.3.1</td>
<td>Numerical modelling of P-wave sources</td>
<td>83</td>
</tr>
<tr>
<td>4.3.2</td>
<td>Seismic data and processing</td>
<td>84</td>
</tr>
<tr>
<td>4.4</td>
<td>Synthetic tests</td>
<td>88</td>
</tr>
<tr>
<td>4.5</td>
<td>Results</td>
<td>91</td>
</tr>
<tr>
<td>4.5.1</td>
<td>Individual arrays</td>
<td>91</td>
</tr>
<tr>
<td>4.5.2</td>
<td>Combined arrays</td>
<td>92</td>
</tr>
<tr>
<td>4.5.3</td>
<td>Weighting by distance</td>
<td>103</td>
</tr>
<tr>
<td>4.6</td>
<td>Discussion</td>
<td>104</td>
</tr>
<tr>
<td>4.7</td>
<td>Conclusions</td>
<td>106</td>
</tr>
<tr>
<td>5</td>
<td>Conclusions and Future Work</td>
<td>109</td>
</tr>
<tr>
<td>5.1</td>
<td>Conclusions</td>
<td>109</td>
</tr>
<tr>
<td>5.2</td>
<td>Future Work</td>
<td>111</td>
</tr>
<tr>
<td>Appendix A</td>
<td>Input files for Fast Marching Surface Tomography Package</td>
<td>115</td>
</tr>
<tr>
<td>A.1</td>
<td>fm2dss.in</td>
<td>116</td>
</tr>
<tr>
<td>A.2</td>
<td>Switch file for WAVEWATCH III</td>
<td>118</td>
</tr>
<tr>
<td>A.2.1</td>
<td>Switch file</td>
<td>118</td>
</tr>
<tr>
<td>A.2.2</td>
<td>Description of switch file</td>
<td>119</td>
</tr>
<tr>
<td>A.3</td>
<td>Input files for WAVEWATCH III</td>
<td>121</td>
</tr>
</tbody>
</table>
### CONTENTS

<table>
<thead>
<tr>
<th>Section</th>
<th>File</th>
<th>Description</th>
<th>Page</th>
</tr>
</thead>
<tbody>
<tr>
<td>A.3.1</td>
<td>ww3_grid.inp.glob_30m</td>
<td>For grid preprocessor (Global 0.5 degree grid) [with comments]</td>
<td>122</td>
</tr>
<tr>
<td>A.3.2</td>
<td>ww3_grid.inp.casc_6m</td>
<td>For grid preprocessor (Cascadia 0.1 degree grid)</td>
<td>130</td>
</tr>
<tr>
<td>A.3.3</td>
<td>ww3_grid.inp.points</td>
<td>For grid preprocessor (Point output)</td>
<td>131</td>
</tr>
<tr>
<td>A.3.4</td>
<td>ww3strt.inp</td>
<td>For initial conditions program [with comments]</td>
<td>132</td>
</tr>
<tr>
<td>A.3.5</td>
<td>ww3_prnc.inp.wind</td>
<td>For NetCDF input field preprocessor (Wind) [with comments]</td>
<td>133</td>
</tr>
<tr>
<td>A.3.6</td>
<td>ww3_prnc.inp.ice</td>
<td>For NetCDF input field preprocessor (Ice)</td>
<td>134</td>
</tr>
<tr>
<td>A.3.7</td>
<td>ww3_multi.inp</td>
<td>For the multi-grid shell [with comments]</td>
<td>135</td>
</tr>
<tr>
<td>A.3.8</td>
<td>ww3_ounf_glob_30m.inp</td>
<td>For the gridded NetCDF output post-processor (Global 0.5 degree grid) [with comments]</td>
<td>143</td>
</tr>
<tr>
<td>A.3.9</td>
<td>ww3_ounf_casc_6m.inp</td>
<td>For the gridded NetCDF output post-processor (Cascadia 0.1 degree grid)</td>
<td>144</td>
</tr>
<tr>
<td>A.3.10</td>
<td>ww3_ounp.inp</td>
<td>For the point output NetCDF post-processor [with comments]</td>
<td>145</td>
</tr>
</tbody>
</table>

**Appendix B**

- B.1 Additional figures and tables for Chapter 2 | 148
- B.2 Published paper and Supporting Information of Chapter 2 | 157

**Appendix C**

- C.1 Additional figures for Chapter 3 | 190
- C.2 Published paper and Supporting Information of Chapter 3 | 192

**List of References**

<table>
<thead>
<tr>
<th>Page</th>
</tr>
</thead>
<tbody>
<tr>
<td>209</td>
</tr>
</tbody>
</table>
# List of Figures

1.1 Maps showing locations of DPG stations and wave buoys. 13
1.2 Average power spectra for DPG station G02B, September 2012-May 2013. 14
1.3 Stacked cross-correlation functions, band-pass-filtered between 60 and 500 s. 15
1.4 Examples of stacked cross-correlation functions for east-west aligned station pairs filtered to 100 s (top row), 150 s (middle row) and 200 s (bottom row). 16
1.5 ETOP01 (Amante and Eakins, 2009) bathymetry. 17
1.6 Group velocity at a) 100 s b) 150 s and c) 200 s. 18
1.7 Raypaths from station G21B (in green) to the locations of each other station for a 150 s wave. 18
1.8 Group travel-time at a) 100 s b) 150 s and c) 200 s for station G21B. 19
1.9 Lag-time between stations J11B (yellow) and G21B (green) at a) 100 s b) 150 s and c) 200 s. 19
1.10 Location of the a) Californian, b) Japanese and c) European arrays along with array configurations. 22
1.11 Velocity seismogram at station STC of the Southern California Seismic Network on 15 February 2013. 23
1.12 Average power spectra for station STC of the Southern California Seismic Network, September 2012-May 2013. 23
1.13 Average power spectra between 0.1 and 0.3 Hz calculated over 35 days during the winter of 2012/13. a) Station STC of the Southern California Seismic Network b) Station OTAH of the Japan Hi-net network c) Station ATTE of the Italian National Seismic Network. 24
1.14 Beamformer output on 2 December 2012 03:00-09:00 at slowness of 0 to 0.1 s/km. 25
1.15 Slowness and ray-paths for body wave phases. 25
1.16 Beampower output on 2 December 2012 03:00-09:00 (Figure 1.14) projected onto a North Pacific grid. 26
2.1 Backprojection of September 2012-May 2013 stack at 150 s. 40
2.2 Beamforming of September 2012-May 2013 stack. 43
2.3 Schematic of propagation and reflection of a simple impulse. 45
2.4 Direction of maximum beam power, September 2012-May 2013. 47
2.5 Infragravity power vs. significant wave height and $\alpha H_s T_{m0}^2 - 2\sqrt{\frac{g}{D}}$. 48
2.6 Same as Figure 2.5d and e but $H_s$ and $\alpha H_s T_{m0}^2 - 2\sqrt{\frac{g}{D}}$ vs. daily seaward:shoreward energy. 49
2.7 $H_s$ and $\alpha H_s T_n^{2} m_{0,2} - 2\sqrt{\frac{g}{D}}$ during January 2013, and backprojection on 18 January 2013. .................................................. 50
2.8 Same as Figure 1.4 but for 18 January 2013. ............................ 51
2.9 Values of $r (\sqrt{IG \text{ seaward:shoreward}})$ obtained from east-west station pairs on (a) 18 January 2013 (b) 25 October 2012 (c) 4 April 2013 (d) 17 January 2013 (e) 11 May 2013 as well as the mean $r$ on each day (f). .... 52
2.10 Same as Figure 2.9f but only using selected clear cross-correlations. .... 53

3.1 The seismic networks used in the study. ................................. 62
3.2 Top row: Slowness vs. azimuth array response functions at frequencies of a) 0.1 and b) 0.3 Hz. White circles are plotted at slowness values of 0.02, 0.04, 0.06 and 0.08 s/km. Bottom row: Array response for a point source located at 35°N 169°E for frequencies of c) 0.1 and d) 0.3 Hz. .... 63
3.3 Example output for 23 December 2012 00:00-03:00. .................... 64
3.4 Modelled source power vs. observed beampower (both integrated between 0.1 and 0.3 Hz). ......................................................... 67
3.5 Locations of beamformer peaks and associated modelled sources. .... 68
3.6 Mean and peak frequencies of the beamformer output and modelled source. 69
3.7 Modelled integrated source power, observed integrated beampower (integrated) and estimated source power. ............................... 71
3.8 Modelled source power, observed beampower, mean shape of modelled source spectrum and mean spectrum of wave-interaction intensity. .... 72
3.9 Modelled $H_s(swell)$ during the first winter (mid-October 2012 to mid-March 2013) vs. $H_s(swell)$ estimated from beamformer output. .... 74
3.10 Modelled and estimated $H_s(swell)$. ........................................ 75
3.11 Causes of error in the estimation of $H_s(swell)$. ......................... 76
3.12 An example when maximum source corresponds to location of maximum $H_s$ (top row) and when maximum source does not correspond to location of maximum $H_s$ (bottom row). ................................. 79

4.1 $C_P$ at 0.207 Hz. ................................................................. 85
4.2 Weights for distance. .......................................................... 87
4.3 Testing how well the California array identifies two neighbouring point sources. ................................................................. 89
4.4 Same as Figure 4.3 but for the Japan array. ................................. 89
4.5 Same as Figures 4.3 and 4.4 but for the Europe array. .................... 90
4.6 Testing how well the combined arrays identify two neighbouring point sources. ................................................................. 91
4.7 Modelled and observed sources on 13 January 2013 18:00-21:00 at 0.188 Hz. ................................................................. 94
4.8 Same as Figure 4.7 but for sources on 15 February 2013 06:00-09:00 at 0.188 Hz. ................................................................. 95
4.9 Same as Figure 4.7 but for sources on 10 October 2012 06:00-09:00 at 0.207 Hz. ................................................................. 96
4.10 Same as Figure 4.7 but for sources on 27 October 2012 06:00-09:00 at 0.207 Hz. ................................................................. 97
4.11 Same as Figure 4.7 but for sources on 19 November 2012 03:00-06:00 at 0.188 Hz. ................................................................. 98
4.12 Same as Figure 4.7 but for sources on 19 November 2012 12:00-15:00 at 0.188 Hz. ............................................. 99
4.13 Same as Figure 4.7 but for sources on 19 January 2013 00:00-03:00 at 0.207 Hz. ............................................. 100
4.14 Same as Figure 4.7 but for sources on 02 October 2012 21:00-24:00 at 0.207 Hz. ............................................. 101
4.15 Same as Figure 4.7 but for sources on 05 October 2012 21:00-24:00 at 0.188 Hz. ............................................. 102
B.1 Station map with lines connecting station pairs that have separation dis-
tances greater than 30 km and less than 120 km. ....................... 148
B.2 Weighted backprojection of September 2012-May 2013 stack. ........ 149
B.3 Beamforming of September 2012-May 2013 stack. ....................... 150
B.4 Backprojection for the days with westerly arrivals. ....................... 151
B.5 Examples of unclear cross-correlations. ................................ 152
B.6 Same as Figure 2.9 but only station pairs which showed clear peaks at ±t_1 are plotted and used in the calculations of the mean. .......... 153
B.7 Wave spectra from buoy 46015 ........................................ 154
B.8 Cross-correlation function from Harmon et al. (2012). ................. 155
C.1 Beamforming output between September 2012 and September 2014. ... 190
C.2 Construction of estimated source spectrum for 15 November 2012 00:00-
03:00. ........................................................................ 191
List of Tables

2.1 Nearshore short-wave conditions for days of westerly infragravity arrivals  51
2.2 Estimates of reflection coefficient ................................. 54
4.1 Comparison of modelled and observed source locations ............ 104
B.1 Stations of the Cascadia Array used in the study. ................. 156
Declaration of Authorship

I, Jennifer F. Ward Neale, declare that this thesis and the work presented in it are my own and has been generated by me as the result of my own original research.

An Investigation into Ocean Wave Sources of Ambient Seismic Noise

I confirm that:

1. This work was done wholly or mainly while in candidature for a research degree at this University;

2. Where any part of this thesis has previously been submitted for a degree or any other qualification at this University or any other institution, this has been clearly stated;

3. Where I have consulted the published work of others, this is always clearly attributed;

4. Where I have quoted from the work of others, the source is always given. With the exception of such quotations, this thesis is entirely my own work;

5. I have acknowledged all main sources of help;

6. Where the thesis is based on work done by myself jointly with others, I have made clear exactly what was done by others and what I have contributed myself;

7. Parts of this work have been published as:


Signed:

Date:
Acknowledgements

I have been fortunate to have had two superb supervisors, Dr. Nicholas Harmon and Prof. Meric Srokosz, who I thank for their guidance, encouragement and wisdom, and for sparking my interest in the subject in the first place. I extend my thanks to my panel chair, Prof. Tim Minshull for making sure I kept on track and for asking insightful questions during my Upgrade. Others who have kindly given their time to increase my understanding of the topic and widened my perspective I also thank; particularly the constructive feedback of Lucia Gualtieri, Sharon Kedar and other anonymous reviewers of my papers, Fabrice Ardhuin for help with WAVEWATCH III, and all those who have commented on my work at conferences. I am grateful to the University and NOC library staff who helped me navigate the system for publishing with the University, and to NERC for funding me. Thanks also go to my parents for their support. And of course my husband, Juan, who has been with me every step of the way.
Chapter 1

Introduction

1.1 Background

Seismometers all around the world record continuous tiny oscillations of the ground, unrelated to earthquakes. This ‘ambient seismic noise’ can be caused in part by human activity near the seismometer (machinery and traffic (Nakata et al., 2011; McNamara and Buland, 2004)) and environmental factors (including wind which interacts with topography, built structures and trees (McNamara and Buland, 2004) as well as rain and river run-off which create noise through transport of sediment (Burtin et al., 2008; Tsai et al., 2012), but by far the largest and most persistent contributor worldwide is ocean waves. Ocean waves produce periodic seismic signals which modulate in amplitude over time depending on wave activity and are known as oceanic microseisms (or just microseisms). Periods of oceanic microseisms range from about 2 seconds to over 100 seconds and amplitudes are on the order of 1-10 microns. They are generated by the coupling of pressure fluctuations beneath ocean waves with the seabed. These pressure fluctuations are also recorded by seismometers and pressure gauges on the seafloor.

The coupling of ocean pressure fluctuations with the seabed to generate propagating seismic waves can only occur if the speed of the ocean forcing matches that of seismic waves (Ardhuin et al., 2015). Two mechanisms are known to match the slower ocean waves with the faster seismic waves. The ‘primary’ mechanism involves the interaction of ocean waves with sloping bathymetry (Hasselmann, 1963) and results in seismic waves with the same frequency as the ocean waves. The interaction of ocean wind and swell waves with the sloping seabed creates what are known as ‘single frequency’ or ‘primary’ microseisms at periods of about 10-20 seconds. This process is limited to shallow water, i.e. depths of no more than one ocean wavelength, due to the exponential decrease in pressure with depth. The interaction of oceanic infragravity...
waves with the seabed creates ‘seismic hum’ at periods of 30-1000 seconds (Rhie and Romanovicz, 2004; Uchiyama and McWilliams, 2008; Webb, 2008). The ‘secondary’ mechanism involves the interaction of ocean waves with similar frequencies but travelling in opposite directions (Longuet-Higgins, 1950), and results in microseisms with frequencies twice that of the ocean waves that generated them. These have peak periods of about 5-7 seconds are known as ‘double frequency’ or ‘secondary’ microseisms. They can be generated at any water depth in the ocean but are amplified at specific depths.

Microseisms can be excited as both seismic surface and body waves. Rayleigh surface waves, especially the fundamental mode, are the dominant mode of energy propagation at periods of 5-20 seconds (Haubrich and McCamy, 1969; Koper et al., 2010) and therefore most microseism studies have focussed on these. Despite an expected weaker coupling between ocean waves and horizontal motions of the seafloor, surface Love waves have also been observed (Lin et al., 2008) and can be a large component at shorter periods (0.25-2 seconds (Haubrich and McCamy, 1969). Compressional body waves (P-waves) have been observed mostly in the double-frequency microseism band (Zhang et al., 2010a) and recently shear body waves (S-waves) have been observed (Nishida and Takagi, 2016).

The variation within the microseismic field in terms of frequency components and propagation type make for a complex picture of noise sources. One of the main areas of research has been in source location, and many studies have now shown that different frequencies are excited in different oceanic regions. Primary microseisms have been found to be generated in coastal waters, as expected from their generation mechanism which requires interaction of the exponentially decreasing pressure term beneath ocean waves with the seabed. Seismic hum sources also appear to be coastal. Particularly large hum signals have been detected from western coasts of northern landmasses (Bromirski and Gerstoft, 2009) because swell from eastward moving storms is transformed to infragravity waves as it impacts the western coastlines.

Double-frequency microseism sources have been located in deep as well as shallow waters but excitation depends critically on the water depth (Kedar et al., 2008). In addition, Rayleigh and P-wave microseisms observed at the same location may arrive from different sources and show differing temporal evolution due to different propagation paths and attenuation characteristics (Gerstoft et al., 2006; Obrebski et al., 2013). At least for Rayleigh waves, the location of the seismometer on the seabed, continent or ocean island seems to affect the sources that the seismometer is sensitive to. Numerous studies have shown that continental sites are dominated by coastal Rayleigh wave sources (Ardhuin et al., 2011; Traer et al., 2012; Gerstoft and Tanimoto, 2007; Bromirski and Duennebier, 2002; Bromirski et al., 1999, 2005) and there has been ongoing debate about whether deep ocean generated Rayleigh wave
microseisms (Obrebski et al., 2012; Cessaro, 1994) can effectively propagate onto land (Tian and Ritzwoller, 2015) or whether apparent deep-ocean generated microseisms can actually be explained by coastal sources (Bromirski et al., 2013). Ocean island stations however seem to detect Rayleigh waves from multiple deep ocean regions (Aster et al., 2008; Bromirski et al., 2005) which has been explained by oceanic islands sharing oceanic basaltic crust, resulting in less energy loss than that associated with propagation onto the continent (Bromirski et al., 2013). P-waves have been observed from both coastal and deep ocean locations (Gerstoft et al., 2006; Haubrich and McCamy, 1969; Zhang et al., 2010a). Understanding source locations is important for geophysicists who use ambient noise to image earth structure and is essential for relating seismic observations to ocean wave activity, which is the focus of this thesis.

1.1.1 Historical context and motivations for studying ambient noise

Microseisms and their connection with storms first captured the attention of seismologists and oceanographers over one hundred years ago. In particular, studies of microseisms were motivated by their potential to warn of approaching storms at a time when other meteorological sources were scarce (Banerji, 1930). Microseisms provided an advantage over another indicator of storms- ocean swell- because seismic waves travel much faster and would arrive more than 24 hours before the swell for a storm 1000 miles away (Deacon, 1949).

However it was not until the mid 20th Century that the mechanism linking storms and seismic signals together to be understood (Longue-Higgins, 1950). Initially, it was theorised that microseisms were generated by the breaking of surf against steep, rocky coasts (Wiechert, 1905) and this seemed to be supported by observations that found correlations between surf in Norway and microseisms recorded in northern and eastern Europe (Gutenberg, 1931, 1936). However, a strengthening of microseisms was sometimes observed whilst a storm was still far offshore (Banerji, 1930; Ramirez, 1940), which could not be explained by the theory. Perhaps the strongest evidence against this theory was found by Gilmore (1946), who found that bearings to a source obtained by tripartite seismic arrays coincided with the tracks of hurricanes, and that the cross-bearings from two arrays located the source at the hurricane centre.

New ideas then emerged to address these observations, implicating pressure changes over the ocean as the cause; either air pressure associated with weather systems (e.g. Gherzi (1932)) or pressure beneath ocean waves (Banerji, 1930). The disturbances in pressure were supposed to propagate outwards from the source as surface seismic waves. Whilst the latter of these theories was correct in identifying the importance of pressure beneath ocean waves in the generation of microseisms, it still could not explain satisfactorily how disturbances at the surface of the ocean could propagate
down to the seabed. According to the water wave theory of the time, pressure fluctuations must decrease exponentially with depth, and would be negligible in the deep ocean where the microseisms were observed to be coming from.

The existence of second-order pressure fluctuations, importantly unattenuated with depth, under standing water waves was described in the work of Miche (1944) and the relevance to microseismic generation was identified by Longuet-Higgins (1950), who explained how this pressure term would be caused by any pair of ocean waves of similar frequency travelling in opposite directions. The wave conditions required for the interaction can take place beneath fast-moving storms, under rapidly changing winds, when two swell systems meet or when swell reflects from the coast. According to the theory, this pressure fluctuation would be twice the frequency of the forcing ocean waves, which supported recent observations that microseisms were just half the period of the corresponding sea states. The amplitude of the pressure fluctuation would be proportional to the product of the two ocean wave amplitudes, and would be amplified at certain water depths due to resonance. The theory could even explain how energy transfer from the ocean wave to the much longer-wavelength seismic wave was possible, as the wavelength of the pressure term (of near-zero wavenumber) would be comparable to a seismic wave of the same frequency. The conditions required for this generation mechanism were expected under a fast moving depression or where waves are reflected from a steep coast (Deacon, 1949). Laboratory tests (Cooper and Longuet-Higgins, 1951) confirmed Longuet-Higgins’ theory of microseism generation.

This theory of microseism generation remains largely unchanged, although Hasselmann (1963) extended the theory to include waves interacting with sloping bathymetry (primary microseisms) and recent authors have expanded it to include wave interactions over the full directional spectrum (Traer and Gerstoft, 2014) which adds an evanescent double-frequency contribution in shallow waters, as well as consideration of bottom effects and generation of body waves (Ardhuin and Herbers, 2013).

The Earth’s seismic hum, the low frequency oscillations between 2 and 7 mHz, was only discovered much more recently (Nawa et al., 1998) and its source was elusive until Rhie and Romanowicz (2004) found evidence linking the free oscillations to oceanic infragravity waves. The preferential excitation of the fundamental mode suggested a source close to the Earth’s surface (Suda et al., 1998), and observed seasonal (Tanimoto and Um, 1999; Ekström, 2001) variations in energy with maxima occurring in January and July pointed to the role of the atmosphere in their generation. Local atmospheric variations above a seismic station could be ruled out because correcting for local barometric pressure actually caused the free oscillations to be brought out more clearly (Roult and Crawford, 2000).

Another hypothesis identified atmospheric turbulence, uniformly distributed over land and sea, as a cause (Kobayashi and Nishida, 1998; Fukao et al., 2002; Tanimoto and
Um, 1999; Nishida et al., 2000; Ekström, 2001). This hypothesis could explain the energy levels of hum (Tanimoto and Um, 1999; Ekström, 2001), but there was no evidence of atmospheric convection at the required scale to confirm it (Rhie and Romanowicz, 2006). Rhie and Romanowicz (2004) found evidence for an oceanic origin at the same time as Tanimoto (2005) was developing a corresponding theory. Rhie and Romanowicz (2004) used seismic arrays in Japan and California to detect the arrival direction of hum, and found that it originated from the North Pacific during northern hemisphere winter, and the Southern oceans during southern hemisphere winter, coinciding with increased wave activity in each hemisphere. They suggested that the conversion of short-period ocean waves to longer-period infragravity waves (Webb et al., 1991), which is thought to occur at the coast by incident swell (Longuet-Higgins and Stewart, 1962; Symonds et al., 1982), and the subsequent interaction of infragravity waves with seafloor topography, was the likely cause. This has since been supported by other studies (Webb, 2007; Rhie and Romanowicz, 2006).

The following sections describe in more detail the theory and previous studies of double-frequency microseisms and infragravity waves, which are the topics of Chapters 2, 3 and 4.

1.1.2 Double frequency P-wave microseisms

The theory of microseism generation by Longuet-Higgins (1950) provides the basis for modern studies of double frequency microseisms. Namely, that when wave trains of similar frequency travelling in opposite directions interact, they create a standing wave with a second-order pressure oscillation $\bar{p}$:

$$\frac{\bar{p} - p_0}{\rho} - gh = -2a_1a_2\sigma^2\cos(2\sigma t)$$

where $p_0$ is ambient hydrostatic pressure, $\rho$ is the fluid density, $h$ is fluid depth, $a_1$ and $a_2$ are the amplitudes of the two waves, $\sigma$ is angular frequency of the ocean waves. It can be seen from the right hand side of equation 1.1 that the pressure oscillation is twice that of the forcing ocean waves and is independent of depth $h$.

Ardhuin et al. (2014) make clear that this pressure pattern is the result of the sum interaction of ocean wave trains with frequencies $f_1$ and $f_2$ and wave number vectors $k_1$ and $k_2$ (i.e. frequency of pressure fluctuation = $f_1 + f_2$ and wave number = $k_1 + k_2$), and that this results in a phase velocity $C_s = 2\pi|f_1 + f_2|/|k_1 + k_2|$ that matches, and hence is capable of exciting, seismic waves when the frequencies of the two ocean waves are nearly equal and their directions nearly opposing ($k_1 + k_2$ goes to zero). They also note that difference interactions ($f_1 - f_2; k_1 - k_2$) proposed in other
studies (Uchiyama and McWilliams, 2008; Traer and Gerstoft, 2014) exist but do not match the velocity of seismic waves.

The pressure forcing term can be considered as applied to the upper surface of the ocean, which propagates downward unattenuated with depth as an acoustic wave. Response kernels of the fluid-solid system are dependent on the ratio $\sigma h/\beta$ where $\sigma$ is the angular frequency, $h$ is water depth and $\beta$ is the elastic shear velocity of a solid half-space (the seabed), and display resonant excitation when depth of water becomes close to $1/4$, $3/4$, $5/4$ etc. of the acoustic wavelength (or more generally, $\frac{1}{2}(m - 1) + \frac{1}{4}$ for integer $m$) (Kedar et al., 2008). Longuet-Higgins (1950) only considered the generation of Stoneley waves, but recently Ardhuin and Herbers (2013) and Gualtieri et al. (2014) have formulated the P-wave site response kernels (see section 1.2.4.3).

The importance of the wave-wave interaction mechanism for the generation of double-frequency microseisms was confirmed by laboratory tests in 1951 (Cooper and Longuet-Higgins, 1951), but the resonance effect resulting from a compressible ocean was not validated until more recently (Kedar et al., 2008). This study by Kedar et al. (2008) tested the theory quantitatively using the ocean wave model WAVEWATCH III (Tolman, 2002b) to hindcast the ocean wave spectrum over the North Atlantic and North Pacific oceans, from which the wave-interaction intensity was calculated and multiplied by the depth-dependent site-effect (they calculated this for Rayleigh waves). They identified regions that were favourable for Rayleigh wave microseism generation due to a combination of wave conditions and ocean depth, and that wave-wave interaction in the North Atlantic could successfully predict observed ground displacements in North America, Greenland, Iceland and Europe. This was the first of an increasing number of studies to utilise the ocean wave model WAVEWATCH III to predict microseism sources and since then the model has been developed to account for better microseism generation at coasts. In particular, the version of the ocean wave model used by Kedar et al. (2008) did not include any coastal reflection, which is of course an important factor in the generation of opposing wave conditions, so they could only account for sources in the deep ocean. Ardhuin et al. (2011) improved the model to include a constant reflection coefficient for three basic shoreface types (continent, ocean island or iceberg), and then for a reflection coefficient that varies with shoreface slope, wave height and wave frequency (Ardhuin and Roland, 2012).

Obrebski et al. (2013) was the first study to model P-wave sources and compare the predictions with observed source locations obtained from beamforming of seismic data. They found a good agreement (by 1.1° to 9.9°) in location between the largest modelled sources and observed sources, which were distributed mostly along the storm tracks of the North Pacific and North Atlantic. In other studies, P-waves were found to be associated with hurricanes and cyclones (Gerstoft et al., 2006; Zhang et al., 2010a; Chen et al., 2015), storms (Haubrich and McCamy, 1969) and regions of large
waves (Gerstoft et al., 2008) both in the deep ocean (Landès et al., 2010; Gerstoft et al., 2008; Zhang et al., 2010a; Euler et al., 2014; Koper et al., 2010) and shallow waters (Gerstoft et al., 2006; Zhang et al., 2010a). In many cases, the P-waves have been found to be generated not at the hurricane centre, but in its wake. This was found by Haubrich and McCamy (1969), Zhang et al. (2010a) and Obrebski et al. (2013) and has been explained by opposing waves created when a turning storm, which excites waves in all directions, moves faster than its waves (Haubrich and McCamy, 1969; Zhang et al., 2010a). As would be expected by changes in wave conditions over summer and winter months, strong seasonality has also been observed in amplitude, source location and P-wave phase e.g. P, PP, PKP (Landès et al., 2010; Gerstoft et al., 2008; Euler et al., 2014; Koper and de Foy, 2008; Reading et al., 2014; Hillers et al., 2012). The relation between P-wave generation and large waves is, however, complex and still unclear. Over monthly or seasonal timescales, P-wave source locations have been found to coincide with regions of high significant wave height (Euler et al., 2014), but the wave-wave interaction mechanism required for microseism generation means that the existence of large waves is necessary but not sufficient to excite microseisms (Kedar et al., 2008; Kedar, 2011). That a larger amplitude seismic signal would be created by moderate, but opposing, waves than by large, monodirectional waves (e.g. Obrebski et al. (2012)) has implications for using P-waves (and microseisms in general) to monitor ocean wave conditions.

1.1.3 Infragravity waves and seismic hum

Infragravity waves are ocean surface waves that have much larger wavelengths (10’s km) and periods (>30 s) than the ocean swell and wind waves that generate them. They are now generally accepted as the cause of the Earth’s seismic hum, yet the process that converts shorter period ocean swell to the long period infragravity waves in the first place is still under discussion. The conversion is known to take place in the coastal zone, but there is uncertainty whether they are generated by a time-varying breakpoint (Symonds et al., 1982), group-bound long waves (Longuet-Higgins and Stewart, 1962), or a combination of both (List, 1992). Observations and modelling studies show that most infragravity wave energy is trapped alongshore by refraction (these are known as edge waves (Herbers et al., 1995b; Munk et al., 1964)), but some energy is reflected from the shoreline as free waves (which means they follow the dispersion relation for linear surface gravity waves) and radiated offshore. These are known as leaky mode waves (Munk et al., 1964). Infragravity wave bursts across ocean basins are associated with large swells arriving at the shoreline, particularly on the eastern side of ocean basins, which are transformed to infragravity waves which then propagate east to west across the ocean (Rawat et al., 2014).
The generation of seismic hum from infragravity waves has also been unclear (Webb, 2008; Bromirski, 2009; Nishida, 2013) but shallow continental shelf waters have been speculated as the dominant source (Webb, 2007) and has most recently been explained by the same mechanism that generates primary microseisms, but results from the interaction of free infragravity waves, rather than short-period waves, with bathymetry (Ardhuin et al., 2015). The mechanism was given by Hasselmann (1963) for primary microseisms and applied to seismic hum by Ardhuin et al. (2015). The following explanation follows Ardhuin et al. (2015). In a simple case, the bottom topography can be thought of as a “frozen wave train”, with \( f_1 = 0 \) and wave number \( k_1 \). The interaction of an ocean wave train with frequency \( f_2 \) and wave number \( k_2 \) with the bottom topography creates a bottom pressure oscillation with a frequency equal to the ocean waves and a wavelength of \( k = k_1 \pm k_2 \) which matches seismic wavelengths when \( k \) goes to zero. At a sloping seabed, ocean wavelengths are shorter and their amplitudes larger where depth is shallower, and are longer with smaller amplitudes in deeper water, assuming that energy is conserved. The horizontal integral of the bottom pressure over half a wavelength is increased by longer wavelengths and decreased by smaller amplitudes. At a specific depth \( D_0 \) where \( kD = 0.76 \), these two effects exactly cancel out over one wavelength: an increase in wavelength of the deeper-water half of the wave is balanced by an increase in amplitude of the shallower-water half. In water depths \( kD > 0.76 \), the pressure change due to the increased amplitude of the shallower-water half dominates over that due to a longer wavelength of the deeper-water half, resulting in a phase shift of average pressure relative to local pressure. In shallower water where \( kD < 0.76 \) wavelength dominates over amplitude and the phase shift changes sign. The effect of this is a large scale pressure field that oscillates around a dipole at \( D_0 \) with a scale on the same order as a seismic wavelength. Seismic sources are therefore located at depths around \( D_0 \) where bottom slope is significant, which for infragravity waves occurs mostly at shelf breaks, although mid-ocean ridges or seamounts (Fukao et al., 2010) may also contribute.

Ardhuin et al. (2015) validated this mechanism for the generation of seismic hum using an extension the ocean wave model WAVEWATCH III for infragravity frequencies up to 300 s (Ardhuin et al., 2014). By modelling the seismic hum sources as fundamental mode Rayleigh waves, they found that the mechanism could explain the power, frequency distribution and variability of observed vertical ground accelerations in the hum frequency band at three European seismic stations.

1.1.4 Contemporary motivations for study

In the last 15 years, interest in microseisms has been renewed by the discovery that stacked cross-correlations of ambient seismic noise between two seismometers
approximates the Rayleigh wave Green’s function between the two stations (Shapiro and Campillo, 2004; Shapiro et al., 2005). One station acts as a virtual source (in place of an earthquake or explosion) whilst the other acts as the receiver. The Green’s function contains information about the Rayleigh wave velocity between the two stations, and utilising an array of seismometers, maps of velocity can be constructed, which can be inverted for shear wave velocity structure in a process called ambient noise tomography (Sabra et al., 2005; Lin et al., 2008; Yao et al., 2008; Nicolson et al., 2012). Similar ambient noise tomography methods for body wave phases have also been developed (Zhan et al., 2010; Poli et al., 2012). Ambient noise helps overcome the limitations of traditional tomography sources, which are usually very distant (so high frequency information is lost) and clustered near plate boundaries (Shapiro et al., 2005).

Theoretically, the technique requires random, diffuse, isotropic wavefields. Microseism sources are typically anisotropic (Stehly et al., 2006) because some areas of the ocean are more conducive for generating microseisms than others, and the dominant directionality at a seismic station may change over time as these areas of the ocean are in turn “illuminated” (Traer et al., 2012) by favourable wave conditions. In practice though, stacking over long (~year-long) time series has been found to be sufficient to bring out the coherent noise for velocity estimates (Yang and Ritzwoller, 2008). However, consideration of the errors introduced from anisotropic noise sources (Harmon et al., 2010) and of the length of time series required for accurate velocity estimates has made the distribution of sources a question of high importance, especially for studies aiming to resolve seismic anisotropy or attenuation (Yao and van der Hilst, 2009; Harmon et al., 2010), or for time-varying imaging which aims to detect very small velocity changes over short timescales (Brenguier et al., 2008). For this reason, a focus of many recent studies has been to locate microseismic noise sources and analyse how they change seasonally.

Therefore, whilst the earliest studies of microseisms were driven by the need for better sea state observations and weather forecasting, recent studies have been largely motivated by geophysicists who use their properties to understand about the Earth’s structure. For the most part, wave buoys, meteorological buoys, satellite observations and numerical models have superseded microseismic indications of sea state and marine weather. However, there are still gaps and limitations in our vast observation networks that microseisms have the potential to fill, such as data gaps when wave buoys go offline (Donne et al., 2014), under-observed parts of the ocean such as mid-ocean regions and the southern hemisphere (Aster et al., 2008), and during the period between the repeat cycles of satellites. Seismic signals could in such a case be used to track and monitor storms and hurricanes across the ocean (Gerstoft et al., 2006; Zhang et al., 2010a; Davy et al., 2014; Chen et al., 2015). Seismic records could also be useful
for calibrating wave model hindcasts (Ardhuin et al., 2012), and historical seismic records may indicate past sea-states when direct observations were limited, and reveal how climate change has affected ocean storminess (Grevemeyer et al., 2000).

Earth’s seismic hum can provide information about deep water infragravity waves which are difficult to measure directly because of their low wave height (Webb et al., 1991; Traer et al., 2012). In the coastal zone, the importance of infragravity waves has been well documented, for example with respect to sediment transport and morphology (Aagaard and Greenwood, 2008; Reniers et al., 2004), harbour oscillations (Okihiro and Guza, 1996) and surface elevation and run-up during storms (Guza and Thornton, 1982). Because they are difficult to measure, they have been less studied in the deep ocean. However they have been given more attention recently because of the realisation of their importance in coupling processes in oceans, ice and atmosphere (Godin et al., 2013), because of the error they may contribute to future high-resolution satellite altimeter measurements of sea level (Aucan and Ardhuin, 2013) and because they have been implicated in the break up of ice shelves in Antarctica (Bromirski et al., 2010). The pressure signal they create at the seabed can therefore be particularly useful for monitoring infragravity waves in the deep ocean, including determining their source regions and their propagation across the ocean basins.

Each of these applications requires a different way of using and understanding microseismic data so the study of microseisms and what they can tell us about the oceans has several branches. For example, in order to use historical seismic data to indicate past sea states, methods applicable to single stations need to be developed as these have operated for the longest amount of time. This may include correlation of currently operating stations with sea states (Bromirski et al., 1999; Bromirski and Duennebier, 2002; Essen et al., 2003; Aster et al., 2008) or polarisation studies to determine the direction of a source from a single station (Schulte-Pelkum et al., 2004; Stutzmann et al., 2009; Schimmel et al., 2011). One of the important questions in this area is whether historic records at particular stations can serve as an indicator for global storminess or whether they are representative of a more local region (Bromirski et al., 2013). Most work suggests that ocean island stations incorporate signals from the whole ocean basin, whilst continental stations are mostly sensitive to waves at nearby coasts. In order to fill data gaps in currently operating coastal wave buoys, we need to understand the recordings on nearby coastal seismometers. Again, whether the station is representative of local coastal sea states, or whether it occasionally records signals from storms further offshore (Ardhuin et al., 2011) is important. For real-time estimation of sea state in remote and under-observed parts of the ocean, and for tracking storms, the large and dense seismic arrays now available can be utilised to locate the microseism source. Seismic arrays allow the direction of the signal to be identified using array beamforming (Traer et al., 2012; Gerstoft and Tanimoto, 2007;
Bromirski and Gerstoft, 2009) which also separates surface wave and body wave arrivals by their slowness. For surface waves, distance to the source is not constrained and back-azimuths often intersect large swaths of ocean (Traer et al., 2012). Triangulation from different arrays has been attempted (Cessaro, 1994; Essen et al., 2003; Chevrot et al., 2007) but different sites may detect unrelated sources and precision is poor (Chevrot et al., 2007). Using P-waves however, the distance to the source can be estimated from the apparent slowness (Gerstoft et al., 2008; Zhang et al., 2010a), due to the fact that P-waves from large distances will propagate up through the Earth towards the seismometer almost vertically, whilst those from nearby will be propagating closer to horizontal. For real-time monitoring of remote ocean waves, the main questions are therefore related to how well microseism sources can be located and what exactly they can tell us about wave activity.

The overall aim of this thesis is to investigate how pressure and seismic records can be used to study or monitor ocean waves. To address this aim the three objectives are 1) to identify infragravity wave propagation on seafloor pressure gauge records, 2) to estimate significant wave height in the deep ocean from seismometers located on land and 3) to investigate how microseism source location can be improved by combining seismic records from multiple arrays.

These objectives are addressed in Chapters 2, 3 and 4. In Chapter 2, sources of infragravity waves are located from cross-correlation of ocean bottom pressure measurements offshore California. The asymmetry of the cross-correlation functions are then used to estimate the reflection coefficient at the coast, which has implications for the propagation of infragravity waves across the ocean and energy levels in the deep ocean. In Chapter 3, the focus is the potential real-time monitoring of double-frequency P-wave microseism sources and ocean wave conditions in the deep ocean, in particular understanding how P-wave observations are related to significant wave height. A large seismic array in California is used to estimate significant wave height in the North Pacific. Chapter 4 extends the work of Chapter 3 to consider how source location can be improved using arrays in Japan and Europe in addition to the array in California.

The rest of this chapter outlines the data and methods used to locate sources of infragravity waves and P-wave microseisms. Because each chapter is written as a self-contained paper, there is some overlap and repetition between the following sections and the Data and Methods sections of each paper. However, here more emphasis is placed on explaining and illustrating the basic methods and identifying common threads within the chapters, with finer details and adjustments kept for each chapter.
1.2 Data and methods

The data and methods used in the Chapter 2 study on infragravity waves are described in sections 1.2.1 to 1.2.2. Sections 1.2.3 to 1.2.4 do the same for the P-wave studies of Chapters 3 and 4.

1.2.1 Ocean bottom pressure

Ocean bottom pressure records are used in Chapter 2 to study deep water infragravity waves. The pressure records used came from Differential Pressure Gauges (DPG’s) deployed off the coast of California/Oregon as part of the Cascadia Initiative array (Toomey et al., 2014). Figure 1.1 shows the locations of the pressure gauges. Records were downloaded for the time period between September 2012 and May 2013 (the start and end dates of second deployment) from the IRIS Data Management Center (http://ds.iris.edu/ds/nodes/dmc/).

Each daily record with original sampling frequencies of 40-50 Hz was band-pass filtered between 0.002 and 0.450 Hz, downsampled to 1 Hz, detrended and tapered with a Hanning taper of width 0.01. 29 of the 39 stations were found to contain good data over the whole time period. The average power spectrum over the whole time period at station G02B is shown in Figure 1.2. The spectrum shows clear energy in the infragravity band as well as the double-frequency microseism peak. Infragravity waves with periods of 100 to 200 seconds, bounded by the red vertical lines, were studied in Chapter 2.

1.2.1.1 Cross-correlation

Cross-correlation of time-series records between station pairs indicates the direction of propagation of coherent signals across the stations. Cross-correlation functions were calculated between each pair of daily ocean bottom pressure records. Figure 1.3 shows cross-correlation functions at each station stacked (summed) over the entire time period. The daily cross-correlation functions were band-pass filtered to a central frequency ± 0.0015 Hz. The central frequencies used were 0.0100, 0.0080, 0.0067, 0.0057 and 0.0050 Hz (periods of 100, 125, 150, 175 and 200 s) and were chosen to cover the peak of infragravity energy shown in Figure 1.2. A Hilbert transform was used to calculate the envelope of each cross-correlation function and each was normalised to a maximum of 1 by dividing by the maximum of the envelope. Figure 1.4 shows examples of stacked filtered cross-correlation functions. The lag time of the envelope corresponds to the wave group travel-time between the two stations.
Figure 1.1: Maps showing locations of DPG stations and wave buoys. a) DPG stations of the Cascadia array shown by inverted triangles. Only stations that returned usable data over the whole data period are shown. b) Enlargement of the boxed region of (a), also showing bathymetry. Red lines connect east-west orientated stations (azimuths of $265^\circ - 275^\circ$). National Data Buoy Center wave buoys 46015, 46022 and 46027 are marked by the blue circles and labelled. This figure has been published along with the content of Chapter 2 as Neale et al. (2015).
Chapter 1 Introduction

1.2.1.2 Backprojection

The envelopes of the cross-correlation functions were backprojected onto a latitude-longitude grid to locate sources of infragravity waves within and crossing the pressure gauge array. The grid covered the array and immediate vicinity, 35 to 50°N, 135 to 124°W, with a spatial resolution of 1 arc-minute. Following Brzak et al. (2009) and Harmon et al. (2012):

\[ P(f_p, l) = \sum_{n=1}^{N} W_n(f_p)env(C_n(f_p, T_n(f, l))) \]  \hspace{1cm} (1.2)

\( C_n \) is the cross-correlation function for station pair \( n \), and \( env(C_n) \) is its envelope. \( f_p \) is the frequency pass-band to which the cross-correlation function has been filtered. \( T_n \) is the theoretical group lag time (positive or negative travel-time) of an infragravity wave with frequency \( f \) (which is the central frequency of the pass-band \( f_p \)) between station pair \( n \) for a source located at \( l \), where \( l \) is the index of each unique latitude and longitude point on the spatial grid. \( W_n \) is a weighting coefficient for station \( n \) to remove the effect of array geometry, assumed 1 unless stated otherwise.
Figure 1.3: Stacked cross-correlation functions, band-pass-filtered between 60 and 500 s. This figure has been published along with the content of Chapter 2 as Neale et al. (2015).

The theoretical group lag time $T_n(f, l)$ is calculated as:

$$T_n(f, l) = t_i(f, l) - t_j(f, l)$$  \hspace{1cm} (1.3)

where $t_i$ and $t_j$ are the group travel-times of a wave with frequency $f$ from a hypothetical source at $l$ to station $i$ and station $j$ respectively.
Figure 1.4: Examples of stacked cross-correlation functions for east-west aligned station pairs filtered to 100 s (top row), 150 s (middle row) and 200 s (bottom row). Vertical lines are plotted at \( \pm \) the theoretical group travel time \( t_1 \) between the two stations for an infragravity wave at the given period traveling along the direct raypath between the two stations. Positive lags represent waves propagating from Station 1 to Station 2, and in all these cases Station 2 is furthest offshore. This figure has been published along with the content of Chapter 2 as Neale et al. (2015).
To calculate the travel-times $t_i$ and $t_j$, the group velocity $v_g$ at frequency $f$ was first calculated at each location on the grid. Group velocity was calculated using the dispersion relation for linear surface gravity waves:

$$\omega^2 = gk \tanh(kh)$$  \hspace{1cm} (1.4)

and

$$v_g = \frac{\partial w}{\partial k} = \frac{g \tanh(kh) + gkh(1 - \tanh(kh)^2)}{2\sqrt{gkh \tanh(kh)}}$$  \hspace{1cm} (1.5)

$\omega = 2\pi f$ is the angular frequency, $g$ is gravitational acceleration (= 9.81 ms$^{-2}$), $k$ is the wavenumber (radians per m) and $h$ is water depth (m). Water depth was taken from ETOP01 bathymetry (Amante and Eakins, 2009), plotted in Figure 1.5. Group velocities calculated at 100, 150 and 200 s (0.0100, 0.0067 and 0.0050 Hz) are shown in Figure 1.6.

Using the velocity grids, travel times from each station to each point on the grid were calculated using the FMST: Fast Marching Surface Tomography Package of Rawlinson and Sambridge (2004) (http://www.iearth.org.au/codes/FMST/). The software
computes a finite-difference solution of the eikonal equation (known as the fast marching method) to calculate the minimum direct travel-time between a source and receiver. Configuration input files used are given in Appendix A.1. Ray paths, computed by following the gradient of the travel-time field, are shown in Figure 1.7 from station G21B to the other stations at 150 s. Travel-time grids for this station at 100, 150 and 200 s are shown in Figure 1.8.
Once travel-time grids $t$ were calculated for each station, it was trivial to calculate the lag time $T_n$ between two stations (equation 1.3) by subtracting one grid from another. Lag-time grids for station $i = J11B$ and station $j = G21B$ are shown in Figure 1.9. These show the lag-time that would be observed between the two stations for a source at each grid point. The convention used is that positive lags correspond to a wave propagating from station 1 ($i$) to station 2 ($j$), whilst negative lags correspond to a wave propagating from station 2 to station 1. Maximum positive and negative lags occur when the source lies on the path that passes through both stations. For each frequency pass-band $f_p$ and grid point $l$, the backprojection is obtained by summing the value of the envelope of the cross-correlation function at $T_n$ over all station pairs (equation 1.2). Regions with the highest summed values are then interpreted as the strongest sources. Results are given in Chapter 2.
1.2.1.3 Incoherent beamforming

The second method used in Chapter 2 is incoherent beamforming of the cross-correlation envelopes. This type of beamforming is incoherent because it is based on the envelopes of the cross-correlation functions. Because envelopes have only positive values, incoherent arrivals do not cancel to zero, and the results are therefore noisier than conventional beamforming when array configuration allows for conventional beamforming. Beamforming is different to backprojection in that it identifies waves crossing the array by their frequency and slowness (reciprocal of velocity). Therefore, where backprojection finds the value of the cross-correlation envelope at the expected lag-time for a source located at \( l \), beamforming finds the value of the cross-correlation envelope at the expected lag-time for a source with group slowness \( s \) and azimuth \( \theta \):

\[
B(f_p, s, \theta) = \sum_{n=1}^{N} \text{env}(C_n(f_p, T_n(s, \theta)))
\]  (1.6)

The expected group travel-time \( T_n \) between the stations in station pair \( n \) is calculated for a synthetic plane wave with group slowness \( s \) and azimuth \( \theta \) as:

\[
T_n(s, \theta) = s \cdot d_n \cdot \cos(\theta - \phi_n)
\]  (1.7)

where \( d_n \) is the interstation distance and \( \phi_n \) is the interstation azimuth. Here sources are treated as plane waves which must have a source some distance outside the array. Therefore sources within the array cannot be identified, but it doesn’t require prior knowledge of the wave velocity as backprojection does. The azimuth of the highest beam power is interpreted as the strongest source direction. Results are given in Chapter 2.

1.2.2 Wave buoys

In Chapter 2, observations of ocean waves by nearshore wave buoys were used to characterise the potential for local infragravity wave generation, based on the relationship between short-period wave activity and infragravity wave height in shallow water (Herbers et al., 1995b; Ardhuin et al., 2014).

1-d ocean wave energy spectra were downloaded from three National Data Buoy Center wave buoys (http://www.ndbc.noaa.gov) off California from the period September 2012 to June 2013. The locations of the wave buoys (numbers 46015, 46022
and 46027) are plotted in Figure 1.1. The energy spectra are given at hourly intervals at frequencies between 0.0325 Hz and 0.4850 Hz.

Significant wave height $H_s$ and average wave period $T_{m0,2}$ (following Rawat et al. (2014)) were calculated from the average energy spectrum for each day:

$$H_s = 4\sqrt{m_0}$$  \hspace{1cm} (1.8)

$$T_{m0,2} = \sqrt{\frac{m_0}{m_2}}$$ \hspace{1cm} (1.9)

$$m_n = \sum_{f=fl}^{fu} f^n S(f)d(f)$$ \hspace{1cm} (1.10)

where $f$ is frequency in Hz, $S(f)$ is the non-directional (1-d) energy spectrum, $d(f)$ is the bandwidth of each frequency band, and $fl$ and $fu$ are the lower and upper limits of the integral ($fl = 0.0325$ Hz and $fu = 0.4850$ Hz).

1.2.3 Seismic records

In Chapters 3 and 4, vertical component passive seismic records from land stations were used to study double-frequency P-wave microseisms in the North Pacific. For Chapter 3 data from California between September 2012 and September 2014 were downloaded from the Southern California Earthquake Data Center (SCEDC, 2013), and for Chapter 4 additional data from Japan’s Hi-net array (Obara et al., 2005) and European arrays http://www.orfeus-eu.org/eida/eida.html were downloaded for particular days during the winter of 2012/13. The stations are plotted in Figure 1.10. The California data had an original sampling rate of 1 Hz, the Japan data 100 Hz and the Europe data 1-100 Hz depending on network and station. All the records were first downsampled to 1 Hz. Records were band-pass filtered between 0.05 and 0.4 Hz, instrument response was removed and records were converted into velocity seismograms (m/s) in Chapter 3 and displacement seismograms (m) in Chapter 4. Each daily record was then detrended and tapered with a Hanning taper width of 0.01. Finally, earthquakes $\geq$ magnitude 5 were removed using earthquake events listed in ISC bulletin (International Seismological Centre, 2013) by setting 1 hour of the waveform to zero if the RMS of that 1-hour window was over 3 times the daily RMS, and daily spectra at each station were then individually examined and bad quality days discarded.
Figure 1.10: Location of the a) Californian, b) Japanese and c) European arrays along with array configurations. Stations are marked by network code (BK= Berkeley Digital Seismograph Network, CI=Southern California Seismic Network, TA=USArray Transportable Array, AZ=ANZA regional network, 0101=Hi-net High Sensitivity Seismograph Network Japan, CH= Switzerland Seismological Network, GU=Regional Seismic Network of North Western Italy, IV=Italian National Seismic Network, NI=North-East Italy Broadband Network, SL=Slovenia Seismic Network. Total number of stations: California: 195, Japan: 201, Europe: 186.

An example of a daily time series for a station in California is shown in Figure 1.11 and the average spectra for this station over September 2012 - September 2014 (between 0.002 and 0.4 Hz) is shown in Figure 1.12. Because this station is recording ground motion on land, the spectrum includes the primary microseism peak at 0.06 Hz resulting from short-period waves in shallow water, that were not recorded at the deep water pressure gauge shown in Figure 1.2. The vertical red lines indicate the frequency range between 0.1 and 0.3 Hz used to study double-frequency P-wave microseisms in Chapter 3. Figure 1.13 shows average spectra over 35 days of winter 2012/13 for a station in each array between 0.1 and 0.3 Hz.

1.2.3.1 Beamforming

Frequency-domain beamforming was used to examine P-wave microseisms in the North Pacific. As mentioned previously, the slowness and azimuth result of beamforming can be projected onto a point on the globe for P-waves as slowness is related to distance. In Chapter 3 we followed the method of Gerstoft et al. (2008) to locate P-wave sources in the North Pacific from the California array. The beamformer output as a function of frequency $f$, slowness $s$, back-azimuth $\theta$ and time step $t_s$ was calculated as:

$$ B(f, s, \theta, t_s) = \frac{1}{N^2} [p(f, s, \theta)H C(f, t_s)p(f, s, \theta)] $$

(1.11)
Figure 1.11: Velocity seismogram at station STC of the Southern California Seismic Network on 15 February 2013.

Figure 1.12: Average power spectra for station STC of the Southern California Seismic Network, September 2012-May 2013. Vertical lines are drawn at frequencies of 0.1 Hz and 0.3 Hz, which is the range studied in Chapter 3. dB relative to 1 (m/s)^2/Hz. The power spectrum was calculated on 86400 s (whole day) windows.
Figure 1.13: Average power spectra between 0.1 and 0.3 Hz calculated over 35 days during the winter of 2012/13. a) Station STC of the Southern California Seismic Network b) Station OTAH of the Japan Hi-net network c) Station ATTE of the Italian National Seismic Network. dB relative to 1 m$^2$/Hz. Power spectra were calculated on 512 s windows.

where $C$ is the cross-spectral density matrix averaged over 3-hour periods, $C = \langle vv^H \rangle$. $H$ denotes the complex conjugate transpose. $v$ is the Fourier transform of the daily time series, which was calculated on 512 s windows and divided by its magnitude to retain only the phase. $p$ is the plane-wave response of the array $p = \exp[-i2\pi fs(\mathbf{r})]$. $e = (\sin \theta, \cos \theta)^T$ contains the directional cosines for a plane wave with azimuth $\theta$, while $\mathbf{r} = (x - x_c, y - y_c)$ contains the $(x,y)$ coordinates of the seismometers with respect to the array centre $(x_c, y_c)$. The beamformer output is normalised by the number of stations $N$ in the array.

An example beamformer output for slowness values relevant to body-wave phases (0.0 to 0.1 s/km) is shown in Figure 1.14. Surface waves propagate with higher slowness; about 0.3 for Rayleigh waves. Slowness can be related to distance for various seismic body-wave phases (Figure 1.15). Slownesses of 0.04 to 0.08 s/km were found to be dominant at the California array (Appendix C Figure C.1), which could correspond to P-waves or PP-waves. Direct P-waves were found to have good correspondence with modelled P-wave sources, which is consistent with other studies (Gerstoft et al., 2008; Euler et al., 2014), and so beampower was projected onto a global grid assuming direct P-wave slownesses (Figure 1.16).

The beamforming equation can be expressed in alternative ways e.g. in Gualtieri et al. (2014):

$$B(f, s, t_s) = \frac{1}{N^2} \left\langle \sum_{j=1}^{N} |S_j(f, t_i)e^{-i2\pi fs(x_j-x_c)}|^2 \right\rangle$$  \hspace{1cm} (1.12)

where $S_j$ is the Fourier transform of the seismogram at station $j$ with start time $t_i$, $x_j$ is the position vector of station $j$, $x_c$ is the position vector of the array centre and $s$ is
Figure 1.14: Beamformer output on 2 December 2012 03:00-09:00 at slowness of 0 to 0.1 s/km (radial axis). Beampower has been normalised to a maximum of 0 dB.

Figure 1.15: Slowness and ray-paths for body wave phases. a) Distance vs. slowness for P-wave phases P (blue), PP (cyan), PKPab (red), PKPbc (black) and PKPdf (purple), plotted using AK135 travel-time tables (Kennett et al., 1995). b) Ray paths for the phases plotted in (a). P: P-wave bottoming in uppermost mantle. PP: free-surface reflection of P-wave. PKPab: P-wave bottoming in upper outer core. PKPbc: P-wave bottoming in lower outer core. PKPdf (also PKIKP): P-wave bottoming in inner core.
the horizontal slowness vector toward the source. The angle brackets denote averaging over each Fourier transform snapshot time and frequency bands to get output at a snapshot time $t_s$. It can also be expressed directly in terms of travel-time $t_P$ between the station and source at $x$ instead of slowness and azimuth (e.g. Farra et al. (2016) and Nishida and Takagi (2016)) and this was used in Chapter 4:

$$B(f, x, t_s) = \frac{1}{N^2} \left \langle \left| \sum_{j=1}^{N} S_j(f, t_i) e^{-i 2\pi ft_P(x)} \right|^2 \right \rangle$$  \hspace{1cm} (1.13)

1.2.4 WAVEWATCH III®

WAVEWATCH III® was used to model ocean wave spectra and the microseism source term on a global grid for Chapters 3 and 4. The model was run on the University of Southampton’s Iridis Compute Cluster.

1.2.4.1 Description of model

WAVEWATCH III is a third generation wind-wave modelling framework under continuing development at NOAA/NCEP (Tolman, 1997, 1999, 2002b, 2009, 2014) along with contributors in the WAVEWATCH III Development Group (WW3DG). The latest pre-release version available at the time of the study was used (version 5.08). The official model version 5.16 has since been released. The following notes on the fundamental workings of the wave model have been summarised from the manual and system documentation (WW3DG, 2016).
Spectral wave components in water with finite depth can be described using several phase and amplitude parameters. For monochromatic waves, the amplitude parameters can be described as amplitude, wave height or wave energy. For irregular waves, the variance of the sea surface is described using variance density spectra, commonly referred to as energy spectra. This variance spectrum, \( F \), is a function of phase parameters which are the wavenumber vector \( k \), wavenumber \( k \), direction \( \theta \), relative angular frequency \( \sigma \) and absolute angular frequency \( \omega \). The distinction between relative and absolute frequency arises in the presence of a mean current depending on whether the frame of reference moves with the mean current \( (\sigma = 2\pi f_r) \) or is fixed \( (\omega = 2\pi f_a) \).

Generally, it is assumed that depth, current and the variance spectrum vary in space and time at scales much larger than individual ocean waves. Linear wave theory can then be applied locally which interrelates the phase parameters \( k \), \( \sigma \) and \( \omega \) (WW3DG, 2016, equations 2.1 and 2.2):

\[
\sigma^2 = gk \tanh kd \quad (1.14) \\
\omega = \sigma + k \cdot U \quad (1.15)
\]

where \( d \) is the mean water depth and \( U \) is the mean current velocity. Equation 1.14 is the well-known dispersion relation for ocean surface gravity waves. As a result, only two independent phase parameters exist and the local variance spectrum becomes two dimensional. Within WAVEWATCH III, the wavenumber-direction spectrum \( F(k, \theta) \) divided by \( \sigma \) is used (= the wave action spectrum \( N(k, \theta) \)) due to the general conservation of wave action as opposed to wave energy in cases with currents, but the output of WAVEWATCH III is the more traditional frequency-direction spectrum \( F(f_r, \theta) \). \( F(f_r, \theta) \) can be calculated from \( F(k, \theta) \) using the transformation (WW3DG, 2016, equation 2.4):

\[
F(f_r, \theta) = \frac{\partial k}{\partial f_r} F(k, \theta) = \frac{2\pi}{c_g} F(k, \theta) \quad (1.16)
\]

where \( c_g \) is the group velocity. The frequency-direction spectrum \( F(f_r, \theta) \) is comparable to the directional ocean wave spectra output by directional wave buoys (Section 1.2.2) and integrated over all directions gives the one-dimensional energy spectrum \( F(f_r) \).

Wave propagation is described by (WW3DG, 2016, equation 2.7):

\[
\frac{DN}{Dt} = \frac{S}{\sigma} \quad (1.17)
\]
where $\frac{DN}{Dt}$ is the total derivative of wave action (moving with the wave component) and $S$ represents the net effect of sources and sinks of the spectrum $F$.

In WAVEWATCH III, this balance equation is given in Eulerian form for propagation in either Cartesian or spherical coordinates. In spherical coordinates defined by longitude $\lambda$ and latitude $\phi$ (for large scale applications of the wave model) the propagation term becomes (WW3DG, 2016, equations 2.9 to 2.15):

$$\frac{\partial N}{\partial t} + \frac{1}{\cos \phi} \frac{\partial}{\partial \phi} \dot{\phi} N \cos \theta + \frac{\partial}{\partial \lambda} \dot{\lambda} N + \frac{\partial}{\partial k} \dot{k} N + \frac{\partial}{\partial \theta} \dot{\theta} N = \frac{S}{\sigma}$$  \hspace{1cm} (1.18)

$$\dot{\phi} = \frac{c_g \cos \theta + U_\phi}{R}$$  \hspace{1cm} (1.19)

$$\dot{\lambda} = \frac{c_g \sin \theta + U_\lambda}{R \cos \phi}$$  \hspace{1cm} (1.20)

$$\dot{\theta}_g = \dot{\theta} - \frac{c_g \tan \phi \cos \theta}{R}$$  \hspace{1cm} (1.21)

$$\dot{\theta} = -\frac{1}{k} \left( \frac{\partial \sigma}{\partial d} \frac{\partial d}{\partial m} + k \cdot \frac{\partial U}{\partial m} \right)$$  \hspace{1cm} (1.22)

$$\dot{k} = -\frac{\partial \sigma}{\partial d} \frac{\partial d}{\partial s} - k \cdot \frac{\partial U}{\partial s}$$  \hspace{1cm} (1.23)

where $R$ is the radius of the Earth and $U_\phi$ and $U_\lambda$ are current components.

The net source term $S$ consists of many individual source terms that account for energy input, interactions and dissipation. From (WW3DG, 2016, equation 2.16):

$$S = S_{ln} + S_{in} + S_{nl} + S_{ds} + S_{bot} + S_{db} + S_{tr} + S_{sc} + S_{ice} + S_{ref} + S_{xx}$$  \hspace{1cm} (1.24)

$S_{ln}$ is a linear input term for model initialisation and initial wave growth. $S_{in}$, $S_{nl}$ and $S_{ds}$ are the main terms for deep water and represent atmosphere-wave interaction ($S_{in}$, dominated by exponential wind-wave growth), non-linear wave-wave interactions ($S_{nl}$) and wave-ocean interaction ($S_{ds}$) that generally contains dissipation (e.g. due to wave breaking). In shallow and extremely shallow water (depths much shorter than the ocean wavelength), wave-bottom interactions $S_{bot}$, depth-induced breaking $S_{db}$, and triad wave-wave interactions $S_{tr}$ become important. $S_{sc}$ is a term for scattering of waves by bottom features, $S_{ice}$ is a term for wave-ice interactions and $S_{ref}$ is a term for reflection off shorelines or floating objects. $S_{xx}$ can contain user defined source terms.

Some of these source terms have more than one option for how they are formulated in the model, and can be chosen depending on the desired trade-off between accuracy and computation time. These different formulations of source terms are chosen or switched
off entirely during compilation of the model, using a ‘switch’ file that contains a set of strings identifying the model options to be selected.

The wave model is run as a series of consecutive programs which include a grid preprocessor, a program to generate initial conditions, the main program and output processors. These programs require input text files that specify particulars such as the geographical grids, time steps, desired output parameters and how the input data files are organised. Descriptions of the switches, input files and input data used in the model run are given below.

### 1.2.4.2 Model setup

The switch file used for the model run is given in Appendix A.2 along with the short description of each switch as given in the WAVEWATCH III manual (WW3DG, 2016). This was based on the switch file used at Ifremer\(^1\) for routine global model runs driven by ECMWF winds (http://www.ecmwf.int/) which is included in the model distribution. An important addition is the switch REF1 for reflection at coasts. The full list of options can be found in the WAVEWATCH III manual (WW3DG, 2016).

The model was set up to run on a 0.5\(^\circ\) global grid from -78\(^\circ\)S to 78\(^\circ\)N with a nested 0.1\(^\circ\) grid covering the Cascadia region offshore California, 38 to 48\(^\circ\)N, -133 to -122\(^\circ\)W (for future work). Grids for bottom depth, sub-grid obstructions and a land-sea mask were generated using the software gridgenv3.0 (Chawla and Tolman, 2007) that comes with the WAVEWATCH III distribution, which uses ETOP01 bathymetry (Amante and Eakins, 2009) and a high resolution shoreline database (GSHHS- Global Self-consistent Hierarchical High-resolution Shoreline (Wessel and Smith, 1996)). The model was run for the period August 2012 to September 2014 on a daily basis, with the first day starting from calm conditions (no waves) and each subsequent day using a restart file generated by the previous run. Coastal reflection was set as a coefficient of 0.1 at the shoreline and 0.2 at sub-grid obstructions (islands) following Ardhuin et al. (2011).

6 hourly 10 m winds (u and v components) and sea ice concentration of ECMWF’s ERA interim data reanalysis (Dee et al., 2011) were used to force the model. The ECMWF data was downloaded in NetCDF format on a 0.5\(^\circ\) grid. Gridded output from the model was requested at hourly intervals for both the global and Cascadia grids. In addition to standard ocean parameters such as significant wave height and period, and 1-d (frequency) energy spectra, the microseism pressure term was output (see below). Frequency-direction spectra were output at selected grid points (although not used in this study). The input files that define the setup are given in Appendix A.3.

\(^1\)Much of the WAVEWATCH III setup was based on the guidance of Fabrice Ardhuin during a WAVEWATCH III workshop.
Validation of WAVEWATCH III output (which is usually done for wave bulk parameters of significant wave height, peak period and direction) is a significant task and has been undertaken by other authors previously (Chawla et al., 2009; Ardhuin et al., 2011; Ardhuin and Roland, 2012). As the inner workings of the model have not been changed here, a full validation of the model output using, for example, satellite data or wave buoy records, was not undertaken.

1.2.4.3 Modelling of P-wave microseisms

Modelling of P-wave microseisms from WAVEWATCH III output was based on work by previous authors, notably Ardhuin et al. (2011), Ardhuin and Herbers (2013), Gualtieri et al. (2014) and Farra et al. (2016).

For the purpose of modelling microseism sources from a wave model, it is more appropriate to use the formulation of the double-frequency pressure term given in Ardhuin et al. (2011) and Farra et al. (2016), which is calculated using the wave energy spectrum, rather than the formula of Longuet-Higgins (1950) based on wave amplitudes. This is:

\[
F_p(x, f_2) = [2\pi]^2[\rho_w g] f_2 E^2(x, f) I(x, f) \tag{1.25}
\]

where \( f \) is the ocean wave frequency, \( f_2 = 2f \) is the seismic frequency, \( x \) is location, \( \rho_w \) is water density, \( g \) is gravitational constant, \( E \) is the ocean wave energy spectrum and \( I \) is a non-dimensional function that depends on the wave energy distribution \( M \) over directions \( \theta \):

\[
I(x, f) = \int_0^\pi M(x, f, \theta)M(x, f, \theta + \pi) d\theta \tag{1.26}
\]

This equation of Ardhuin et al. (2011) is based on equation 2.15 of Hasselmann (1963) who also expressed the pressure in terms of the ocean wave energy spectrum, but the spectral density \( F \) has been recast as a function of frequency \( f \) instead of angular frequency as this is provided by numerical ocean wave models like WAVEWATCH III, and an integral from 0 to \( 2\pi \) in Hasselmann (1963) has been replaced with an integral from 0 to \( \pi \), as the integrand has a periodicity of \( \pi \).

This pressure term is calculated within and output from WAVEWATCH III. \( F_p \) was multiplied by the site effect for P-waves, \( C_P \), to obtain a map of P-wave sources at each time step:
\[ P(x, f_2) = F_p(x, f_2) \times [2|CP(x, f_2)|^{\frac{\rho_c(x)}{\rho_w}}]^2 \] (1.27)

where \( \rho_c \) is crustal density. The values of \( \rho_w \) and \( \rho_c \) were taken from the water and upper crustal layers of the global crustal model CRUST1.0 (http://igppweb.ucsd.edu/~gabi/rem.html), with \( \rho_w = 1020 \text{ kg m}^{-3} \) and \( \rho_c \) varying spatially.

The site effect term for P-waves, \( CP \), was calculated using the formulation of Gualtieri et al. (2014) which is based on the transmission and reflection coefficients at the seafloor of downgoing P-waves in the water column:

\[ \Phi_w(h, \omega, \theta_{Pw}) = 2\omega \frac{\cos \theta_{Pw}}{\alpha_w} h = 2\omega q_{w} h \] (1.34)

where \( \alpha_w \) is the compressional wave velocity of the water column, \( \alpha_c \) is the compressional wave velocity of the crust and \( \beta_c \) is the shear wave velocity of the crust.

\( CP \) was calculated at each grid point with depth \( h \), giving \( CP \) as a function of seismic frequency and water (ocean surface) take-off angle at each location \((CP(x, f_2))\). The site effect is strongest for vertically propagating P-waves (take-off angle of 0°) (Ardhuin and Herbers, 2013). For P-waves recorded at a station \( s \), the relevant take-off angle depends on the distance between the location of the source and \( s \).
Because the aperture of each array is much smaller than the distances between the station and the source locations, the distance between the centre of the array and the location of the source was used to find the relevant take-off angle for each array. Using the ObsPy (Beyreuther et al., 2010) Python package, crustal take-off angles $\theta_{pc}$ for distances of 15 to 99° were found. These crustal take-off angles were then interpolated onto the distances between the centre of the array and each grid point. The take-off angle at the ocean surface, $\theta_{pw}$, at each grid point was then calculated using Snell’s law (Gualtieri et al., 2014):

$$\sin \theta_{pw} = \frac{\alpha_w \sin \theta_{pc}}{\alpha_c}$$  \hspace{1cm} (1.35)

$C_P$ was then obtained for the relevant ocean surface take-off angle at each grid point, and multiplied by the pressure source term to calculate the P-wave source (equation 4.3). An average site effect independent of take-off angle can also be calculated by integrating $C_P$ over all take-off angles.
Chapter 2

Source regions and reflection of infragravity waves offshore of the USA’s Pacific Northwest.

This chapter has been published as ‘Neale, J., Harmon, N. and Strokosz, M. (2015), Source regions and reflection of infragravity waves offshore of the U.S.s Pacific Northwest, Journal of Geophysical Research: Oceans, 120, 6474-6491, doi:10.1002/2015JC010891’ and is included in Appendix B.2. The co-authors listed in this publication directed and supervised the research that forms the basis for the thesis.

2.1 Abstract

Infragravity waves are oceanic surface gravity waves but with wavelengths (10’s km) and periods (>30 s) much longer than wind waves and swell. Mostly studied in shallow water, knowledge of infragravity waves in deep water has remained limited. Recent interest in deep-water infragravity waves has been motivated by the error they may contribute to future high-resolution satellite radar altimetry measurements of sea level. Here, deep-water infragravity waves offshore of the Pacific Northwest of the USA were studied using Differential Pressure Gauges which were deployed as part of the Cascadia Initiative array from September 2012-May 2013. Cross-correlation of the records revealed direction of infragravity wave propagation across the array, from which source regions were inferred. The dominant source was found to be the coastline to the east, associated with large wind waves and swell incident on the eastern side of the basin. The source shifted southward during northern-hemisphere summer, and on several days in the record infragravity waves arrived from the western side of the Pacific. Asymmetry of cross-correlation functions for five of these westerly arrivals was
used to calculate the ratio of seaward to shoreward propagating energy, and hence estimate the strength of infragravity wave reflection at periods of 100-200 s. Reflection of these remote arrivals from the west appeared to be strong, with a lower bound estimate of $r=0.49\pm0.29$ (reflection coefficient ± standard error) and an upper bound estimate of $r=0.74\pm0.06$. These results suggest that reflection at ocean boundaries may be an important consideration for infragravity waves in the deep ocean.

2.2 Introduction

Low frequency infragravity waves are associated with wave groups of the higher frequency sea waves and swell in the coastal zone. Two mechanisms have been proposed for the generation of infragravity waves from the short-wave groups. One is that the interaction of shoreward propagating swell creates ‘bound’ or ‘forced’ infragravity waves (Longuet-Higgins and Stewart, 1962; Herbers et al., 1995a). As the swell waves break, the forced infragravity waves are released as free infragravity waves and are reflected from the beach. Free infragravity waves satisfy the dispersion relation for surface gravity waves and have longer wavelengths than forced waves of the same frequency (Webb et al., 1991). The second mechanism is that infragravity waves are generated by a time-varying breakpoint, with standing waves shoreward of the breakpoint and progressive infragravity waves radiating seawards (Symonds et al., 1982).

In either case, the seaward-propagating free infragravity waves can have two fates: those that travel seaward at oblique angles can become refractively trapped along the shoreline as ‘edge waves’ by a sloping beach or shelf (Herbers et al., 1995a; Munk et al., 1964); those that propagate directly seaward can escape into the open ocean as ‘leaky waves’ (Munk et al., 1964). Due to their long wavelength (Aucan and Arduin, 2013), only a small fraction of the infragravity energy escapes from the coast into the open ocean (<1% (Webb et al., 1991)), with most being trapped within a few hundred meters of the shore (Webb et al., 1991). The amount of energy leaked into the open ocean for a given short-wave spectrum and coastline is poorly understood (Aucan and Arduin, 2013). Variation in alongshore topography may be partly responsible (Uchiyama and McWilliams, 2008), although the model of Arduin et al. (2014) produces a good prediction of measured infragravity wave levels assuming a locally straight coast.

Infragravity waves that make it into the open ocean propagate with very little attenuation (Godin et al., 2013), and it is possible to observe infragravity waves that have been generated from coasts thousands of kilometres away on the other side of an ocean basin (Herbers et al., 1995a; Harmon et al., 2012).
Most studies of infragravity waves have been undertaken in shallow water on continental shelves where they are most energetic (Webb et al., 1991) and instrumentation is more accessible. Here they are also known as ‘surf beat’ or ‘swash’, and they are important for sediment transport and nearshore morphology (Aagaard and Greenwood, 2008; Reniers et al., 2004) and harbour oscillations (Okihiro and Guza, 1996). The first studies were undertaken by Munk (1949) and Tucker (1950).

Infragravity waves in the deep ocean have received less attention than shallow water infragravity waves partly due to their very small amplitudes in the deep ocean (<1 cm (Webb et al., 1991), several cm at most (Aucan and Ardhuin, 2013)). However, there has recently been a resumed interest in infragravity waves in the deep ocean as they have been recognised as important for coupling processes in the ocean, ice, atmosphere, and solid earth (Godin et al., 2013). Aucan and Ardhuin (2013) have shown that infragravity waves in the deep ocean may add significant error to sea level measurements associated with sub-mesoscale currents, which are due to be collected by future satellite radar altimetry missions. Bromirski et al. (2010) have recently shown that infragravity waves generated along the Pacific coast propagate transoceanic distances and can be implicated in the flexure and subsequent break up of Antarctic ice shelves. Infragravity waves at frequencies below 0.004 Hz may transfer energy from the ocean to the atmosphere (Livneh et al., 2007; Godin et al., 2015). The deformation of the seafloor under the pressure of infragravity waves is used in measurements of seafloor compliance to determine the shear velocity structure of the shallow oceanic crust (Crawford et al., 1998), and the propagation of infragravity waves over a sloping seafloor are thought to create low frequency seismic noise known as Earth’s seismic hum (Rhie and Romanowicz, 2006; Ardhuin et al., 2015).

Pressure sensors (or seismometers) deployed on the seafloor have been the most widely used approach to observe infragravity waves in the deep ocean (Godin et al., 2013). Using an array of pressure gauges in the southwestern Pacific off the South Island of New Zealand, Godin et al. (2014) observed strong directionality of the infragravity wavefield with the northwest coast of the South Island acting as a net source of infragravity wave energy. Webb et al. (1991) studied deep-water infragravity waves in the Pacific during November 1988 and identified infragravity waves originating from the Gulf of Alaska, the northwest Pacific and the southern tip of South America, but little from the southern ocean or tropical western Pacific.

A further and more comprehensive study of infragravity waves in the deep ocean was undertaken by Aucan and Ardhuin (2013). They analysed pressure records from 40 locations in the Pacific and Atlantic oceans to determine spatial and temporal variability of infragravity wave energy at depths of 3-6 km. Their inferred infragravity significant wave heights were found to reach larger values than estimated in previous work, reaching over 4 cm in episodic events. Energy levels in the Atlantic and Pacific
were found to be similar, and mid to high latitudes in both oceans displayed strong seasonal cycles associated with seasonal variability of wind-waves.

Other studies have noted the arrival of infragravity waves which seem to have been generated right across the other side of the ocean basin (Harmon et al., 2012), and a combined observational and modelling study (Rawat et al., 2014) has shown the coherent propagation of large infragravity wave bursts from one side of the basin to the other. The latter study made use of a global numerical model of free infragravity wave generation and propagation that has been under development recently (Ardhuin et al., 2014).

The aim of this chapter was to determine how strongly free infragravity waves reflect when they reach the coastline or shelf of an ocean basin. As far as the authors are aware, no estimate of deep-water infragravity reflection has yet been made, although reflection from the shoreline has previously been estimated from pressure gauges in shallow waters <13 m in depth (e.g. Herbers et al. (1995a); Sheremet (2002)) and from laboratory data (Battjes, 2004). Studies such as these have found that infragravity waves reflect strongly from the shore with reflection coefficients above 0.6. Considering that infragravity waves are capable of propagating right across the oceans, reflection at the ocean boundary, whether at the shoreline or shelf, may be important for infragravity energy in the deep ocean. Here we present estimates of the directionality and reflection coefficient for the infragravity wavefield offshore of the Pacific Northwest of the USA.

### 2.3 Data and methods

To measure deep water infragravity waves we used differential Pressure Gauge (DPG) records from the Cascadia Initiative array (Toomey et al., 2014) between September 2012 and May 2013, downloaded from the IRIS Data Management System (http://ds.iris.edu/ds/nodes/dmc/data/types/). The array consisted of 39 DPGs between depths of 107 and 4462 m offshore of the Pacific Northwest of the USA (Figure 1.1 in Chapter 1). The locations were considered far enough offshore (50-500 km) to be removed from the effects of infragravity edge waves at the coast, which are trapped within a few hundred meters of the shore (Webb et al., 1991).

Monthly spectra of the records were used to identify bad data. In total 29 stations returned usable data over the whole data period and these stations were used in the study. The station locations are listed in Appendix B Table B.1. For most of our analysis we exclude the most northerly stations (above 44°N), which fall outside the main cluster of stations.
The daily pressure records were band-pass-filtered between 0.002 Hz and 0.45 Hz using a 2nd order Butterworth filter prior to decimation to 1 Hz, then detrended and tapered.

In order to characterise local infragravity wave generation we also examined nearshore short-wave parameters using data from the National Data Buoy Center’s data buoys 46015, 46022 and 46027 (http://www.ndbc.noaa.gov), which are also shown in Figure 1.1 in Chapter 1. Daily significant wave height, $H_s$, and average wave period, $T_{m0,-2}$, were calculated for each buoy from the daily average spectra as follows:

$$H_s = 4\sqrt{m_0} \quad (2.1)$$

$$T_{m0,-2} = \sqrt{\frac{m_0}{m_2}} \quad (2.2)$$

$$m_n = \sum_{f=fl}^{fu} f^n S(f) d(f) \quad (2.3)$$

where $f$ is frequency in Hz, $S(f)$ is the non-directional wave spectrum, $d(f)$ is the bandwidth of each frequency band, $fl=0.0325$ Hz and $fu=0.4850$ Hz. We use $H_s$ as a proxy for local infragravity wave generation because previous studies have found a high correlation between $H_s$ and infragravity wave height in shallow water (Herbers et al., 1995b). In addition, we calculated $\alpha H_s T_{m0,-2}^2 \sqrt{\frac{g}{D}}$ for each buoy where $g=9.81$ m s$^{-2}$, $D$ is water depth (m) and $\alpha$ is a dimensional constant with units of s$^{-1}$, because this parameter has been found to improve the correlation between infragravity waves and short-wave conditions (Ardhuin et al., 2014). With $\alpha=12 \times 10^{-4}$ s$^{-1}$ the parameter empirically models the observed free infragravity wave height (Ardhuin et al., 2014; Rawat et al., 2014). For our purposes, the value of $\alpha$ does not matter as we are interested in the relative change of infragravity wave generation with short-wave conditions rather than absolute wave heights, but we used $\alpha=12 \times 10^{-4}$ s$^{-1}$ so that the modelled infragravity wave heights can be compared with other studies. We used both these measures, $H_s$ and $\alpha H_s T_{m0,-2}^2 \sqrt{\frac{g}{D}}$, as proxies for local infragravity wave generation, and we averaged $H_s$ and $\alpha H_s T_{m0,-2}^2 \sqrt{\frac{g}{D}}$ over the three wave buoys.

### 2.3.1 Cross-correlation

A cross-correlation function was computed between each DPG station pair on each day to aid in identifying coherent signals between each station pair. A stack over the data period was computed by summing the daily cross-correlation functions to get a sense
of the long term average wavefield (Figure 1.3 in Chapter 1). Each cross-correlation function was band-pass-filtered to a central frequency \( \pm 0.0015 \) Hz (a wider range was found to make the records too spiky, and a narrower range too smooth, to identify the main peaks), and the envelope of the signal was calculated using a Hilbert transform. The central frequencies used were 0.0100, 0.0080, 0.0067, 0.0057 and 0.0050 Hz corresponding to periods of 100 s, 125 s, 150 s, 175 s and 200 s which covered the main infragravity band. The differential pressure gauges were uncalibrated, so each cross-correlation function was normalised by the maximum of its envelope. Figure 1.4 in Chapter 1 shows examples of these band-pass-filtered cross-correlation functions.

Asymmetry of the cross-correlation functions gives information on the direction of infragravity wave propagation across the array. For example, if energy travels from Station 1 to Station 2, the time series at Station 2 will lag the time series at Station 1 by \( t_1 \) seconds, where \( t_1 \) is the time it takes for energy to travel between the two stations, and a peak in the cross-correlation function at \( t_1 \) will result. Likewise, if energy is traveling from Station 2 to Station 1, the time series at Station 1 will lag Station 2 by \( t_1 \), resulting in a peak at \( -t_1 \) in the cross-correlation function. If energy is traveling perpendicular to the Station 1-Station 2 alignment, a peak at zero lag would be expected, as both stations receive the signal at the same time. Two methods were used to combine the information contained in all the individual cross-correlation functions from the array: backprojection and beamforming.

Beamforming did not require calculation of unique travel time grids for each frequency, as was necessary for backprojection, so it allowed us to examine directionality over many frequencies, and was more appropriate for analysing temporal changes in wave propagation over the array from sources outside of the array. However, due to uncertainties about the quality of beamforming over varying bathymetry, backprojection was used to verify the beamforming results, and was useful for examining sources along the coast close to the array.

2.3.2 Backprojection

Backprojection of the infragravity wave energy allows us to examine the spatial distribution of wave generation inside and outside of the array and to determine the direction of energy propagation across the array.

For each frequency band of interest, \( f_p \), the enveloped cross-correlation functions were backprojected onto a spatial grid of latitude \( \varphi \) and longitude \( \gamma \) using a method similar to that used by Harmon et al. (2012) and Brzak et al. (2009), and given in equation 2.4:
Chapter 2 Infragravity waves off Pacific Northwest

\[ P(f_p, l) = \sum_{n=1}^{N} W_n(f_p) \text{env}\left(C_n(f_p, T_n(f, l))\right) \]  

(2.4)

\( P \) is the backprojection as a function of frequency pass-band \( f_p \) and location index \( l \), \((\varphi_l, \gamma_l)\), where \( l \) is the index of each unique latitude and longitude point on our spatial grid (1,2...L grid points). \( \text{env}(C_n) \) is the enveloped band-pass-filtered cross-correlation (with centre frequency \( f \pm 0.0015 \) Hz) for station pair \( n \). The envelope was calculated using a Hilbert transform, and the maximum value for the envelope was normalised to 1. \( W_n \) is a weighting coefficient for station pair \( n \), described below, to reduce the effect of array geometry on the projection. For an unweighted backprojection, \( W_n = 1 \) for all \( n \) and \( f \). \( T_n \) is the theoretical group lag time for station pair \( n \) for a hypothetical source at \( l \) with frequency \( f \). \( T_n \) is calculated using:

\[ T_n(f, l) = t_i(f, l) - t_j(f, l) \]  

(2.5)

where \( t_i \) (\( t_j \)) is the group travel time from the source at \( l \) with frequency \( f \) to station \( i \) (\( j \)) of the station pair. Group travel times between the source and station are minimum direct travel times calculated using a ray theoretical approach following Harmon et al. (2012). This approach is more accurate than a calculation from interstation distance/average group velocity along the great circle path between the two stations, as it takes into account bathymetry and the effects of non-great circle propagation paths. \( t_i \) (\( t_j \)) were calculated in the following way: First, group velocity of infragravity waves at each frequency of interest at each grid point, \( v_g(f, l) \), was calculated using ETOP01 bathymetry (Amante and Eakins, 2009) and the dispersion relation:

\[ \omega^2 = gk \tanh(kh) \]  

(2.6)

\[ v_g = \frac{\partial \omega}{\partial k} \]  

(2.7)

where \( \omega=2\pi f = \) angular frequency (radians per second), \( g = \) acceleration due to gravity (ms\(^{-2}\)), \( k = \) wavenumber (radians per m), \( h = \) water depth (m), \( v_g = \) group velocity (ms\(^{-1}\)).

Second, the group velocity grids \( v_g(f, l) \) and station locations were input into an Eikonal travel time solver (Rawlinson and Sambridge, 2004) which output travel times, \( t(f, l) \), from each station to each grid point at each frequency of interest.
Chapter 2 Infragravity waves off Pacific Northwest

(a) Stack 150s  
(b) Isotropic 150s  
(c) Weighted 150s

Figure 2.1: Backprojection of September 2012-May 2013 stack at 150 s. a) Backprojection of stack. b) Backprojection for isotropic source distribution. c) Weighted stack.

To obtain our backprojection $P(f_p, l)$ we sum the envelopes of the individual band-pass-filtered cross-correlations, multiplied by $W$, at their respective travel time for hypothetical source at $l$ (equation 2.4).

Backprojection was computed over a spatial grid of $35^\circ$N $< \varphi < 50^\circ$N, $135^\circ$W $< \gamma < 124^\circ$W at a spatial resolution of $0.0167^\circ$ (1 arc minute), and at frequencies of 0.0100, 0.0080, 0.0067, 0.0057 and 0.0050 Hz (100, 125, 150, 175, 200s). Daily cross-correlations became less clear at station separation distances below 50 km and above 120 km, so we only used station pairs within this range (shown in Appendix B Figure B.1). The output of the backprojection technique is a map for each frequency, as shown in Figure 2.1a.

2.3.2.1 Backprojection for isotropic source distribution

In order to be certain that the backprojection results were not an artefact of the array geometry, we calculated what the backprojection results would be if the array was subjected to an isotropic wavefield. Backprojection for an isotropic source distribution, $I(f_p, l)$, was computed from the theoretical cross-correlation function, $R(f, t)$, which has two symmetrical impulses at the positive and negative lags corresponding to the group arrival time of the ray path between the two stations for the given frequency. The theoretical isotropic cross-correlation for each station pair (band-pass-filtered, enveloped and normalised to a maximum of 1) can then be backprojected using equation 2.4 but replacing $C_n$ with $R_n$ and using $W=1$: 
\[ I(f_p, l) = \sum_{n=1}^{N} W_n(f_p) \text{env} \left( R_n \left( f_p, T_n(f, l) \right) \right) \] (2.8)

From this, we can see which locations the array would illuminate as sources even if all sources were equal. Figure 2.1b shows the isotropic response for a period of 150 s.

### 2.3.2.2 Backprojection weighting

The theoretical isotropic cross-correlations were also used to calculate a weighting coefficient \( W_n \) for each unique station pair \((n=1, 2, \ldots, N)\) for the backprojection using a least squares regression, seeking to minimise equation 2.9 below (Widrow et al., 1967; Applebaum and Chapman, 1976):

\[
\min_w \| G W - P \|_2^2, \quad w \geq 0 \tag{2.9}
\]

\( P_l = 1 \) for each unique latitude and longitude point on our map \((l=1, 2, \ldots, L)\). This characterises an ideal isotropic backprojection for all sources being equal. \( G_{ln} \) contains the isotropic backprojection of each station pair \( n \) (i.e. equation 2.8 before the summation over all station pairs). \( W_n \) contains the resulting weighting coefficient for each station pair:

\[
G_{ln} = \text{env} \left[ R_n \left( T_n(l) \right) \right] \tag{2.10}
\]

\[
P = \begin{bmatrix} p_1 \\ p_2 \\ \vdots \\ p_L \end{bmatrix} = \begin{bmatrix} 1 \\ 1 \\ \vdots \\ 1 \end{bmatrix}
\]

\[
W = \begin{bmatrix} w_1 \\ w_2 \\ \vdots \\ w_N \end{bmatrix}
\]

The backprojection can then be calculated with the solution \( W \) using equation 2.4 which we call a weighted backprojection. An example is given in Figure 2.1c.
2.3.3 Beamforming

Beamforming was another method used to estimate the direction of energy propagation across the array. It is similar to backprojection but gives beamformer output as a function of slowness (reciprocal of velocity) and azimuth rather than location, and identifies waves propagating across the array from outside sources.

At each frequency of interest, we generate our beamformer output, $B$, as function of angular frequency pass-band $\omega_p$, group slowness $s$, and plane wave back azimuth $\theta$, in the following way and given in equation 2.11.

$$B(\omega_p, s, \theta) = \sum_{n=1}^{N} \text{env} \left( C_n(\omega_p, T_n(s, \theta)) \right)$$  \hspace{1cm} (2.11)

$C_n$ is the observed cross-correlation for the station pair $n$ band-passed at a centre frequency $\omega$. We generate $T_n$, a synthetic plane wave group travel time for station pair $n$, at each slowness and back azimuth of interest. $T_n$ is calculated using:

$$T_n(s, \theta) = s \cdot d_n \cdot \cos(\theta - \phi_n)$$  \hspace{1cm} (2.12)

(in terms of interstation distance $d_n$ and azimuth $\phi$)

We sum the envelopes (env in equation 2.11, calculated using a Hilbert transform) of the individual band-pass-filtered cross-correlations with a centre frequency, $\omega$ at their respective synthetic travel time. We use a Gaussian band pass filter $\exp(-\alpha * (\omega - \omega_c)^2/\omega_c^2)$, where $\alpha = 100.0 * \sqrt{\delta}/1000$ where $\delta$ is interstation distance, $\omega$ is angular frequency and $\omega_c$ is the centre angular frequency. We examined from 0.05-0.167 Hz, 0 - 360° back azimuth, and from 20-100 s/km slowness. An example of beamformer output is shown in Figure 2.2a.

The maximum of the beamformer output identifies the slowness and azimuth of the dominant wavefield for each frequency of interest. However, we note the wavefield may be more complicated than a single plane wave for a given day, with a curved wavefront or multiple components. In the beamformer output, complications present themselves as very broad maxima in azimuth centred on a given slowness or multiple local maxima at the best fitting average array slowness.

The slowness is not so important for our purposes and only verifies that the beamforming is picking up waves that are traveling at the expected velocity for a given frequency and water depth. Our results focus on the azimuth of the dominant wavefield at each frequency of interest.
Chapter 2 Infragravity waves off Pacific Northwest

Figure 2.2: Beamforming of September 2012-May 2013 stack. a) Beamforming of stack at 146 s. Radial axis is slowness in s/km, angular axis is azimuth of wave arrival. b) Beamforming of isotropic source distribution of plane waves with period 146 s and slowness of 8.22 s/km. c) Beamform output of (a) at slowness of 8.22 s/km (which is the slowness of maximum beam) projected onto map of Pacific. Rays originate from the array centre.

Our beamforming method works best when there are no large changes in bathymetry (hence slowness) across the array, so we limited the beamforming to stations in deep water west of 126°W where changes in velocity across the array are small (maximum 10% between station pairs).

We also performed the beamforming on theoretical cross-correlations for an isotropic distribution of plane waves of a given frequency $R_n(\omega, t)$. $R_n(\omega, t)$ was generated as two symmetrical peaks of value 1 at $t = \pm(s \times d_n)$ and 0 at all other $t$. $s$ was obtained as the slowness of maximum power from our backprojection results for the given frequency. The theoretical isotropic cross-correlation $R_n$ for each station pair (band-pass filtered, enveloped and normalised to a maximum of 1) was then input into equation 2.11 in place of the observed cross-correlation $C_n$. The ideal result in this case would be an equal distribution of energy across all azimuths at the given slowness. Inevitably the array geometry results in some azimuths being more sensitive than others, but by comparing our results to the isotropic case, we can make sure our results are not an artefact of the array and the processing. The isotropic response at 146 s is shown in Figure 2.2b.

2.3.4 Effects of normalisation on directionality estimates

Since the amplitudes of the individual DPG’s are unknown, we normalised each record by the envelope of each trace prior to cross-correlation, effectively only retaining the
instantaneous phase information. This procedure will tend to amplify weaker sources and mute the strongest sources. Therefore, our results will tend to pick out the most coherent wavefield, not necessarily the largest so our directionality estimates are biased in this sense. However, the relative amplitudes on a given cross-correlation should accurately represent the relative amount of energy propagating in one direction versus another.

2.3.5 Calculating reflection coefficients

2.3.5.1 Seaward:shoreward wave propagation

For an offshore station pair aligned perpendicular to a coastline, each side of the cross-correlation function represents either shoreward propagating energy or seaward propagating energy. Therefore, the ratio of the positive and negative enveloped cross-correlation at \( \pm t_1 \) gives the ratio of seaward: shoreward energy propagation. In this case, the coastline lies approximately north-south and so east-west aligned station pairs (azimuth 265° - 275° or 85° - 95°, shown in Figure 1.1 in Chapter 1) were used in the analysis of seaward:shoreward infragravity wave propagation. Again, only station pairs separated by distances >30 km and <120 km were used as quality of the daily cross-correlation functions became much reduced outside this range. This limited the analysis to adjacent stations, and all remaining station pairs were between 60 and 75 km apart.

2.3.5.2 Assumption of no local infragravity wave generation

Shoreward propagating infragravity waves were considered to be remote arrivals, whilst seaward propagating infragravity waves may have been due to reflections at the coast or may have been locally generated ‘leaky’ infragravity waves. In practice it is difficult to separate the cause of the seaward propagating waves- but previous studies have found that infragravity wave generation is small when either nearshore \( H_s \) (Herbers et al., 1995b) or \( \alpha H_s T_{m0}^2 \sqrt{\frac{g}{D}} \) (Ardhuin et al., 2014) are small. Here we have used both measures, \( H_s \) and \( \alpha H_s T_{m0}^2 \sqrt{\frac{g}{D}} \), and assumed that when these were low, the amount of leaky infragravity waves was minimal. In this case, with local generation assumed to be zero, seaward propagation is reflected shoreward incident energy only (see Figure 2.3). This reflection may involve both specular reflection and scattering, as well as loss of energy through bottom friction and other processes, so the reflection coefficients we calculate contain the net effect of these processes.
Figure 2.3: Schematic of propagation and reflection of a simple impulse. a) A simple impulse (I) propagates shoreward towards the coastline at normal incidence, passes through Station 2 and Station 1, gets reflected, and propagates back seaward (R). \( t_1 \) is the travel time between Station 1 and Station 2. \( t_2 \) is the travel time between Station 1 and the coast. b) the time series observed at Station 2. c) The time series observed at Station 1. d) Cross-correlation of records at Station 1 and Station 2.

Since the amplitude of the reflected wave \( R \) equals the amplitude of the incident wave \( I \) multiplied by the reflection coefficient \( r \), i.e. \( R = Ir \):

\[
r^2 = \frac{R^2}{I^2} = \frac{\text{seaward}}{\text{shoreward}}
\]

(2.13)

and

\[
r = \sqrt{\frac{\text{seaward}}{\text{shoreward}}}
\]

(2.14)

Equations 2.13 and 2.14 apply to the case illustrated in Figure 2.3 for a single wave arrival and reflection. The peaks of observed cross-correlations result from waves propagating from multiple directions (Snieder, 2004), but the result still holds and has been shown formally by Wapenaar and Thorbecke (2013) and Godin et al. (2014).

This assumption of no local generation at low \( H_s \) or \( \alpha H_s T_{m0}^2 \sqrt{\frac{g}{D}} \) may not be perfect, so the reflection coefficients calculated should be considered an upper bound.
### 2.3.6 Cases of reflection

To calculate a reflection coefficient, the data were scanned for days when 1) nearshore $H_s$ or $\alpha H_s T^2 m_{0.2}$ were low (so that locally generated leaky infragravity waves could assumed to be negligible and 2) a strong arrival of infragravity energy from the west ($240^\circ$ to $300^\circ$) was observed (so that the arrival and reflection are both observable on an east-west aligned station pair). The three days that best matched these conditions were 17 January 2013, 18 January 2013 and 11 May 2013. These days were analysed along with another two days, 25 October 2012 and 4 April 2013, when arrivals from the west were present but less clear, and $H_s$ and $\alpha H_s T^2 m_{0.2}$ were (mostly) higher.

For each of the five days, the reflection coefficient was calculated from each east-west station pair using equation 2.14. Values for seaward and shoreward infragravity energy were taken as the peak in the enveloped cross-correlation function at around $\pm t_1$. To account for a wave arrival slightly oblique ($\pm 30^\circ$) to the station pair alignment ($270^\circ$) and velocity errors/scattering, which would result in a peak slightly off $\pm t_1$, the maximum value in a 200 s window around $\pm t_1$ was used as the peak. Varying this window between 100 and 300 s made very little difference to the results.

### 2.4 Results

#### 2.4.1 Infragravity wave energy and directionality

Figure 1.3 in Chapter 1 shows the stacked cross-correlations filtered between 60 and 500 seconds, and Figure 1.4 in Chapter 1 shows examples from east-west aligned station pairs filtered to 100, 150 and 200 seconds. Peaks at $\pm$ the theoretical travel time $t_1$ of an infragravity wave group between the two stations confirms that over the year there is coherent infragravity wave energy propagating both seaward and shoreward, with more going seaward.

From backprojection and beamforming, the dominant source of infragravity energy was found to be from the coastline to the east/northeast, consistent with local generation, as seen in Figures 2.1c and 2.2c (plots for other periods are similar can be seen in Appendix B Figures B.2 and B.3). Backprojection (Figure 2.1c) highlighted the stretch of coastline between $40^\circ$N and $44^\circ$N as the dominant source, whilst beamforming (Figure 2.2c) identified the region between $42^\circ$N and $46^\circ$N as the main source. The differences are small but may be due to the fact that backprojection used additional stations nearer the coast and was weighted to remove effects to array geometry. There was however a notable change in direction with time of year (Figure
Chapter 2 Infragravity waves off Pacific Northwest

2.4). The source to the east/northeast dominated from mid-September through to March, but in April and continuing through May the dominant source shifted to the south, coinciding with the switch to austral winter and perhaps indicating a switch to a remote infragravity wave source.

The change in source direction to the south was accompanied by a reduction in power at infragravity wave frequencies (Figure 2.5a). Power showed a positive correlation with local short-wave $H_s$ and $\alpha H_s T_{m0, -2}^2 \sqrt{\frac{g}{D}}$ (Figure 2.5b-e), indicating the importance of leaky infragravity waves generated at the local coastline to infragravity power offshore in this part of the ocean. Scatter in the relationship could be partly due to remote arrivals. For example, power on 17 January 2013, 18 January 2013 and 11 May 2013 (marked by the black crosses in Figure 2.5d and e) was higher than expected for the local $H_s$ or $\alpha H_s T_{m0, -2}^2 \sqrt{\frac{g}{D}}$, probably due to the infragravity arrival observed from the west.

2.4.2 Reflection of infragravity waves

The infragravity seaward:shoreward ratio was calculated for each daily cross-correlation in the record. Throughout most of the record, the infragravity seaward:shoreward ratio was $>1$ (i.e. offshore propagation). A ratio $>1$ cannot be explained by incident arrivals and their reflections, but again indicates that for most days of the record the deep water infragravity waves were generated at the local coastline to the east. As with infragravity power (Figure 2.5d and e) the seaward:shoreward ratio increased as significant wave height $H_s$ and $\alpha H_s T_{m0, -2}^2 \sqrt{\frac{g}{D}}$ increased at the coast (Figure 2.6a and b). Again, scatter in the relationship may be caused by the shoreward propagating infragravity waves (which have no reason to correlate with $H_s$ or $\alpha H_s T_{m0, -2}^2 \sqrt{\frac{g}{D}}$) as
Figure 2.5: Infragravity power vs. significant wave height and $\alpha H_s T_{m0,-2}^2 \sqrt{g/D}$. a) Infragravity (IG) power at 171 seconds at stations G21B, G30B, G02B and M14B. Each time series has been normalised by its maximum value. b) Daily significant wave height $H_s$ (m) at Buoy 46015, 46022 and 46027. c) $\alpha H_s T_{m0,-2}^2 \sqrt{g/D}$ at Buoy 46015, 46022 and 46027. d) Mean $H_s$ vs IG power at G30B. Grey points show daily $H_s$ (averaged over the three buoys) vs IG power. The daily data points were binned into eight bins according to $H_s$ with centres of 1, 2, 3, 4, 5, 6, 7 and 8 m, and the red points mark the mean of each bin. Error bars have a length of 2 standard errors (1 s.e. positive and 1 s.e. negative), and the corresponding numbers give the number of data points in that bin. 17 January 2013, 18 January 2013 and 11 May 2013 are marked by the black crosses. e) is same as d) but using $\alpha H_s T_{m0,-2}^2 \sqrt{g/D}$ and bin centres of 0.010, 0.020, 0.030, 0.040, 0.050, 0.075, 0.100 and 0.150 m.
Figure 2.6: Same as Figure 2.5d and e but $H_s$ and $\alpha H_s T_{m0, -2}^2 \sqrt{g/D}$ vs. daily seaward:shoreward energy. Daily seaward:shoreward energy was calculated as the mean from the 7 east-west station pairs over all periods 100-200 s. A linear regression was calculated through the first four (for $H_s$) or five (for $\alpha H_s T_{m0, -2}^2 \sqrt{g/D}$) binned points and the equation of best fit, $R^2$ value and $p$-value are shown.

well as the likely imperfect relationship between $H_s$ or $\alpha H_s T_{m0, -2}^2 \sqrt{g/D}$ and seaward propagating waves. With the data binned according to $H_s$ or $\alpha H_s T_{m0, -2}^2 \sqrt{g/D}$, the correlation with seaward:shoreward energy became much clearer (Figure 2.6a and b red circles). At $H_s$ of about 4 m and below, or $\alpha H_s T_{m0, -2}^2 \sqrt{g/D}$ of about 0.05 m and below, the relationship between $H_s$ and $\alpha H_s T_{m0, -2}^2 \sqrt{g/D}$ and the seaward:shoreward ratio was approximately linear. For $H_s$, a linear regression calculated from the first four binned data points (which all had standard error of $\leq 0.14$) had a slope of $0.69 \pm 0.04$ m$^{-1}$ and intercept of $0.21 \pm 0.10$. For $\alpha H_s T_{m0, -2}^2 \sqrt{g/D}$, the first five binned data points all had standard error of $\leq 0.10$ and had a slope of $33.27 \pm 1.48$ m$^{-1}$ and intercept of $0.55 \pm 0.04$. The square root of the intercept is an estimate of the reflection coefficient if the seaward:shoreward ratio is linear, therefore giving an estimate of $r \approx 0.49 \pm 0.29$ for $H_s$ or $r \approx 0.74 \pm 0.06$ for $\alpha H_s T_{m0, -2}^2 \sqrt{g/D}$. At higher $H_s$ or $\alpha H_s T_{m0, -2}^2 \sqrt{g/D}$, the data suggested a potential levelling off of seaward-propagating (leaky) infragravity waves, but the number of data points at these higher values was limited.

Figure 2.7 shows one of the examples when $H_s$ and $\alpha H_s T_{m0, -2}^2 \sqrt{g/D}$ were low (Table 2.1), and a strong infragravity arrival from the west was observed. The reflection coefficient was calculated from each of the seven east-west aligned station pairs on this day, at periods of 100, 125, 150, 175 and 200 s. Some examples of cross-correlations on this day, from which the reflection coefficient was calculated, are shown in Figure 2.8. The reflection coefficient calculated varied between the different station pairs. Each cross in Figure 2.9a shows the reflection coefficient calculated for each station pair at
Chapter 2 Infragravity waves off Pacific Northwest

Figure 2.7: \( H_s \) and \( \alpha H_s T_{m0, -2}^2 \sqrt{2/\sigma} \) during January 2013, and backprojection on 18 January 2013. a) \( H_s \) and b) \( \alpha H_s T_{m0, -2}^2 \sqrt{2/\sigma} \) at Buoys 46015, 46022 and 46027 during January 2013. 18 January is bounded by the dashed lines. Backprojection on 18 January at c)100 s, d)150 s and e)200 s shows the westerly infragravity wave arrival.

The mean varied between 0.83 at 100 s and 0.51 at 200 s, with a mean over all periods and station pairs of 0.66. A visual inspection of the cross-correlation functions found that for some station pairs the two peaks were not evident at all, and so a calculation based on the values at these two peaks was essentially meaningless (see Appendix B Figure B.5) for some examples of unclear correlations). For this reason, any unclear cross-correlations were discarded, and the results were replotted (Appendix B Figure B.6a). In this case the mean varied between 0.90 at 100 s to 0.59 at 200 s, with a mean over all periods of 0.71 (the mean at each period is plotted in Figure 2.10). It is understood that this approach risks a bias towards larger reflection coefficients, although an inspection of the differences found that overall both anomalous low and high values were removed, and the remaining data points became less scattered.

The reflection coefficient was calculated in the same way for the other four dates (25 October 2012, 17 January 2013, 4 April 2013 and 11 May 2013) when \( H_s < 2 \) m (and \( \alpha H_s T_{m0, -2}^2 \sqrt{2/\sigma} < 0.03 \) m) and westerly arrivals were observed (Figure 2.9b-e and Figure 2.10). The number of stations considered to have clear cross-correlations on these other four days were generally less than on 18 January 2013 (numbers along
<table>
<thead>
<tr>
<th>Date</th>
<th>$H_s$ (m)</th>
<th>Average Wave Period (s)</th>
<th>$\alpha H_s T_{m0,-2} \sqrt{\frac{g}{D}}$</th>
<th>Dominant Wave Period (s)</th>
<th>Dominant Wave Direction (°)</th>
</tr>
</thead>
<tbody>
<tr>
<td>4-Apr-2013</td>
<td>1.83</td>
<td>6.11</td>
<td>0.019</td>
<td>8.37</td>
<td>238</td>
</tr>
<tr>
<td>25-Oct-2012</td>
<td>1.63</td>
<td>6.55</td>
<td>0.021</td>
<td>8.84</td>
<td>301</td>
</tr>
<tr>
<td>17-Jan-2013</td>
<td>1.36</td>
<td>6.50</td>
<td>0.018</td>
<td>12.90</td>
<td>286</td>
</tr>
<tr>
<td>18-Jan-2013</td>
<td>1.06</td>
<td>8.82</td>
<td>0.026</td>
<td>12.90</td>
<td>288</td>
</tr>
<tr>
<td>11-May-2013</td>
<td>0.58</td>
<td>6.88</td>
<td>0.009</td>
<td>14.81</td>
<td>213</td>
</tr>
</tbody>
</table>

*aAverage of NDBC wave buoys 46015, 46022 and 46027. Spectra are shown in Appendix B Figure B.7.*
Figure 2.9: Values of $r$ ($\sqrt{IG$ seaward:shoreward$}$) obtained from east-west station pairs on (a) 18 January 2013 (b) 25 October 2012 (c) 4 April 2013 (d) 17 January 2013 (e) 11 May 2013 as well as the mean $r$ on each day (f). The red points in (a-e) mark the mean value of $r$ from the seven east-west station pairs at each period, with error bars of 2 standard deviations (1 s.d. positive and 1 s.d. negative). The number of station pairs used in the calculation of the mean are given along the bottom of the plot. N/A means no estimate of $r$ was made because back projection showed no clear arrival from the west for this period. The mean value of $r$ over all periods is given by the number in the top right corner of each plot. (f) The mean values at each period from plots a-e. The axes for all plots are equal to those given for plot (a).
bottom of plots in Appendix B Figure B.6). In particular, the arrivals on 25 October 2012 and 4 April 2013 were much less clear and not over all periods (see Appendix B Figure B.4 for backprojection plots). Therefore we focus on the ‘best cases’ of 17 January 2013, 18 January 2013 and 11 May 2013.

The mean reflection coefficient at each period from each of the five days (plots a-e of Figure 2.9) are plotted together in Figure 2.9f, and the means from plots a-e of Appendix B Figure B.6 are plotted together in Figure 2.10. The means were closely matched, especially for 18 January 2013 and 11 May 2013, in Figure 2.10. It is difficult to explain why the values obtained for the 17 January 2013 are the lowest obtained despite not having the lowest $H_s$ or $\alpha H_s T_{m0}^2 - 2 \sqrt{\frac{g}{D}}$ (Table 2.1). The mean reflection coefficient obtained over the three best cases of 17 January 2013, 18 January 2013 and 11 May 2013 is given in Table 2.2, for both the calculations where all seven station pairs’ cross-correlations were used (Figure 2.9) and for the calculation using only selected clear cross-correlations (Figure 2.10). The means were $0.65 \pm 0.02$ and $0.66 \pm 0.02$ respectively. These estimates lie within the estimates of $r$ from the linear regressions of seaward:shoreward energy with $H_s$ and $\alpha H_s T_{m0}^2 - 2 \sqrt{\frac{g}{D}}$, which are also shown in Table 2.2.
### Table 2.2: Estimates of reflection coefficient

<table>
<thead>
<tr>
<th>Reflection Coefficient, r</th>
<th>Standard Error</th>
<th>Standard Deviation</th>
</tr>
</thead>
<tbody>
<tr>
<td>All cross correlations from 3 best cases&lt;sup&gt;a&lt;/sup&gt;</td>
<td>0.65&lt;sup&gt;b&lt;/sup&gt;</td>
<td>0.02&lt;sup&gt;c&lt;/sup&gt;</td>
</tr>
<tr>
<td>Subselected cross correlations from 3 best cases</td>
<td>0.66</td>
<td>0.02</td>
</tr>
<tr>
<td>Regression estimate from $H_s$</td>
<td>0.49&lt;sup&gt;d&lt;/sup&gt;</td>
<td>0.29</td>
</tr>
<tr>
<td>Regression estimate from $\alpha H_s T_{m0}^{2} \sqrt{D}$</td>
<td>0.74&lt;sup&gt;d&lt;/sup&gt;</td>
<td>0.06</td>
</tr>
</tbody>
</table>

<sup>a</sup>Best cases are 17 January 2013, 18 January 2013 and 11 May 2013

<sup>b</sup>Mean reflection coefficient over east-west stations pairs at periods 100-200s.

<sup>c</sup>Standard deviation/$\sqrt{n}$ where $n$ is number of cross-correlations used in the calculation of mean $r$.

<sup>d</sup>Square-root of intercept shown in Figure 2.5.

### 2.5 Discussion

The dominant infragravity wave sources found here are in agreement with previous studies. The northwest coast of America has consistently showed up as a source of infragravity waves (Webb et al., 1991; Rhie and Romanowicz, 2006; Aucan and Ardhuin, 2013; Rawat et al., 2014). In general, the eastern boundaries of basins provide stronger sources of infragravity waves than western boundaries due to larger wave heights and wave periods incident on these coasts (Rawat et al., 2014). However, a recent study by Crawford et al. (2015) on deep ocean infragravity waves in the Atlantic has shown that other factors such as short-wave incidence angle or spread and coastal morphology may be more important for infragravity wave generation than wave heights and periods. This study has shown some rare examples where infragravity waves arrived from the western side of the Pacific.

The results suggest that infragravity waves that propagated west to east across the Pacific Ocean reflected at the North American Pacific coastline strongly. Perhaps the main limitation of the method used is that is problematic to distinguish between reflected infragravity waves and locally generated infragravity waves. Here the reflection coefficient was estimated by assuming that there were no locally generated infragravity waves on the days investigated because nearshore significant wave height $H_s$ or the parameter $\alpha H_s T_{m0}^{2} \sqrt{D}$ was low. An alternative method which extrapolated the observed trend of the seaward:shoreward ratio with $H_s$ and $\alpha H_s T_{m0}^{2} \sqrt{D}$ to zero complemented the results obtained on the individual days investigated.

One approach to tackling the problem of locally generated infragravity waves might be to use calibrated pressure gauges (only uncalibrated gauges were available in this study). Then, the expected seaward-propagating infragravity wave energy for a particular sea state could be estimated from an empirical model between sea state and infragravity wave height, for example such as that described in Ardhuin et al. (2014), and then removed from the observed seaward propagating energy. This would have
other problems associated with it such as the accuracy of the model for particular coastlines and sea states. However, it might offer some more insight as to the magnitude of the reflection, and allow estimates of reflection to be made on days when wave activity at the coast is larger so local generation cannot be assumed to be zero. Another approach could be to use modelled infragravity energy for the whole ocean basin (Ardhuin et al., 2014) and see if there is a mismatch in modelled and observed energy on days when reflection is expected to be large.

Another way to distinguish reflection from local generation— that should be possible with the method used here— is by identifying a peak in energy that arrives at a greater lag than the main arrival ($t_1 + 2t_2$ peak labelled $IR$ in Figure 2.3d). This peak has not been clearly identified in this dataset. This might be because the schematic in Figure 2.3 is too simple for the real complicated wavefield. The complexity of the wavefield can be seen by the early arrivals and multiple late arrivals in the cross-correlations (Figures 1.4 in Chapter 1 and 2.8), indicating waves propagating obliquely to normal incidence and multiple reflections respectively. The IR peak would not be clear or symmetrical if reflection occurred from a source unaligned with the station pair, from multiple sources, or occurred gradually over a large geographical area (such as the continental shelf). The beamforming output does show a late arrival in the direction that would be expected from a reflection, but the isotropic responses in Figure 2.2 suggest that this is due to aliasing by the array geometry. However, a good example of this reflection peak can be seen in the data of Harmon et al. (2012) for pressure gauges southwest of Sumatra, which were dominated by remote arrivals rather than local generation. The stacked cross-correlation between their stations 42 and 45 shows symmetrical secondary peaks at -2000 and +2000 second lags (Appendix B Figure B.8a). A quick calculation of the reflection coefficient from the ratio of the IR peak to the $I^2$ peak gives $r$ of approximately 0.3, which is much lower than that found here, but is not based only on arrivals at normal incidence to the coast. This result may indicate that reflection coefficients depend strongly on the bathymetry and configuration of coastlines. Indeed, if all coastlines reflected at approximately 0.7, then infragravity energy in the middle of the ocean would be substantially higher than a case with no reflection. However Webb et al. (1991) found little infragravity waves originating from large parts of the Pacific, which led them to suggest that reflection from coastlines is small. This perhaps suggests that the North American Pacific coastline is an example of a particularly strong reflector for infragravity waves reflecting at normal incidence to the coast (as indeed it is also a strong source (Ardhuin et al., 2014)).

The lag time of this late arrival (IR peak at $t_1 + 2t_2$) can also offer information on where the reflection occurs— whether at the continental slope or at the shoreline, or a combination of both. A backprojection of the Sumatra lag onto a $t_1 + 2t_2$ travel-time grid (i.e. $T_n$ in equation 2.4 = $(t_i(f,l) - t_j(f,l)) + 2 \cdot t_j(f,l)$) puts the reflection off the
island of Batu at depths of 25-100 m (Appendix B Figure B.8b and c). There are no other studies (that the authors are aware of) that estimate reflection of transoceanic propagating infragravity waves in deep water, and therefore their potential reflection from a continental shelf, but previous studies of infragravity waves in shallow water have calculated reflection at the shoreline. These studies found reflection at the shoreline to be similar or higher to those found here: \( r^2 = 0.6 \pm 0.11 \) and \( r^2 = 0.65 \pm 0.25 \) (corresponding to \( r=0.77 \) and \( r=0.81 \)) for two sites in De Bakker et al. (2014), or \( r^2 = 0.8 \) (\( r= 0.89 \)) in Sheremet (2002). If reflection did occur at the shoreline rather than at the shelf, the ‘reflection coefficient’ obtained in deep water would not strictly be only a reflection coefficient, but would contain the effects of dissipation on two crossings of the shelf. Considering dissipation can occur by bottom friction, triad interactions or breaking (De Bakker et al., 2014), and any reflected waves could become trapped, the reflection coefficients obtained in this study would seem quite large, although perhaps explained by a narrow shelf. On the other hand, the continental shelf is fairly linear (north-south) and it could be because of this that we observe a very coherent reflected wavefield, more than what would be expected for a more irregular shoreline. A simple calculation for long wave reflection at a step, \( r = (1 - \sqrt{h_1/h_0})/(1 + \sqrt{h_1/h_0}) \) from Lamb (1932), where we use \( h_1=100 \) m and \( h_0=3000 \) m based on the bathymetry of our study region, gives a value of \( r=0.69 \), which is similar to our estimates. Of course, this oversimplifies the problem as the bathymetry is more complex than a simple step, so the agreement may be merely fortuitous. As infragravity waves refract towards normal incidence as they cross the shelf (Herbers et al., 1995b), reflection from the shoreline should depend less on the angle of incidence of the remote arrival than reflection at the continental shelf, and this may be another way to distinguish the location of the reflector.

### 2.6 Conclusions

The North American Pacific coastline was found to be the dominant source of infragravity waves observed offshore, although energy and source were found to change seasonally. During northern-hemisphere winter, infragravity wave energy was higher and the waves mostly originated from the nearby coastline to the east/northeast, whilst during northern-hemisphere summer, energy decreased and arrivals mostly came from the south. The seasonal pattern can be explained by the relationship between infragravity waves and short-period (2-30 s) wave activity. Infragravity wave energy in the deep ocean (i.e. leaky free infragravity waves) increased with short-period wave activity at the coast (indicated by \( H_s \) and \( \alpha H_s T_{m0}^{-2} \sqrt{g/\rho} \)), resulting in higher infragravity wave energy during winter when short-period wave activity is highest.
Remote arrivals coming from the west, propagating eastward, were observed but rare. The strength of reflection of these remote arrivals was estimated using the asymmetry of cross-correlation functions calculated between station pairs perpendicular to the coastline. Whilst the method is limited to the assumption of no local infragravity wave generation when short-period wave activity is low, reflection did appear to be strong, with a lower bound estimate of $r=0.49\pm0.29$ (reflection coefficient ± standard error) and an upper bound estimate of $r=0.74\pm0.06$ for this particular coastline. These results indicate that reflection has the potential to be an important factor to account for infragravity wave energy in the deep ocean.
Chapter 3

Monitoring remote ocean waves using P-wave microseisms

This chapter has been published as ‘Neale, J., Harmon, N. and Srokosz, M. (2017), Monitoring remote ocean waves using P-wave microseisms, Journal of Geophysical Research: Oceans, 122, 470-483, doi:10.1002/2016JC012183’ and is included in Appendix B.2. The co-authors listed in this publication directed and supervised the research that forms the basis for the thesis.

3.1 Abstract

Oceanic microseisms are generated by the interaction of opposing ocean waves and subsequent coupling with the seabed, so microseisms should contain information on the ocean conditions that generated them. This leads to the possibility of using seismic records as a proxy for the ocean gravity wavefield. Here we investigate the P-wave component of microseisms, which has previously been linked to areas of high wave interaction intensity in mid-ocean regions. We compare modelled P-wave microseismic sources with those observed at an array in California, and also investigate the relationship between observed sources and significant wave height. We found that the time-varying location of microseism sources in the North Pacific, mapped from beamforming and backprojection of seismic data, was accurate to $\leq 10^\circ$ in 90% of cases. The modelled sources were found to dominate at $\sim 0.2$ Hz which was also reflected in the seismic observations. An empirical relationship between observed beampower and modelled source power allowed sources during an independent data period to be estimated with a correlation coefficient of 0.63. Likewise, significant wave height was also estimated with a correlation coefficient of 0.63. Our findings suggest that with improvements in resolution and amplitude retrieval from beamforming,
correlations up to 0.78 should be possible between observed P-wave microseisms and significant wave height in remote ocean regions.

3.2 Introduction

Oceanic microseisms are tiny, continuous oscillations of the ground caused by the interaction of ocean waves with the solid earth beneath them. The most energetic microseisms are generated when ocean wave trains of similar frequency but opposite direction interact, producing a pressure fluctuation that is unattenuated with depth, has twice the frequency of the forcing waves, and a near-zero wave number (Longuet-Higgins, 1950; Ardhuin and Herbers, 2013; Traer and Gerstoft, 2014; Ardhuin et al., 2015). This pressure fluctuation couples to the seabed to produce microseisms with typical peak frequencies of about 0.14-0.20 Hz (5-7 seconds), which propagate as seismic surface waves and body waves (e.g. Toksöz and Lacoss (1968); Haubrich and McCamy (1969); Roux et al. (2005); Gerstoft et al. (2006); Koper and de Foy (2008); Koper et al. (2010); Reading et al. (2014)).

The requirement of opposing waves to generate double frequency microseisms means that microseism energy cannot be explicitly related to ocean wave height (Kedar et al., 2008; Ardhuin et al., 2011), as it depends on the directional characteristics of the wave trains as well as the wave energy. However, in cases where the opposing wave trains are related, as a result of directional spreading of one wave system (classified as Class I in Ardhuin et al. (2011, 2012)) or reflection from a coastline (Class II), a strong dependence on the ocean wave height is expected. As such, empirical relationships between microseism energy and significant wave height recorded at nearby coastal wave buoys have been identified (Ardhuin et al., 2012; Bromirski et al., 1999; Ferretti et al., 2013). These relationships were based on the full seismic spectrum which is dominated by seismic surface waves, and work well for seismic stations that are most sensitive to local sea states or coastal reflection sources.

This study investigates the P-wave component of seismic noise, for which origins in the deep ocean have been inferred (Toksöz and Lacoss, 1968; Haubrich and McCamy, 1969; Landès et al., 2010; Gerstoft et al., 2008; Zhang et al., 2010a; Euler et al., 2014; Koper et al., 2010). Here Class I situations do occur, but the opposing wavefield could also result from the interaction of two independent wave systems: either the crossing of two swells, or the interaction of swell with locally generated wind waves (Class III in Ardhuin et al. (2011, 2012)). Given that two low-energy wave systems directly opposing could produce the same microseismic energy as two high-energy wave systems with more oblique interaction, the correspondence between microseism energy and wave heights is expected to be weaker. Over long averaging times, sources of P-wave
microseisms do coincide with regions of high wave heights (Euler et al., 2014), but it is unclear what microseisms can tell us about ocean wave energy on shorter time periods when the degree of opposing wavefield is unknown and likely to be highly variable. This has implications for using P-wave microseisms as a proxy for wave heights.

In this study we investigate the relationship between P-wave microseisms and significant wave height at short (3-hour) time periods. Significant wave height is a common and useful parameter to describe sea state, but measurements by wave-buoys far from the coast are very sparse, and coverage by satellites insufficient. Locating P-wave microseisms to these deep ocean regions, and being able to infer wave heights, would allow for the development of real-time monitoring of sea state from seismic records, adding a valuable new record to our current measurements of significant wave height in the deep ocean.

We first compare observed P-wave microseisms with modelled microseism sources in terms of location, frequency and amplitude, and derive an empirical relationship between the observed and modelled source. Then, we reconstruct the ocean wave energy spectrum from the estimated microseism source spectrum by applying assumptions about the degree of wavefield opposition. Finally, the significant wave height calculated from the estimated ocean wave energy spectrum is then compared with modelled significant wave height.

3.3 Data and processing

Two years of vertical component seismic data (channel LHZ) from an array in California (Figure 3.1), were downloaded from the Southern California Earthquake Data Center (SCEDC, 2013) covering the period September 2012-September 2014. The Southern California Seismic Network (SCSN) that makes up the bulk of the array has been used previously to successfully locate P-wave microseisms (Gerstoft et al., 2006, 2008; Zhang et al., 2010a,b; Obrebski et al., 2013). Instrument response was removed, and each daily record was band-pass filtered between 0.002 and 0.400 Hz. Earthquakes were removed using an automated detection and removal algorithm. Earthquake events were identified using the ISC bulletin (International Seismological Centre, 2013) and removed by setting 1 hour of the waveform to zero if the RMS of that window was over 3 times the daily RMS. Daily spectra for each station were then individually examined and any bad quality data were discarded.

A similar beamforming procedure to Gerstoft and Tanimoto (2007) and Gerstoft et al. (2008) was used to examine microseisms as a function of frequency, azimuth and slowness. Each daily time series was split into 512 s chunks which were Fourier transformed to give a complex valued vector $v(f_2, t_i)$ containing the response from all
Figure 3.1: The seismic networks used in the study. White triangles=Berkeley Digital Seismograph Network, black triangles=Southern California Seismic Network, light grey triangles=ANZA regional network, white triangle with black border=USArray Transportable Array.

stations in the array, where $f_2 = 2f$ is the seismic frequency in Hz (twice the ocean wave frequency $f$) and $t_i$ refers to the start time of the Fourier transform. Only phase was retained by dividing $v$ by its magnitude to reduce the influence of local site amplification effects (Gerstoft and Tanimoto, 2007). The cross-spectral density matrix, $C = \langle vv^H \rangle$, where the brackets indicate temporal averaging over a 3-hour period and $H$ denotes the complex conjugate transpose, was calculated. The plane-wave response of the array is given by $p(f_2, s, \theta) = \exp[-i2\pi f_2 s(r_e)]$ where $e = (\sin\theta, \cos\theta)^T$ is the directional cosines for a plane wave with given azimuth $\theta$, $r$ is the coordinates of the seismometers with respect to their mean and $s$ is slowness. The beamformer output is then $b(f_2, s, \theta, t_s) = p(f_2, s, \theta)^H C(f_2, t_s) p(f_2, s, \theta)$ for each 3-hour snapshot $t_s$ which we adjusted for the number of stations $N$ in the array by dividing by $N^2$ (Euler et al., 2014). Array response functions at frequencies of 0.1 and 0.3 Hz are shown in Figure 3.2a and b.

For each 3-hour time period we therefore have beamformer output as a function of frequency, slowness and azimuth. An examination of the full 2 years of data revealed a large contrast between the strongest arrivals during winter and summer (Appendix C Figure C.1). During winter months, the strongest arrivals came from the North Pacific and North Atlantic with slownesses corresponding to P or PP phases, whereas during summer months the strongest arrivals came from southerly azimuths with slownesses of PKP phases. Gerstoft et al. (2008) previously found P-waves from the Pacific to dominate at the SCSN network during winter, so here we limit our analysis to P-waves.
from the North Pacific from mid-October to mid-March. Additionally we reject any
days when the number of stations falls below 170 to minimise errors and biases
associated with too few stations. At each frequency of interest, we back-projected
azimuth and slowness onto a 2° latitude by 2° longitude geographical grid assuming
slownesses of direct P-waves with a source at the Earth’s surface using the ak135
travel time tables of Kennett et al. (1995) for a spherically symmetric Earth model.
We projected for distances between 30° and 90° from the array centre which is the
typical range for teleseismic P-waves (Obrebski et al., 2013). A synthetic test for a
point source located at 35°N 169°E is shown in Figure 3.2c and d. An example output
for 23 December 2012 00:00-03:00 is shown in Figure 3.3a.

Double frequency P-wave microseism sources were modelled over the same 2-year
period as the seismic data. The method of Ardhuin et al. (2011) and Ardhuin and
Herbers (2013) based on the numerical ocean wave model WAVEWATCH III (Tolman,
2014) was followed to calculate the vertical ground displacement associated with
P-waves at each source location. From Ardhuin et al. (2011) and Farra et al. (2016),
the second-order pressure spectrum at near-zero wavenumber and twice the ocean
Figure 3.3: Example output for 23 December 2012 00:00-03:00. a) Beamformer output integrated between 0.1 and 0.3 Hz. b) Modelled source integrated between 0.1 and 0.3 Hz. c) Watershed regions identified from the beamformer image in (a).
wave frequency $f$, due to the interaction of similar frequency waves traveling in opposite directions (Hasselmann, 1963), is given by:

$$F_P(x, f) = [2\pi]^2 [\rho_w g]^2 f_2 E^2(x,f) I(x,f)$$

which has units of Pa²m²s and where $\rho_w$ is the density of water, $g$ is gravitational acceleration, $E(x,f)$ is the ocean wave frequency spectrum at location $x$ and $I(x,f)$ is a non-dimensional function that depends on the wave energy distribution $M$ over the directions $\theta$:

$$I(x,f) = \int_0^\pi M(x,f,\theta) M(x,f,\theta + \pi) d\theta$$

WAVEWATCH III was forced with 6-hourly winds and sea-ice cover from ECMWF’s ERA-interim reanalysis (Dee et al., 2011) and the second order pressure spectrum was output hourly on a 0.5° longitude by 0.5° latitude global grid from 78° North to 78° South. The output was smoothed using a 7-by-7 Gaussian low-pass spatial filter and resampled to a 2°-by-2° grid and averaged over 3 h time periods for direct comparison with the beamforming. Coastal wave reflection, which is function of wave amplitude, wave frequency and beach slope, can significantly affect the value of $F_P$ near the coast. Because we are mainly interested in deep-water events where reflection has less influence, we chose the reflection coefficient simply as $R^2=0.1$ for continents and large islands and 0.2 for small islands (Ardhuin et al., 2011). A second model run over a two-month period with lower reflection coefficients of $R^2=0.02$ for continents and large islands and 0.04 for small islands, which more closely correspond to the estimates of Stutzmann et al. (2012), did not produce significantly different values of $F_P$. The map of P-wave sources (Pa²m²s) is calculated by multiplying the second order pressure spectrum at each grid point by the squared source site effect $[2|C_P|\rho_c/(\rho_w)]^2$ (Farra et al., 2016):

$$P(x, f_2) = F_p(x, f_2) \times [2|C_P(x, f_2)|\rho_c/(\rho_w)]^2$$

$$= [2\pi]^2 [\rho_w g]^2 f_2 E^2(x,f) I(x,f) \times [2|C_P(x, f_2)|\rho_c/(\rho_w)]^2$$

$C_P$ is a non-dimensional amplification coefficient dependent on frequency, water-depth and P-wave take-off angle (distance from source to receiver) (Ardhuin and Herbers, 2013; Gualtieri et al., 2014). $C_P$ was calculated using the water depth and take-off angle appropriate for each grid point using the formulation of Gualtieri et al. (2014),
taking water-depth from the ETOP01 dataset (Amante and Eakins, 2009) and using the ObsPy package (Beyreuther et al., 2010) to calculate take-off angle based on the distance from the array centre. Values for spatially-varying crustal density $\rho_c$ and water density $\rho_w$ were taken from the upper crust layer and water layer of the global crustal model CRUST1.0 (http://igppweb.ucsd.edu/~gabi/rem.html). An example of the modelled source on 23 December 2012 00:00-03:00 is shown in Figure 3.3b. A study by (Obrebski et al., 2013) has previously validated the location of modelled sources with seismic P-wave observations.

We also output the modelled ocean wave energy spectrum $E(f)$ from which we calculated significant wave height, $H_s$, as observations of $H_s$ (e.g. from wave buoys or satellites) are not available at the temporal resolution or spatial extent required for comparison with the seismic data. Again we smoothed and resampled the WAVEWATCH III output of $E(f)$ to a 2°-by-2° grid and averaged over 3-hour time periods. Note that for convenience we now leave out the explicit notation for dependence of $E$ and $I$ on location $x$.

### 3.4 Results

#### 3.4.1 Comparison between observed and modelled source

For each 3-hour time period over the first winter (mid-October 2012 to mid-March 2013) the beamformer output and modelled source were integrated over the main double-frequency microseism band between 0.1 and 0.3 Hz. In order to identify multiple peaks in the beamformer maps, caused by more than one source acting in the basin at a given time, we used a standard watershed algorithm that identifies peaks in the image and the regions associated with each of these peaks (see Figure 3.3c). For each beamformer peak identified we found the largest modelled source within its watershed region to examine how the beampower varies with source power. By associating the beampower peak to a source located anywhere within its watershed region, rather than using a direct pixel comparison, we allow for errors in the beamformer location which result in it being offset from the true source location. We considered beamform peaks in the North Pacific (above 0° N).

The plot of modelled source power vs. beampower in Figure 3.4 shows that up to a certain limit, an increase in modelled source power had little to no effect on the observed beampower, but beyond this limit beampower increased rapidly with source power, although with a lot of scatter. The correlation coefficient calculated using Spearman’s Rank was 0.71 (strong correlation, >99% significance). After binning the data (shown in red) we defined the noise level as the mean of the first bin plus 2
standard deviations, which corresponds to a beampower of -28.68 dB. The binned data was found to be best described by an exponential function (shown in green), which indicated that the beampower only exceeded the noise level when source power \((F_P|2|C_P|\frac{\rho_c}{\rho_w}|^2)\) reached approximately 102 dB \((1.6\times10^{10} \text{ Pa}^2\text{m}^2)\).

Next we calculated the distance between the beamformer peaks and their associated modelled sources, excluding all peaks below the noise threshold of -28.68 dB. The locations of the beamformer peaks and the sources are plotted in Figure 3.5a and b with colour indicating the number of occurrences at each grid point. It can be seen from Figure 3.5a and b that both the observed and modelled sources occurred over the whole region, although the most energetic 100 events were concentrated around 40-50°N, 160-180°E (Figure 3.5c and d), which corresponds well with the locations observed by Obrebski et al. (2013) in the year 2010. The beamformer peak and associated modelled source were found to be located \(\leq10^\circ\) (5 grid points) apart in 90% of occurrences, were \(\leq6^\circ\) (3 grid points) apart in 67% of occurrences but only matched the same grid point in 4% of occurrences. Obrebski et al. (2013) found similar offsets (between 1.1° and 9.5°) for the 54 strongest modelled events in their study.

Discrepancies may be due to 3-d velocity structure that we are not taking into account in the backprojection (up to 4° (Euler et al., 2014)), contamination from other phases or earthquakes, contamination from array response (Gal et al., 2016), or inaccuracies in the modelled source location (either due to inaccuracies in the wave model or site effect). In some cases the discrepancy was found to arise from the resolution limits of the array, where two neighbouring sources merged into one observed source centred on...
Figure 3.5: Locations of beamformer peaks and associated modelled sources. a) Located peaks in beamformer output. Beamformer output was integrated between 0.1 and 0.3 Hz before applying a watershed algorithm at each time step to identify the peaks. Only peaks that were above -28.68 dB (the noise threshold) are plotted. b) Locations of corresponding modelled source peaks. The modelled source was also integrated between 0.1 and 0.3 Hz and the maximum source within each beamformer watershed region of the peaks shown in (a) was identified as the source peak. Colour indicates how many times the peak was located at each grid point. c) Same as (a) but for the 100 most energetic peaks. d) Same as (b) but only for the 100 events shown in (c).

To compare the frequency content of the modelled and observed sources, we examined the non-integrated output of each of the beamformer peaks and associated source peaks. Again we only considered beamformer peaks that were above the noise threshold, and in addition only the cases when the observed and modelled sources were ≤10° apart to exclude any spurious observations.

The mean spectra of the observed and modelled sources were calculated (Figure 3.6a and b), and the mean frequency from these spectra was calculated as $\sqrt{\frac{m_2}{m_0}}$, where $m_n = \int_{f_2=0.1Hz}^{0.3Hz} (S(f_2) \times f_2^n) df$ and $S(f_2)$ is either the beamform spectrum $b$ or the modelled source spectrum $P$. The mean frequency of the beamformer was 0.21 Hz, whilst the mean frequency of the modelled source was 0.19 Hz, indicating that the
seismic observations represented well the frequency content of the sources acting in the ocean basin. To examine the variability amongst individual spectra, the peak frequency of each spectrum was picked and plotted as a histogram with colour corresponding to number of occurrences in Figure 3.6c. Again it can be seen that both the beamform peak and modelled source peak occurred most often at about 0.2 Hz, with rare instances where the peak occurred down to 0.1 Hz or up to 0.3 Hz.

The results indicate that the seismic beamforming observations reflect the location, frequency content and amplitude of microseismic sources, with location accurate to 10° in 90% of occurrences and with a strong correlation coefficient between beampower and source power of 0.71.

Figure 3.6: Mean and peak frequencies of the beamformer output and modelled source. a) Mean beamformer (left) and modelled source (right) spectrum. Instances when the beamform peak and modelled source peak were > 10° apart were excluded from the calculation of the mean. b) Peak frequency of beamform spectrum vs. peak frequency of modelled source spectrum. The grayscale shading indicates the number of time steps the peak occurred at each given frequency.
Next we used the relationship between beampower and modelled source power found during the first winter to estimate source power over the second winter (mid-October 2013 to mid-March 2014) using the observed beampowers over that period.

To estimate source power from observed beampower, we fitted another curve between beampower and source power this time with beampower as the predictor and source power as the response variable (Figure 3.7a), and only considering beampowers over the noise threshold. Again we also excluded from the regression any cases when the observed and modelled sources were >10° apart. The best-fitting logarithmic function was applied to the seismic observations of the second winter to estimate sources during the second winter (a linear fit between beampower and source power for source powers >102 dB was attempted but resulted in overestimation of sources at large beampowers). The time series of the modelled source amplitude \(10^\left(P_{source}/10\right)\) is plotted along with estimated source amplitude in Figure 3.7b, where only the largest estimated source at each time step is plotted if there were more than one source acting simultaneously. Pearson’s linear correlation coefficient between the estimated source power and modelled source power was 0.63 (>99% significance).

### 3.4.2 Relating microseism source to significant wave height

The results presented have shown that we are able to estimate the integrated microseism sources in the North Pacific from seismic observations with some confidence. If we now consider observations at each frequency, we can reconstruct the full source spectrum, from which the ocean wave energy spectrum can be estimated, and thus significant wave height. This is a rearrangement of equation 3.3:

\[
E(x, f) = \sqrt{\frac{P(x, f_2)}{2\left|CP(x, f_2)\right|\rho_c^2\rho_w^2\rho_g^2 f_2 I(x, f)}}
\]  

(3.4)

where the source power spectral density spectrum \(P(x, f_2)\) is estimated from the observed beampower spectrum, values of \(\rho_c\), \(\rho_w\) and \(C_p\) are taken at the grid points where peaks in beampower are located, and \(I(x, f)\) is unknown.

Significant wave height can then be calculated from the estimated ocean wave energy spectrum:

\[
H_s(swell) = 4\sqrt{\int_{0.05Hz}^{0.15Hz} E(f)df}
\]

(3.5)
Figure 3.7: Modelled source power, observed beampower and estimated source power.

a) Modelled source power vs. observed beampower for the first winter. Both are integrated between 0.1 and 0.3 Hz. Values of beampower below the noise threshold of -28.68 dB and cases when the observed and modelled source locations were >10° apart were excluded. The red circles are the mean of each data bin with error bars of ± two standard deviations. The green curve is the best-fitting logarithmic function of the form $y = a \log(b(x-x_0)) + y_0$ where values of $a, b, x_0$ and $y_0$ are given in the figure. b) Modelled source vs. estimated source for the second winter (mid-October 2013 to mid-March 2014). The estimated source was calculated from the observed beampower using the regression in (a). The time series has been split over two plots for clarity.

Here we distinguish our estimate of $H_s$ as $H_s(\text{swell})$ because we are only considering the spectrum between seismic frequencies of 0.1-0.3 Hz (0.05-0.15 Hz ocean wave frequency which equals 6.7-20.0 s periods) whereas traditionally significant wave height is calculated from a wider band (e.g. 2-30 ocean wave periods).

To estimate the source spectrum $P(x, f_2)$ from the observed beampower, beampower and source power were correlated in a similar way as previously, but instead of plotting the integrated powers, the power at each frequency was plotted separately (Figure 3.8a). From Figure 3.8a it was found that values at the different frequencies all lay along the same curve, so we applied one relationship to all frequencies. Again we binned the data points according to source power and fitted an exponential function. The noise level of the beampower (as defined previously) was found to be -20.16 dB
Figure 3.8: Modelled source power, observed beampower, mean shape of modelled source spectrum and mean spectrum of wave-interaction intensity. a) Similar to Figure 3.4 but with modelled source power vs. observed beampower plotted separately at each frequency. Only cases where the integrated value of the beampower was above the noise threshold of -28.68 dB are plotted, and only cases where the observed and modelled source were apart by \( \leq 10^3 \). The red circles are the mean of each data bin with error bars of \( \pm \) two standard deviations. The dashed line is placed at the top of the error bar of the first bin and is equal to -20.16 dB, which we define as the noise threshold. The green curve is the best-fitting exponential of the form \( y = ae^{b(x-x_0)} + y_0 \) where values of \( a, b, x_0 \) and \( y_0 \) are given in the figure. b) Same as (a) but with the axes switched and only including beampowers(\( f_2 \)) \( \geq -20.16 \) dB. The green curve is the best-fitting logarithmic function of the form \( y = a \log_b(b \times (x-x_0)) + y_0 \). c) Mean shape of modelled source spectrum as a function of \( f/f_p \) which is the seismic frequency divided by the peak seismic frequency of the source spectrum. The data is plotted in black and the smoothed data in red (almost identical). d) Mean spectrum of the wave-interaction intensity.

and the source exceeded the noise level at a value of 117 dB \( (5.0 \times 10^{11} \text{ Pa}^2 \text{ m}^2 \text{s}) \).

Taking only beampower values above this noise threshold, we replotted using beampower as the predictor variable (Figure 3.8b) and found that source power at each frequency is best estimated from the observed beampower using a logarithmic function given by: \( P_{\text{estimated}}(dB) = a \log_b(b \times (\text{beampower} - x_0)) + y_0 \) with \( a=10.3564, b=0.89105, x_0=-20.8629 \) and \( y_0=108.5138 \).
The estimated source spectrum was constructed from the beampower spectrum using the identified logarithmic relationship in the following way. If the peak of the beampower spectrum was above the noise level of -20.16 dB, the source power of this peak, \( P_{\text{estimated}} \) (dB), was estimated, and the source amplitude as
\[
P_{\text{estimated}} = 10^{\frac{P_{\text{estimated}}(\text{dB})}{10}}.
\]
It was found that by setting the estimated source amplitude at all other frequencies (below the noise threshold) to zero, the source spectrum and resulting ocean wave spectrum became too narrow. Instead, the mean shape of the modelled source spectrum as a function of \( f/f_p \) (where \( f_p \) is the peak frequency of the spectrum) was calculated (Figure 3.8c), and \( P_{\text{estimated}} \) was multiplied by this. An example of the construction of one spectrum in this way is shown in Appendix C Figure C.2. An assumption here is that the source spectrum only has one peak, and indeed it was found that the modelled source spectrum had a single peak (counting peaks as those above the mean + one standard deviation) in 99% of cases whilst the beamformer spectrum had a single peak in 80% of cases. We limit our estimation of \( P_{\text{estimated}} \) to cases when both the beamformer spectrum and source spectrum had a single peak.

\( I(f) \) depends on the degree of wave interaction at each frequency. When the opposing wavefield is a result of coastal reflection this can be related to the reflection coefficient, or when the opposing wavefield is a result of directional spread this be related to the wave energy (Ardhuin et al., 2012). However for two opposing wavefields in mid-ocean regions \( I(f) \) cannot necessarily be related to the ocean wave energy but depends on the (unknown) characteristics of both wavefields. For example, a given \( I(f) \) could result from two low energy swells that directly oppose each other, or from one large energy swell that meets a low energy wind-sea. The assumption made about \( I(f) \) is therefore expected to be one of the largest sources of error in the estimation of ocean wave energy and significant wave height from seismic observations. We calculated the mean \( I(f) \) from the modelled cases as a function of \( f/f_p \) (Figure 3.8d) and assumed this \( I(f) \) spectrum in the calculations.

Because the previous results indicated that location was only accurate up to 10° in the majority of cases, the map of \( \left[ 2|C_p| \frac{\rho_c}{\rho_w} \right]^2 \) was first smoothed over 10° using a 11-by-11 low-pass spatial Gaussian filter before taking the value at the beampower location.

The wave energy spectrum was estimated from the beamforming observations of the first winter using equation 3.4, and significant wave height estimated using equation 3.5. Figure 3.9 shows the modelled \( H_s(\text{swell}) \) vs. the estimated \( H_s(\text{swell}) \) which had a Pearson’s linear correlation coefficient of 0.48. The scatter points are shaded by point density and binned by modelled \( H_s(\text{swell}) \). It can be seen from the figure that at the most common wave heights (\( \sim 2-3 \text{ m} \)) the estimation was most accurate, whereas higher modelled wave heights were underestimated. We used the mismatch between the estimated \( H_s(\text{swell}) \) bins and the line \( y = x \) to define a calibration factor, which
Figure 3.9: Modelled $H_s$\textit{(swell)} during the first winter (mid-October 2012 to mid-March 2013) vs. $H_s$\textit{(swell)} estimated from beamformer output. The red circles are the mean of each data bin with error bars of ± two standard errors. The grayscale shading indicates the point density (number of points per m$^2$). The offset between the mean of each data bin and the line $y = x$ was used as a calibration factor for the second winter.

was then applied to estimates of $H_s$\textit{(swell)} during the second winter (a calibration factor based on a curve which smoothed out the uncertainties at the larger wave heights did not make any significant difference to the results). Figure 3.10 shows modelled vs. estimated $H_s$\textit{(swell)} for the second winter which had a Pearson’s linear correlation coefficient of 0.63. With the calibration factor, the means of each data bin lie much closer to the line $y = x$, with underestimation only occurring at $H_s$\textit{(swell)} values over about 8m.
Figure 3.10: Modelled and estimated $H_s$ swell. a) Modelled $H_s$(swell) during the second winter (mid-October 2013 to mid-March 2014) vs. $H_s$(swell) estimated from beamformer output. The red circles are the mean of each data bin with error bars of ± two standard errors. The grayscale shading indicates the point density (number of points per m$^2$). b) Time series of modelled and estimated $H_s$(swell) during the second winter. Only the maximum estimated $H_s$(swell) at each time step is plotted.

3.5 Discussion

Errors in the estimate of $E(f)$ using equation 3.4 and consequently $H_s$ will arise for three main reasons: 1) Inaccurate estimate of source amplitude $P(f_2)$ from beamformer amplitude 2) Inaccurate location of beamformer peak, therefore inaccurate site effect 3) inaccurate estimate of $I(f)$. Another expected source of error is that is it not possible to extract information about the ocean wave spectrum at frequencies for which the wavefield is unidirectional ($I(f) = 0$), because the observed beampower is only sensitive to opposing waves. Consequently, these parts of the ocean wave energy spectrum would be underestimated as would the derived significant wave height.

In order to understand how each of these errors influence the correlation between estimated and modelled $H_s$(swell), we looked at each of these sources of error in turn. Figure 3.11a shows the ideal case when the source spectrum, site effect and $I(f)$ are known exactly for the estimation of $E(f)$ using equation 3.4 (i.e. the modelled values were used). The correlation coefficient in this case between modelled $H_s$(swell) and estimated $H_s$(swell) was equal to 1.00. Underestimation would be expected if there was a large portion of unidirectional wave energy in addition to the opposing
Figure 3.11: (Continued on the following page)
Figure 3.11: Causes of error in the estimation of $H_s(\text{swell})$ during the first winter. In all plots modelled $H_s(\text{swell})$ is on the x-axis and estimated $H_s(\text{swell})$ is on the y-axis. $r$ is the Pearson’s linear correlation coefficient. ‘Modelled source’ means that modelled $P(f_2)$ was used in the calculation. ‘Modelled location’ means that the modelled source location was used for the value of the site effect. ‘Modelled $I(f)$’ means that the exact (modelled) value of $I(f)$ for each case was used. ‘Estimated source’ means that $P(f_2)$ was estimated from the beampower. ‘Estimated location’ means that the beamformer location was used to obtain the value of the smoothed site effect. ‘Estimated $I(f)$’ means that the $I(f)$ of Figure 3.8d was used. The plots therefore represent: a) The ideal case. b) error in $I(f)$. c) error in location. d) error in $I(f)$ and location. e) error in source amplitude. f) error in source amplitude and $I(f)$. g) error in source amplitude and location. h) error in source amplitude, $I(f)$ and location.

wavefields, but this wasn’t the case as the modelled $I(f)$ was rarely (<1% of the time) exactly 0. Plot b examines the error due to inaccurate estimate of $I(f)$. In this estimate of $H_s(\text{swell})$ the mean value of $I(f)$ shown in Figure 3.8d was used in the calculation. There was still moderately strong correlation of 0.78 but this was much lower than the ideal case (a). There also appears to be a tendency for underestimation which increases with modelled $H_s(\text{swell})$. This would be caused by the assumed $I(f)$ being too large for higher wave heights. An explanation for this may be related to the case for waves generated by local winds, in which $I(f)$ generally decreases with increasing $E(f)$ (Ardhuin et al., 2012). Plot c examines the error due to inaccurate location (site effect). In this estimate of $H_s(\text{swell})$ the location of the beamformer peak was used to obtain the value of the smoothed site effect. The correlation of 0.90 indicates that the error introduced from inaccurate location is less than the error introduced by inaccurate $I(f)$. Plot d examines the combined error from $I(f)$ and location and as expected the correlation drops further, to 0.71.

Plots e-h are the same as plots a-d but with the addition of error caused by inaccurate estimate of the source amplitude. Plot e) shows that inaccurate amplitude estimation causes a larger reduction in correlation than errors in $I(f)$ and location combined. With inaccuracies in $I(f)$ and location correlation goes down to 0.48 (Figure 3.11h).

The large amount of scatter between source amplitude and beampower is unsurprising given the amplitude removal during the beamforming process and, as mentioned by Obrebski et al. (2013), because the beampower also depends on the size of the area the source is acting over as well as energy losses along the propagation path (including geometric spreading, attenuation and transmission through Earth structure such as the Moho, 410km and 660km discontinuities (Nishida and Takagi, 2016)). Nevertheless, our results show that a relationship between source amplitude and observed beampower does exist. Furthermore, Figure 3.8b suggests that with improvements in beampower and location estimation correlations of up to 0.78 are possible even with the uncertainty surrounding $I(f)$. A direction for these improvements may be found in
the recent work of Farra et al. (2016) and Nishida and Takagi (2016). Farra et al. (2016) used a ray-theoretical approach to estimate P-wave ground displacement for a given source including site, receiver and propagation effects, and Nishida and Takagi (2016) used a similar formulation to estimate the pressure source by minimising the squared difference between observed and modelled ground displacement.

Finally, throughout the study, the modelled source has been considered the ‘true’ value. Scatter between the beampower and modelled source may be caused by inaccuracies in the wave model itself, for example due to the wind input or parameterisation within the model. Although there is currently no other way of estimating wave-interaction intensity over such spatial and temporal scales with which to validate the model output, an idea of error within the WAVEWATCH III model could be found by analysing the spread of results obtained from multiple runs with different wind inputs and parameterisation. Scatter may also be caused by inaccuracies in the calculation of the site effect, which may not well represent each 2 by 2° pixel in regions of large bathymetric variability (Hillers et al., 2012), and for which we have not taken into account the effect of sediments or earth structure below the upper crust (Gualtieri et al., 2014). A thick sediment layer at the source results in reduced amplitudes of land-recorded microseisms (Gualtieri et al., 2015), and may be important for sources close to the coast where sediments are thicker.

It is important to remember however that estimates about significant wave height can only be made where there is wave interaction occurring. Sometimes this does correspond to the largest wave heights in the ocean basin (e.g. Figure 3.12a,b), but this is not necessarily the case (Figure 3.12c,d).
Chapter 3 Monitoring ocean waves from P-waves 79

Figure 3.12: An example when maximum source corresponds to location of maximum $H_s$ (top row) and when maximum source does not correspond to location of maximum $H_s$ (bottom row). a) Modelled source on 22 December 2012 15:00-18:00. b) Modelled $H_s$ on 22 December 2012 15:00-18:00. c) Modelled source on 27 December 2012 15:00-18:00. d) Modelled $H_s$ on 27 December 2012 15:00-18:00. The modelled source has been integrated over $f^2 = 0.1-0.3$ Hz whilst $H_s$ has been calculated from the modelled ocean wave spectra over the full ocean wave frequency range $f = 0.0300-0.5758$ Hz.

3.6 Conclusions

Observed seismic P-waves in California were found to relate to modelled microseismic sources in the North Pacific Ocean in terms of location, frequency content and amplitude. The observed P-waves were located through beamforming and backprojection, and were found to match the location of strong modelled sources by $\leq 10^\circ$ (5 grid points) in 90% of cases. Both the modelled sources and observed P-waves were dominated by microseisms with a frequency of approximately 0.20 Hz. Beampower was moderately to strongly correlated with the power of the modelled sources, and only exceeded the noise threshold when sources were $>1.6 \times 10^{10}$ Pa$^2$m$^2$ (integrated between 0.1 and 0.3 Hz) or $>5.0 \times 10^{11}$ Pa$^2$m$^2$s. The empirical relationship between beampower and source power allowed sources during the second winter to be estimated from observed beampower. The resulting estimated sources were found to correlate with the modelled sources with a correlation coefficient of 0.63.
After reconstructing the source spectrum from the beamformer spectrum, and making an assumption about the directional characteristics of the wavefield, the ocean wave energy spectrum was estimated, and from that, significant wave height. During the first year, the modelled and estimated significant wave height correlated by 0.48. An underestimation of wave height at higher modelled values appeared to be introduced in the assumption of the directional wave characteristics. A calibration factor between modelled and estimated significant wave height calculated from the first winter’s results was applied to the second year and the underestimation was largely removed, giving a correlation of 0.63 between modelled and estimated significant wave height. Inaccuracy in beampower was found to be the largest source of error, followed by inaccuracy in directional assumption. With improvements in the location and amplitude estimation of sources from beamforming, it should be possible to obtain estimates of significant wave height that correlate with modelled wave heights by up to 0.78.
Chapter 4

Improving microseismic P-wave source location with multiple seismic arrays

4.1 Abstract

Using array analysis, the direction and distance to a seismic P-wave source can be determined. However, individual arrays are limited in their geographical coverage and by their resolving capability, which is determined by the array aperture, configuration and number of stations. We demonstrate these limitations on three large seismic arrays located in Japan, Europe and California, and find that all give a unique but imperfect insight into the P-wave sources acting in the North Pacific. We then combine the data from all three arrays in order to get a more complete and comprehensive view. By considering all three arrays together, a wider geographic range of sources is captured, location accuracy is improved, and neighbouring sources become more easily identified. Next we weight each array for distance in order to optimise the result. Being able to resolve and accurately locate source regions is an important step in being able to use seismic records to monitor ocean wave activity and track storms in real time.

4.2 Introduction

Double-frequency, or ‘secondary’, microseisms are continuous vibrations observed by seismometers worldwide, caused by the interaction of opposing ocean waves and subsequent coupling of the pressure signal with the seabed (Longuet-Higgins, 1950). They are dominated by surface waves (Landès et al., 2010; Roux et al., 2005; Sabra
et al., 2005; Shapiro and Campillo, 2004), but many studies have now reported observations of compressional body waves (P-waves) in the double frequency band (Backus et al., 1964; Toksöz and Lacoss, 1968; Haubrich and McCamy, 1969; Roux et al., 2005; Koper and de Foy, 2008; Gerstoft et al., 2006, 2008; Zhang et al., 2010a; Landès et al., 2010; Obrebski et al., 2013; Euler et al., 2014; Gualtieri et al., 2014; Reading et al., 2014; Neale et al., 2017), and more recently, S-waves (Liu et al., 2016; Nishida and Takagi, 2016).

P-waves have a distinct advantage over surface waves for sources in the deep ocean because they can be located by distance as well as direction (Haubrich and McCamy, 1969; Gerstoft et al., 2006). Early studies (Toksöz and Lacoss, 1968; Haubrich and McCamy, 1969) found that they originated in the deep-ocean near storms. Many studies have since confirmed a deep-ocean origin (Landès et al., 2010; Gerstoft et al., 2008; Zhang et al., 2010a; Euler et al., 2014; Koper et al., 2010) as well coastal sources (Gerstoft et al., 2006; Zhang et al., 2010a), with most sources located in the 30-60° latitude band associated with extra-tropical cyclones and corresponding to regions of elevated wave heights (Euler et al., 2014). Strong seasonality has been observed in amplitude, source location and P-wave phase e.g. P, PP or PKP (Landès et al., 2010; Gerstoft et al., 2008; Euler et al., 2014; Koper and de Foy, 2008; Reading et al., 2014; Hillers et al., 2012). The link between P wave microseisms and sea-state leads to exciting possibilities of using P-wave microseisms to track storms and monitor deep-ocean wave activity in real-time (Gerstoft et al., 2006; Davy et al., 2014; Zhang et al., 2010a; Reading et al., 2014), complementing oceanographic and satellite observations especially in regions with poor data coverage, as well as to calibrate wave model hindcasts (Ardhuin et al., 2012) and supplement earthquake sources to improve P-wave seismic tomography (Zhang et al., 2010b).

Accurately locating microseism sources is essential for relating the measured seismic noise to wave activity, and is also important for tomography studies which may be biased by a non-isotropic distribution of sources (Harmon et al., 2010). However, individual arrays are limited by their coverage and also by their resolution in their ability to resolve closely spaced sources, as noted previously by Hillers et al. (2012) and Euler et al. (2014). In such instances, two closely spaced sources will appear in the beamforming result as a single source centred on the average location. Using multiple seismic arrays will improve coverage and may improve the ability to resolve neighbouring P-wave sources. Previous authors (Euler et al., 2014; Hillers et al., 2012; Landès et al., 2010; Pyle et al., 2015) have combined arrays with a focus on detecting common, robust sources over monthly or seasonal time periods, but here we combine arrays with the purpose of locating closely-spaced sources over short (3-hour) time periods, which is the typical output frequency of operational ocean wave models. This builds on our previous work (Neale et al., 2017) where we located P-wave sources in...
Chapter 4 Locating P-wave microseisms with multiple arrays

the North Pacific with the same temporal resolution using just the California array, and which found location accuracy to be up to 10° is most cases. An improvement in source location will be beneficial for possible ocean or seismic monitoring, ocean wave model validation, and data assimilation into operational wave models. We show here that a combination of arrays can improve the resolution of multiple sources not visible by individual arrays alone.

4.3 Data and methods

4.3.1 Numerical modelling of P-wave sources

P-wave microseism sources were modelled following the method of Ardhuin et al. (2011), Ardhuin and Herbers (2013) and Farra et al. (2016). The second order pressure spectrum $F_p$ (with units of $\text{Pa}^2\text{m}^2\text{s}$) resulting from the interaction of opposing ocean waves of similar frequency $f$ (Hasselmann, 1963) is given by:

$$F_p(x,f_2) = [2\pi]^2(\rho_w g)^2f_2E^2(x,f)I(x,f) \quad (4.1)$$

where $f_2 = 2f$, $x$ is location, $\rho_w$ is the density of seawater, $g$ is gravitational acceleration, $E(x,f)$ is the ocean wave frequency spectrum and $I(x,f)$ is a non-dimensional function that depends on the wave energy distribution $M$ over the directions $\theta$:

$$I(x,f) = \int_0^\pi M(x,f,\theta)M(x,f,\theta + \pi)d\theta \quad (4.2)$$

$F_p$ was calculated on a global 0.5° longitude by 0.5° latitude grid using the numerical ocean wave model WAVEWATCH III (Tolman, 2014) with 6-hourly wind and sea-ice concentration from ECMWF’s ERA-interim reanalysis (Dee et al., 2011). $F_p$ was then resampled to a 2°-by-2° grid (smoothed first using a 7-by-7 Gaussian low-pass filter) and averaged over 3-hour time periods between October 2012 and March 2013.

The P-wave microseism source at each grid point was calculated by multiplying the second order pressure spectrum by the squared source site effect $|\frac{2|C_p|\rho_c}{\rho_w}|$ (Farra et al., 2016):
We used $\rho_w = 1020 \text{ kg m}^{-3}$ and spatially varying values of crustal density $\rho_c$ were taken from the global crustal model CRUST1.0 (http://igppweb.ucsd.edu/~gabi/rem.html). $C_P$ is an amplification coefficient that depends on frequency, water-depth and P-wave take-off-angle (Ardhuin and Herbers, 2013; Gualtieri et al., 2014). We used the formulation of Gualtieri et al. (2014) and Farra et al. (2016) to calculate $C_P$ at each frequency using water-depth from ETOP01 bathymetry (Amante and Eakins, 2009). Because the take-off angle applicable for each array depends on the distance between the source grid point and the array, the maps of source site effect would be slightly different for each array. Instead a non-angle-dependent site effect was calculated by averaging over the full range of take-off angles, which is shown at 0.207 Hz in Figure 4.1a. The similarity between the non-angle-dependent amplification coefficient and the angle-dependent amplification coefficients for each of the three arrays can be seen in the Figures 4.1 b,c and d. These site effect maps are like those shown previously in Gualtieri et al. (2014) and Farra et al. (2016).

We considered P-wave microseism sources at frequencies of 0.188, 0.207 and 0.227 Hz to focus on the peak frequency of double-frequency P-wave sources in the North Pacific (Neale et al., 2017) and searched for time periods with multiple or closely-spaced sources acting in the basin.

### 4.3.2 Seismic data and processing

Seismic vertical component data were downloaded for the days identified from the modelled sources. Continuous waveforms were obtained from the Southern California Earthquake Data Center (SCEDC, 2013), the Data Management Center of Japan’s National Research Institute for Earth Science and Disaster Resilience (Obara et al., 2005) and the European Integrated Data archive (http://www.orfeus-eu.org/eida/eida.html). Figure 1.10 in Chapter 1 shows the networks and distribution of stations within the arrays.

Data were downsampled to 1 Hz if necessary, instrument response was removed, and were band-pass filtered between 0.002 and 0.400 Hz. Earthquakes events over magnitude 5 were identified using the ISC bulletin (International Seismological Centre, 2013) and removed from the data by setting a 1-hour window of the waveform to zero.
Chapter 4 Locating P-wave microseisms with multiple arrays

Figure 4.1: $C_P$ at 0.207 Hz. a) Averaged for all take-off angles relevant for distances of 15-99°. b) Calculated at the Californian array. c) Calculated at the Japanese array. d) Calculated at the European array.

if the RMS of that window was over 3 times the daily RMS. Any remaining bad quality data were identified by visual examination of the daily spectra and discarded.

For each array separately, P-wave sources were located using the beamforming method (Gerstoft et al., 2006, 2008; Obrebski et al., 2013; Gualtieri et al., 2014; Farra et al., 2016) direct onto a 2-by-2° geographical grid based on P-wave travel times, $t_P$, from each grid point $x$ to each station:

$$B_Z(f_2, x, t_s) = \frac{1}{N^2} \left\langle \sum_{j=1}^{N} S_j(f_2, t_s) w_j(x) e^{-i2\pi f_2 t_2(x)} \right\rangle^2$$

(4.4)

$N$ is the total number of stations, $S_j$ is the complex spectrum of the record at station $j$, $f_2$ is the seismic frequency, $t_s$ refers to the start time of the Fourier transform, $w_j$ is a weighting coefficient for station $j$. At first we use $w_j = 1$. $S_j$ was calculated on 512-s segments overlapping by 50% and was divided by its magnitude to retain only phase, thereby making the arrays more comparable in amplitude. We used $f_2 = 0.188 \pm 0.01$, $0.207 \pm 0.01$ and $0.227 \pm 0.01$ Hz. P-wave travel times for a source at the surface were calculated using the global ‘AK135’ travel time tables (Kennett et al., 1995). The
angle brackets denote averaging over each frequency band and over 3-hour periods to get the output at a 3-hour snapshot time $t_s$.

Each beam projection is limited to a geographical range of 99° from the array as beyond this the P-waves will start to be effected by the Earth’s outer core. In addition, P-waves from locations too close to the array will be regional and the global velocity model we use is unlikely to be accurate enough, so we project for a minimum distance of 15°. The arrays in California and Japan both give good coverage over the North Pacific, each filling the gaps of the other, whilst the European array offers coverage in the northern half of the North Pacific.

Looking at multiple images to pick out the best features from each array when the true source is unknown would be time-consuming and would probably require input and judgement from a human operator. This is not ideal in the context of using the observed sources for ongoing and possibly near-real-time monitoring, so in order to synthesise the results from multiple arrays we combined the data from all arrays to produce one beamforming image. Because $t_P$ of equation 4.4 is independent between each station and each grid point, we simply included the stations from all arrays in the summation.

Finally, the empirical weight $w$ was adjusted for each grid point and each station based on the distance ($d$) in degrees from that grid point to the station.

$$w = \frac{1/d}{\sum 1/d}$$

This is an empirical weighting based on the observation that sources closest to Japan in particular were much better resolved using the Japan array, and therefore acts to boost the influence of the arrays closest to each grid point. The term $\sum 1/d$ equalises the total weighting at each grid point. To reduce the effects associated with small numbers at the very edges of the range, the weighting for grid points 15° to 18° from a station was reduced to equal the average weighting at 18° from the station. Maps of $1/d$ and $w$ for three stations (one in each array) is shown in Figure 4.2.
Figure 4.2: Weights for distance. Top row: one station in the California array, middle row: one station in the Japan array and bottom row: one station in the Europe array. The first column shows the result of the first half of the weighting calculation \((1/d)\). The second column gives the final weight \((1/d)/\sum 1/d\).
4.4 Synthetic tests

The ability of an array to locate sources and resolve neighbouring sources depends on its resolution. The angular resolution of an array is determined by the array aperture, and depends on the wavelength (hence frequency) of the wave being examined (Rost and Thomas, 2002). As we are beamforming onto a geographical grid, angular resolution corresponds to a spatial resolution that also depends on distance from the source. The spatial resolution, $R$, of the array can be approximated by the intensity pattern width associated with a point source (Hillers et al., 2012). This is equal to $R = \lambda \Delta / D$ where $\lambda$ is seismic wavelength, $\Delta$ is distance from array to source and $D$ is the aperture of the array. For example, for the Californian array, a 0.2 Hz source at 43°N 171°E ($\lambda=28$ km, $\Delta=6100$ km, $D=1400$ km) has a theoretical spatial resolution $R \sim 120$ km. For the Japan array, the same source at the same location ($\Delta=3400$ km, $D=610$ km) has a theoretical spatial resolution of 160 km and for the Europe array ($\Delta=10,000$ km, $D=1300$ km) it is 210 km. In practice, the ability of an array to resolve neighbouring sources will also depend on the number of stations (resolution of velocity) and station configuration (spatial aliasing) (Rost and Thomas, 2002, 2009). (Hillers et al., 2012) found that neighbouring sources with a separation distance much further apart than their estimated resolution could not be resolved.

We tested the resolving capability of each array using synthetic point sources of amplitude $A = 1$, by replacing $S_j$ with $\text{Syn}_j$, where $\text{Syn}_j = \sum_{k=1}^{N_p} A e^{-i2\pi f_j t_k}$ i.e. the sum of $N_p$ phase delays for sources located at grid points $k_1...k_{N_p}$. For each array, we placed a source at 43°N 171°E and a second source some distance to either the north or the east. The synthetics were tested at $f_2 = 0.207$ Hz. For the California array (Figure 4.3) the two sources separated at a distance of 670-1100 km to the north or 1300-1500 km to the east. For the Japan array (Figure 4.4) they separated at a distance of 670-890 km to the north or 650-2100 km to the east. For the Europe (Figure 4.5) array it was 890-1300 km to the north or 820-1000 km to the east. These distances are all much larger than the theoretical resolutions calculated previously, and for real sources the resolution may be even lower as sources may not act as point sources. The synthetic tests also reveal the importance of viewing direction when separating two neighbouring sources. For both the California array and Japan array, north-south separated sources would be identified from their different azimuths, whereas for the Europe array north-south separated sources would be identified by their different slownesses (take-off angle). In each case, sources separated azimuthally from the array are better resolved than those separated by distance, meaning that the California and Japan arrays are better able to resolve north-south separated sources, whereas the Europe array is better able to resolve those separated east-west.
Figure 4.3: Testing how well the California array identifies two neighbouring point sources. In each plot one synthetic source is placed at 43°N, 171°E and all plots are normalised to a maximum of 0 dB. In a) b) and c) the second synthetic source is placed at increasing distances to the north and in d) e) and f) it is placed at increasing distances to the east. a) Separation = 445 km b) Separation = 667 km c) Separation = 1112 km d) Separation = 815 km e) Separation = 1303 km f) Separation = 1465 km.

Figure 4.4: Same as Figure 4.3 but for the Japan array. a) Separation = 445 km b) Separation = 667 km c) Separation = 889 km d) Separation = 489 km e) Separation = 652 km f) Separation = 2112 km.
Figure 4.5: Same as Figures 4.3 and 4.4 but for the Europe array. a) Separation = 667 km b) Separation = 889 km c) Separation = 1335 km d) Separation = 652 km e) Separation = 815 km f) Separation = 978 km.

The ability of array combination to improve the resolving capability was also tested using synthetic point sources. The first point source was again placed at 43°N 171°E and the second 667 km to the north. Previously we found that the California and Japan arrays were only partly able to separate these sources (Figures 4.3b and 4.4b) and the Europe array was unable to at all (Figure 4.5a). For the combined arrays (Figure 4.6a) two source regions emerge. Similarly, a second point source was placed 489 km to the east. Previously we found that none of the arrays were able to separate these sources. Combining the arrays, two source regions do emerge (Figure 4.6b). These tests demonstrate the concept of combining arrays to improve the resolution, however they are very idealised with point sources and no noise and the resolution of the real data is not expected to be as high.

The distance weighting scheme was tested for sources close to Japan, with two sources placed just 370 km apart to the northeast of Japan. Combining the arrays with distance weighting (Figure 4.6f) resulted in a better resolution of the two sources than without the weighting (Figure 4.6c), although they were separated in both cases.
Figure 4.6: Testing how well the combined arrays identify two neighbouring point sources. a) One synthetic source is placed at 43°N, 171°E and the second is placed 667 km to the north. b) One synthetic source is placed at 43°N, 171°E and the second is placed 489 km to the east. c) Two closely spaced sources near Japan at 49°N 153°E and 47°N 157°E. d) e) f) are same as (a, b, c) but with distance weighting.

4.5 Results

4.5.1 Individual arrays

Sources observed at each array on selected days are shown along with modelled sources in Figures 4.7 to 4.15. The figures show how each array offers a unique view of the sources acting at any given time, and that limitations of an individual array can be overcome by considering the results from multiple arrays. On the 13 January 2013 18:00-21:00 (Figure 4.7) there is one source just south of the westernmost Aleutian Islands and another (the largest) source about 2500 km to the southeast along with a couple of other sources on this northwest to southeast line. Both the California array (Figure 4.7b) and Japan array (Figure 4.7c) observe the southeast source, with a tail towards the northwest, whereas the Europe array (Figure 4.7d) highlights the source in the northwest whilst the southerly source is out of range. Figures 4.8a-d on the 15 February 2013 06:00-09:00 show a case where one of the sources (southeast of Japan) is only observed, although very weakly, by the Japan array. Figures 4.9a-d on 10 October 2012 06:00-09:00 shows a case when it is only the Europe array that observes a clear
source in the north (again just south of the Aleutian Islands) but misses the other two sources to the southeast and southwest. The California array observes both the southeast and southwest source, whilst the Japan array observes a strong southwest source and a weak blurring of the north and southeast sources.

A benefit of using multiple arrays is therefore better coverage, and the potential of observing sources that would otherwise be missed. Another, related, advantage is that different arrays are able to separate out neighbouring sources in different situations. One example of this is seen on 27 October 06:00-09:00 (Figure 4.10), where there are two east-west aligned sources separated by 1800 km. The Europe array distinguishes two sources, but they are merged and centred on the average location between them in the outputs from California and Japan. This also indicates that the synthetic point tests overestimated the ability of the arrays to resolve neighbouring sources, as both the California and Japan arrays could begin to separate the synthetic sources at distances of about 1300 km.

In other examples, sources near the east coast of Japan were very well resolved by the Japan array but poorly or unresolved by the other arrays. This is seen in the examples on 19 November at 03:00-06:00 (Figure 4.11) and 12:00-15:00 (Figure 4.12). At 03:00-06:00, the California array merges all of the sources on a roughly north-south line east of Japan. The Japan and Europe arrays both distinguish two source regions but miss the source further south below 30°N. The resolution of the Japan array for the two sources it does see is extremely high. At 12:00-15:00, what is very blurred output by the California and Europe arrays becomes highly resolved when viewed from Japan. On 19 January 00:00-03:00 (Figure 4.13) it can again be seen that the Japan array is able to resolve north-south separated sources which are merged by the California array. In other examples (Figures 4.14 and 4.15) it is only the California array that is able to separate neighbouring sources.

### 4.5.2 Combined arrays

The results presented so far have shown just how differently the same sources are imaged using different arrays. Therefore, in order to get a fuller picture of the sources acting in the ocean basin at any given time, it is beneficial to consider the results from more than one array. An ideal combined image would better represent the sources than any of the single images, but for automation of source monitoring, it would also be advantageous if it was only as good as the best single image.

Results of array combination for each of the previous examples using the real data are shown in plot (e) of Figures 4.7 to 4.15. For the first example on 13 January 2013 18:00-21:00 (Figure 4.7), the combined image shows a much better representation of
the sources than any of the individual array images, although the northern source observed only by the European array is not completely separated and the image is dominated by the southeast source. In the second example (15 February 2013 06:00-09:00, Figure 4.8) the northern source seen by all arrays appears in the combined image as expected. The weak source observed by the Japan array is not visible in the combined image. In the third example (10 October 2012 06:00-09:00, Figure 4.9) the combined image is worse than the California array output as the eastern source is not as well resolved, but it does better represent the source area than the individual Japan or Europe images. Similar situations can be seen in the example of 02 October 2012 21:00-24:00 (Figure 4.14) where the northwest-southeast sources nicely separated by the California array are not as clear in the combined image, also on 05 October 2012 21:00-24:00 (Figure 4.15) where the southwest source partially separated in the California output is not visible in the combined image, and on 19 January 2013 00:00-03:00 (Figure 4.13) where the individual Japan output is the best. In the forth example on 27 October 2012 06:00-09:00 (Figure 4.10) there is some hint of source separation in the east-west direction, as there was in the Europe image, but overall the image is dominated by the California and Japan arrays which place one source in between the two east-west aligned sources. In the fifth, sixth and seventh examples (Figures 4.11-4.13) for sources very close to Japan, the combined image offers a big improvement over the individual California and Europe arrays but does not add any improvement to the individual Japan image.

Overall, it was found that the combined image generally gave a good representation of P-wave sources visible in the single arrays, balancing advantages and disadvantages of each. This meant that the combined image was sometimes less clear than the best single image, but was able to pick up to some extent sources only visible in one array, or was able to better resolve sources compared to individual arrays. However, in some cases, including where the combined image failed (was worse than the best single image), a source observed by only one array was overshadowed by the other arrays and did not appear in the combined image. In particular this happened for sources in the north seen by the Europe array which were overshadowed by the Japan and California arrays (Figure 4.7 and Figure 4.9). In other cases it was found that the location quality of a single array image was reduced after combining with data from stations that were much further from the source location. This was most visible in the sources close to Japan which were captured best by the Japan array (Figures 4.11, 4.12 and 4.13). This may be because the model is unphysical, in that distance (and therefore effects of geometric spreading and attenuation) is not taken into account. In order to retain the high resolution of source locations near Japan imaged by the Japan array, we adjusted the weighting $w_j$ to enhance the relative contribution of those stations that are closer to the source.
Figure 4.7: Modelled and observed sources on 13 January 2013 18:00-21:00 at 0.188 Hz. a) Modelled source P (Pa²m²s). Colour bar above figure. b) Beamformer output from California array, c) Beamformer output from Japan array d) Beamformer output from Europe array. e) Combined beamformer output f) Combined beamformer output with distance weighting. All beamformer outputs are normalised to 0 dB with colour bar above (b).
Figure 4.8: Same as Figure 4.7 but for sources on 15 February 2013 06:00-09:00 at 0.188 Hz.
Figure 4.9: Same as Figure 4.7 but for sources on 10 October 2012 06:00-09:00 at 0.207 Hz.
Figure 4.10: Same as Figure 4.7 but for sources on 27 October 2012 06:00-09:00 at 0.207 Hz.
Figure 4.11: Same as Figure 4.7 but for sources on 19 November 2012 03:00-06:00 at 0.188 Hz.
Figure 4.12: Same as Figure 4.7 but for sources on 19 November 2012 12:00-15:00 at 0.188 Hz.
Figure 4.13: Same as Figure 4.7 but for sources on 19 January 2013 00:00-03:00 at 0.207 Hz.
Figure 4.14: Same as Figure 4.7 but for sources on 02 October 2012 21:00-24:00 at 0.207 Hz.
Figure 4.15: Same as Figure 4.7 but for sources on 05 October 2012 21:00-24:00 at 0.188 Hz.
4.5.3 Weighting by distance

The results of the distance weighting for the data are shown in plot (f) of Figures 4.7 to 4.15. As expected, the greatest improvement of the distance weighting can be seen for the sources close to the coast of Japan. In Figure 4.8 the weaker source observed by the Japan array and missing from the original combined image is now visible. In Figure 4.11 on 19 November 2012 03:00-06:00 the sources are resolved in the new combined image as well as they are in the individual Japan image. In Figure 4.13 for 19 January 2013 00:00-03:00 the source to the north originally only visible in the Japan image is much better resolved although still weaker than it should be, and the source to the east is also better resolved. There are artefacts close to Japan and in the southeast of the grid. In Figure 4.9 for 10 October 2012 06:00-09:00 a source gets separated out in the east much more clearly although the northern source is still missing and there are artefacts around the edges of array ranges. In Figure 4.14 for 02 October 2012 21:00-24:00 the two sources in the north become better separated, and in Figure 4.15 for 05 October 2012 21:00-24:00 the two sources close to Japan, one of which was previously only observed by the California array, become much better resolved. Therefore, by combining the arrays and using a distance based weighting, the benefits of each individual array were aggregated into one image which helped resolve and separate out multiple P-sources.

We also tested whether the location accuracy of the strongest source was improved by combining the arrays. For each event, the location of the maximum modelled source and the location of the corresponding beamformer peak was found. Because the strongest source may not create the strongest beamformer peak, and here we are interested in location, this corresponding beamformer peak was not necessarily the strongest but was chosen based on proximity to the modelled source location. For example, in Figure 4.13c the northern source was chosen to correspond with the largest modelled source shown in plot (a) even though the beamformer peak to the south was stronger. For each array and combined image, the distance between the modelled source and the corresponding beamformer peak was calculated in degrees (Table 4.1). Distances ranged between 0.0 and 16.6°. The largest distances were found when an array merged two separate sources into one averaged location, such as the California array on 19 January 2013 (distance = 10.0°) and the Japan array on 02 October 2012 (distance = 16.6°). The average distance offset was smaller for the combined and weighted combined images (3.9 and 3.6° respectively) than for the California and Japan arrays (each 4.6°). They were slightly larger than for the Europe array (3.2°) although on three events the source was out of range for the Europe array and for one event the Europe output was too noisy for a location to be determined. These errors are similar to the those expected from the use of a 1-d global velocity model (Euler et al., 2014).
### Table 4.1: Comparison of modelled and observed source locations

<table>
<thead>
<tr>
<th>Date</th>
<th>( f_2 ) (Hz)</th>
<th>( \lambda_P ) (°)</th>
<th>( \phi_P ) (°)</th>
<th>( \lambda_{ca} ) (°)</th>
<th>( \phi_{ca} ) (°)</th>
<th>( d_{ca} ) (°)</th>
<th>( \lambda_{jpn} ) (°)</th>
<th>( \phi_{jpn} ) (°)</th>
<th>( d_{jpn} ) (°)</th>
<th>( \lambda_{eu} ) (°)</th>
<th>( \phi_{eu} ) (°)</th>
<th>( d_{eu} ) (°)</th>
<th>( \lambda_{c} ) (°)</th>
<th>( \phi_{c} ) (°)</th>
<th>( d_{c} ) (°)</th>
<th>( \lambda_{w} ) (°)</th>
<th>( \phi_{w} ) (°)</th>
<th>( d_{w} ) (°)</th>
</tr>
</thead>
<tbody>
<tr>
<td>13-Jan-13 T7</td>
<td>0.188</td>
<td>37</td>
<td>193</td>
<td>33</td>
<td>195</td>
<td>4.0</td>
<td>n/a</td>
<td>n/a</td>
<td>n/a</td>
<td>n/a</td>
<td>n/a</td>
<td>n/a</td>
<td>n/a</td>
<td>n/a</td>
<td>n/a</td>
<td>n/a</td>
<td>n/a</td>
<td>n/a</td>
</tr>
<tr>
<td>15-Feb-13 T3</td>
<td>0.188</td>
<td>41</td>
<td>184</td>
<td>41</td>
<td>177</td>
<td>3.0</td>
<td>43</td>
<td>179</td>
<td>2.5</td>
<td>45</td>
<td>183</td>
<td>4.3</td>
<td>43</td>
<td>177</td>
<td>3.6</td>
<td>n/a</td>
<td>n/a</td>
<td>n/a</td>
</tr>
<tr>
<td>10-Oct-12 T3</td>
<td>0.207</td>
<td>35</td>
<td>163</td>
<td>35</td>
<td>165</td>
<td>1.6</td>
<td>37</td>
<td>165</td>
<td>2.6</td>
<td>n/a</td>
<td>n/a</td>
<td>n/a</td>
<td>n/a</td>
<td>n/a</td>
<td>n/a</td>
<td>n/a</td>
<td>n/a</td>
<td>n/a</td>
</tr>
<tr>
<td>27-Oct-12 T3</td>
<td>0.207</td>
<td>45</td>
<td>201</td>
<td>47</td>
<td>189</td>
<td>8.6</td>
<td>49</td>
<td>187</td>
<td>10.3</td>
<td>47</td>
<td>197</td>
<td>3.4</td>
<td>47</td>
<td>191</td>
<td>7.2</td>
<td>n/a</td>
<td>n/a</td>
<td>n/a</td>
</tr>
<tr>
<td>19-Nov-12 T2</td>
<td>0.188</td>
<td>49</td>
<td>151</td>
<td>47</td>
<td>147</td>
<td>3.3</td>
<td>51</td>
<td>151</td>
<td>2.0</td>
<td>51</td>
<td>149</td>
<td>2.4</td>
<td>49</td>
<td>155</td>
<td>1.3</td>
<td>51</td>
<td>151</td>
<td>2.0</td>
</tr>
<tr>
<td>19-Nov-12 T5</td>
<td>0.188</td>
<td>45</td>
<td>165</td>
<td>45</td>
<td>161</td>
<td>2.8</td>
<td>47</td>
<td>161</td>
<td>3.4</td>
<td>43</td>
<td>167</td>
<td>2.5</td>
<td>47</td>
<td>161</td>
<td>3.4</td>
<td>47</td>
<td>161</td>
<td>3.4</td>
</tr>
<tr>
<td>19-Jan-13 T1</td>
<td>0.207</td>
<td>51</td>
<td>163</td>
<td>41</td>
<td>165</td>
<td>10.0</td>
<td>51</td>
<td>163</td>
<td>0.0</td>
<td>n/a</td>
<td>n/a</td>
<td>n/a</td>
<td>n/a</td>
<td>n/a</td>
<td>n/a</td>
<td>n/a</td>
<td>n/a</td>
<td>1.3</td>
</tr>
<tr>
<td>02-Oct-12 T8</td>
<td>0.207</td>
<td>39</td>
<td>189</td>
<td>43</td>
<td>191</td>
<td>4.3</td>
<td>47</td>
<td>169</td>
<td>16.6</td>
<td>n/a</td>
<td>n/a</td>
<td>n/a</td>
<td>n/a</td>
<td>n/a</td>
<td>n/a</td>
<td>39</td>
<td>193</td>
<td>3.1</td>
</tr>
<tr>
<td>05-Oct-12 T8</td>
<td>0.188</td>
<td>47</td>
<td>165</td>
<td>45</td>
<td>161</td>
<td>3.4</td>
<td>47</td>
<td>161</td>
<td>2.7</td>
<td>45</td>
<td>161</td>
<td>3.4</td>
<td>45</td>
<td>157</td>
<td>5.9</td>
<td>45</td>
<td>157</td>
<td>5.9</td>
</tr>
</tbody>
</table>

Mean d 4.6 4.6 3.2 3.9 3.6

Maximum d 10.0 16.6 4.3 7.2 5.9

---

\( a \) The location of the maximum modelled source is given along with the location of the beamformer peak associated with this source. This is not necessarily the maximum valued beamformer peak in the output. The first two columns give the date, time step and frequency of the events in order as they are presented in Figures 4.7 to 4.15. Columns 3 and 4 give the latitude and longitude of the modelled source. Columns 5, 6 and 7 give the location of the California array beamformer output and the distance between this location and the modelled source location in degrees. Columns 8, 9 and 10 give the location and distance for the Japan beamformer output. Columns 11, 12 and 13 give the location and distance for the Europe beamformer output. n/a is shown where the maximum modelled source was outside the array range or (for 19-Jan-2013) when the output was too noisy to pick a location. Columns 14, 15 and 16 give the location and distance for the combined output and Columns 17, 18 and 19 give the location and distance for the weighted combined output. n/a in these columns is used when there was no clear peak in the beamformer output to pick. The last two rows give the mean distance calculated for each array/combination and the maximum distance calculated for each array/combination.

### 4.6 Discussion

Accurately locating P-wave microseismic sources is an important step in relating seismic records to oceanic and atmospheric conditions. The relationship has implications for using seismic records to monitor wave activity in real-time and track storms across the ocean. Previous studies have shown promise for such applications; P-wave microseisms have been located to specific storms such as Hurricane Katrina (Gerstoft et al., 2006), Super Typhoon Ioke (Zhang et al., 2010a) and tropical cyclone Dumile (Davy et al., 2014), as well as from regions of high wave activity associated with winter storms (Gerstoft et al., 2008; Euler et al., 2014; Neale et al., 2017).
Geographical coverage of an array is determined by its location, and constraints of array aperture, configuration and number of stations limit the ability of a single array to resolve and accurately locate microseismic sources. We have shown examples where neighbouring sources become merged or sources are missed entirely by individual arrays, whereas other arrays resolve them.

With a view to using microseisms to monitor ocean wave activity and seismic sources, a single image that collates the information from each array would be more suitable than looking at many images from individual arrays. We therefore combined the stations from each array into one image and adjusted the weighting of each station so that those closest to each grid point were weighted higher, which is a more physical model. The previous studies already mentioned have combined data from multiple arrays but with notable differences from our study, especially in the temporal resolution. Landès et al. (2010) combined three arrays located in the USA, Turkey and Kyrgyzstan to locate P-wave microseisms at 0.1-0.3 Hz, but with the aim of improving the accuracy of the main global sources over seasonal timescales. They combined their arrays by multiplying the individual output of each array. A comparison of the sources with a wave-interaction model was not attempted until the later study by (Hillers et al., 2012) and this again looked at sources over 13-day to seasonal time periods. Euler et al. (2014) used four arrays in Africa to locate P-wave sources at 0.1-0.2 Hz, but the main difference here was that none of the arrays were deployed at the same time, but were each deployed over a 1-2 year period between 1994 and 2007. They manually picked peaks in the output for each array and plotted these onto one map to find persistent sources over the entire 13 year period. They compared these source locations with a map of bathymetric excitation coefficients rather than with sources modelled from a combination of wave-wave interaction and bathymetry. Other studies have focussed on higher frequency P-waves using small aperture arrays. Koper et al. (2010) used multiple arrays with apertures of 2-28 km to locate microseisms with frequencies of 0.4-4 Hz. They picked only the maxima of each array output averaged over one year, and again plotted the P-wave locations on one map to find a common source. Pyle et al. (2015) combined arrays in Asia, Australasia and North America with apertures of 10-25 km to find robust locations of P-wave sources at 0.67-1.33 Hz. They used a ranking scheme and histograms to locate sources seen by all arrays over a whole month, and compared with maps of significant wave height. Our study builds on this previous work by directly comparing observed sources at three different arrays with modelled P-wave sources (instead of significant wave height) over shorter (3-hour) time periods, with a focus on how multiple arrays can be used to help reveal multiple neighbouring sources.

In some cases, sources that were visible by a single array were not visible or resolved strongly in the combined image, and east-west separated sources were not well
resolved. This was especially true for the Europe array, which had fewer stations and was often overpowered by the California and Japan arrays and was further reduced by the distance weighting. Because the Europe array views the North Pacific from a perpendicular viewpoint with respect to the California and Japan arrays, the influence of the Europe data could be increased by using a weighting that is dependent on azimuth as well as distance. In addition, an algorithm such as the CLEAN algorithm used by Gal et al. (2016), which iteratively removes the array response associated with a point source, may offer further improvement on the separation of weaker sources from the stronger sources. This method is dependent on the dominant source, so by itself cannot correctly separate two closely spaced sources when the beamformer output is centred between the two sources. However, after separating or partly separating the two sources by combining multiple arrays as we have shown here, the algorithm may help separate out and resolve the sources better. Other methods that could be used may include high resolution beamforming techniques such as sparse beamforming (Elad, 2010), MUSIC (Van Trees, 2002) and MVDR (Capon, 1969) used in ocean acoustics. The inclusion of more stations and arrays is expected to improve the ability to separate closely spaced sources even further. For this feasibility study, we used only a subset of 202 of the 800+ stations of Japan’s Hi-net network. This subset over the Chugoku district was chosen to match that used by Nishida and Takagi (2016) because of weak crustal heterogeneity there, and because the number of stations closely matches the number of stations in the California (242) and Europe (186) arrays. The optimal number and location of stations to use in the process would need to be the subject of future work. Improvement of source location accuracy would be expected if crustal and mantle structure were accounted for in the back-projection. Euler et al. (2014) found that accounting for heterogeneities corrected source locations by 0-4° (mostly below 2°) which may reduce our current offsets of about 3.2-4.6° between observed and modelled sources (Table 4.1). Additionally, a better comparison between amplitudes of multiple sources would be obtained if amplitude information in the seismic displacement spectra was retained, and backprojecting phases such as PP and PKP may provide further information and increased coverage.

4.7 Conclusions

We used three large seismic arrays to image P-wave microseism sources in the North Pacific on occasions when multiple sources were present at any given 3-hour time period. The three arrays used here often gave quite different pictures of the sources acting in the North Pacific at any given time. The Europe array was best at resolving neighbouring sources in the very north and those separated in an east-west direction. The California array was generally good at resolving sources in the central regions of
the North Pacific but was not good at separating closely spaced sources aligned east-west and had poor resolution for sources in the west of the North Pacific near Japan. The Japan array resolved sources best in the western half of the North Pacific and imaged to exceptional resolution sources to the east of Japan and north into the Sea of Okhotsk. Therefore no single array gave the best representation of sources and a fuller understanding of the sources acting in the basin can be obtained by using data from multiple different arrays placed at different locations around the basin. Most studies, with a few exceptions such as Koper et al. (2010), Landès et al. (2010), Hillers et al. (2012), Euler et al. (2014) and Pyle et al. (2015), currently use only a single array to locate P-wave sources and would benefit from including data from other arrays.

We combined the data from each array into one beamforming image. In most cases the combined images were found to give a comprehensive view of sources observed from each individual array. Especially after weighting by distance, the combined images were mostly as good or better than the best single image. The combined images offered improved coverage and revealed source locations that would have been missed by looking at output from only one array. In addition to being able to better resolve multiple sources, the combined images gave on average a more accurate location of the main source. The mean offset (arc-length in degrees) between the maximum modelled source location and the beamformer location was 4.6° for the California array, 4.6° for the Japan array, 3.2° for the Europe array, 3.9° for the array combination and 3.6° for the array combination with distance weighting. The greatest improvements in accuracy occurred when a single array merged two separate sources to an average location, while the multiple arrays separated them.
Chapter 5

Conclusions and Future Work

5.1 Conclusions

This thesis aimed to enhance our understanding of the ocean wave sources of ambient seismic noise. In particular, where the ocean wave noise sources are located and how well we can locate them, and whether we can relate seismic observations to useful indications of ocean wave activity. The focus was on infragravity waves (Chapter 2), which are the source of the low-frequency component of seismic noise known as Earth’s seismic hum, and ocean swell waves in the deep ocean which interact to create seismic noise at twice their frequency known as ‘double-frequency’ microseisms (Chapters 3 and 4).

Infragravity waves in deep water (‘leaky waves’) have been poorly studied partly due to the difficulties arising from their very low wave heights. However interest in their sources, propagation and seasonal characteristics has arisen because of their potential importance in the breaking up of Antarctic ice shelves (Bromirski et al., 2010), ocean and atmosphere coupling (Livneh et al., 2007; Godin et al., 2015) and contribution of error to future altimeter measurements of sea level (Aucan and Ardhuin, 2013). The pressure fluctuations beneath infragravity waves are recorded on pressure gauges and seismometers on the seabed. In Chapter 2, the source regions of infragravity waves propagating over a pressure gauge array in deep water offshore California were investigated. The dominant source region was found to be the coastline to the east (the western coast of the USA), in line with previous studies that associated strong infragravity wave sources with high wave heights and swells incident on eastern coasts (Rawat et al., 2014). During winter months the waves came from the south, consistent with seasonal changes in wave activity in each hemisphere. Occasional arrivals from the west provided a unique opportunity to estimate the reflection of infragravity waves at the coastline. Strong reflection has implications for infragravity wave energy in the
deep ocean, as infragravity waves propagate with little attenuation (Godin et al., 2013). Previous estimates of the reflection coefficient at the shoreline had been made in shallow water (Herbers et al., 1995b; Sheremet, 2002; De Bakker et al., 2014) and from laboratory data (Battjes, 2004) but not for deep water infragravity waves which may get reflected at the continental shelf. The estimates found here for deep water infragravity waves were made using the asymmetry of cross-correlation functions between pairs of pressure gauges aligned perpendicular to the coast. Reflection coefficients of 0.49-0.74 compared favourably with the previous estimates of reflection at the shoreline in shallow water, and are considerably higher than reflection coefficients for wind waves and swell.

Sources of double-frequency P-wave microseisms in the North Pacific Ocean were investigated in Chapter 3 using land seismometers in California and an ocean wave model. P-wave microseisms were studied instead of the more dominant surface Rayleigh waves because it is possible locate P-wave sources to a point on the globe using array observations, which is important for relating seismic records to wave activity across the ocean basin. Observed P-wave sources were compared with those modelled using the ocean wave model on short (3-hour) time steps. Locations of observed and modelled sources matched by 10° or less in 90% of cases, and were concentrated around 40-50°N, 160-180°E, which is consistent with previous observations of large events (Obrebski et al., 2013) and can be related to storm tracks in the North Pacific. Sources at ~0.2 Hz were found to dominate the observations and modelled sources. A regression between modelled source power and observed beampower allowed a time series of source strength to be estimated from the seismic observations. However, for purposes of using P-wave microseisms as a proxy for wave conditions, it is the relationship between observed P-waves and ocean wave parameters such as significant wave height that is important. The relationship is complicated because large ocean waves are not by themselves sufficient to generate strong microseisms, but require an opposing wavefield and favourable ocean depths for resonant amplification. Previous studies found links between sources of P-wave microseisms and broad regions of high significant wave heights over long averaging times (Euler et al., 2014) but the possibility of being able to infer ocean wave heights from P-waves over shorter timescales remained unclear. Here, the ocean wave energy spectrum was reconstructed from the estimated source power by dividing out the site amplification effect at the observed source location and assuming a constant spectrum for the degree of the wavefield opposition. Significant wave height was calculated from the reconstructed spectrum and compared to significant wave heights of the ocean wave model. After applying a calibration factor (calculated from the offset between modelled and estimated significant wave height in the first year), estimated and modelled significant wave heights correlated with a coefficient of 0.63. Results suggested that the estimation of source power from beampower was the largest source
of error. With improvements in this estimation, correlations of up to 0.78 would be possible even with the assumption about the unknown wavefield opposition.

The study in Chapter 3 revealed occasions when closely-spaced modelled P-wave sources could not be resolved from the seismic observations. This limitation had been recognised previously by Hillers et al. (2012) and Euler et al. (2014). The resolving capability was expected to be improved by combining observations from multiple seismic arrays and this was investigated in Chapter 4. Previous studies had combined arrays, but only on longer averaging timescales with the aim of finding persistent and common sources, rather than for revealing finer details on shorter timescales. In Chapter 4, seismic networks in Japan and Europe were used in addition to the array in California to locate double-frequency P-wave sources in the North Pacific. It was found that each gave a unique but mostly imperfect view of the sources at any 3-hour time step. Often, closely-spaced sources that were merged in the output of one array could be resolved in the output from another. The Europe array was best at resolving sources separated in an east-west direction, whereas the Japan array performed best for sources in the western North Pacific and the California array was generally best for sources in the central North Pacific. Combining data from all the arrays improved geographical coverage and incorporated the advantages of the individual arrays in resolving sources within one image.

These three chapters therefore addressed the problems set out at the beginning of the thesis; to locate ocean wave sources of seismic noise, to determine how well they can be located, and how they relate to ocean wave conditions and characteristics. The investigations also highlighted a number of areas where further research would be beneficial. Some ideas for further work are given in the following section.

5.2 Future Work

In each chapter (2-4) ideas about further work that could improve upon the results were discussed. In this section the possible next steps, identified as the most promising or interesting of these ideas, are outlined.

The estimate of the coastal reflection coefficient of infragravity waves was taken from just one array on the eastern side of the North Pacific. The nearby coastline was found to be the dominant source of infragravity waves propagating over the array, which meant that most infragravity energy propagated away from the coast. There were only five days in the dataset when arrivals propagating towards the coast coincided with low coastal wave heights (and therefore low amounts of local infragravity wave generation) from which estimates of reflection could be made. A better opportunity for estimating reflection of remote arrivals would be expected at an array on the western
side of the ocean basin, where remote infragravity wave arrivals would be common. A search of publicly available seismic data from deployments of OBSIP instruments (Ocean Bottom Seismograph Instrument Pool, \url{http://www.obsip.org/experiments/experiment-map/}) highlights a temporary array to the east of Taiwan deployed in 2008-2009 (the TAIwan Integrated GEodynamic Research (TAIGER) experiment (Okaya et al., 2006)) and the Eastern North America Community Seismic Experiment (Gaherty et al., 2014) as potential candidates for some extra seismic and pressure data on the eastern sides of the Pacific and Atlantic respectively. A deployment off New Zealand in 2009-2010 for the Marine Observations of Anisotropy Near Aotearoa (MOANA) Seismic Experiment, which has previously been used to study infragravity waves (Godin et al., 2013, 2014) may not be so suitable for estimates of reflection despite remote arrivals from the northeast Pacific being a significant contribution to the infragravity wavefield southeast of the South Island, because the propagation direction of these arrivals is along the coastline (Godin et al., 2014). A model of infragravity wave generation and propagation, such as used by Ardhuin et al. (2014) and Rawat et al. (2014), would need to be used to find the most suitable sites for deployments to measure perpendicular infragravity wave arrivals. Another way forward, and perhaps more feasible with respect to current data availability, would be to vary the reflection coefficient at the coast used in the wave model of Ardhuin et al. (2014) and compare the modelled infragravity energy at deep water sites with observed energy. This would not require arrays but only single stations. Rawat et al. (2014) compared the modelled wave heights with observed wave heights in this way to track infragravity waves as they propagated across the Pacific. Focusing on shoreward arrivals and their reflections may give some indication as to whether the model overestimates or underestimates the reflected energy, and the coefficient could then be adjusted and compared for different coastlines.

The findings of Chapter 3 indicated that the estimation of source power from beampower was the largest source of error in estimating significant wave height in the deep ocean from P-wave microseisms. A clear direction for future work therefore is improving this estimation. In Chapter 3, amplitudes of the observed ground displacement spectra were normalised by dividing by the magnitude. This was necessary to get a good correlation between the observed beampower and modelled sources, but introduces an error because the spectrum will be normalised by not only P-wave energy, but by all other microseismic waves at this frequency, of which Rayleigh waves are likely to be important. Rayleigh waves are known to be dominated by coastal sources and therefore unrelated to the P-wave amplitudes. A better way of estimating the source amplitudes, instead of simply using a regression between observed beampower and modelled source power, may be to use a similar approach of Nishida and Takagi (2016) and Farra et al. (2016), using non-normalised displacement spectra. In this method, ground displacement at a station is forward modelled from a
given vertical ocean microseism source accounting for the ocean site effect, geometric spreading and energy loss at discontinuities. In addition, a 3-D velocity model and individual station corrections to account for local seismic structure could be incorporated to enhance the source location. The value of the source can then be estimated by minimising the squared difference between observed and estimated ground displacements. Therefore, following from Chapter 3, the beampower would be used as it was to find the locations of the strong ocean sources, and then the source amplitudes would be estimated assuming vertical sources at each of those locations. To improve the source location, the work of Chapter 4 could be expanded to include more stations. This would be most easily achieved by including more of the available stations in North America and Japan. The improved source locations could then be used to estimate source amplitudes using the method described above. This should improve the ability to estimate significant wave heights from P-wave microseisms. Furthermore, this work could be applied globally by extending it to the Atlantic and Indian oceans. Another potential future direction of the work is the incorporation of streamed seismic data into a real-time monitoring system to produce information useful for the shipping industry. My work has shown that simple empirical relationships can be developed between the seismic data and significant wave height, dominant wave period and wave interaction, however more sophisticated techniques from machine learning (such as those under development for surface wave microseisms in Ireland (Donne et al., 2014)) might enhance probabilistic determination of likely ocean wavefield characteristics based on historical seismic data and wave action models.

This thesis has explored various avenues for obtaining infragravity and gravity wave observations from seismic and ocean bottom measurements. While some promising results have been obtained and potential ways of improving the results discussed, obtaining accurate wave observations on short timescales from seismic data continues to be a challenging problem.
Appendix A

Section A.1 gives the input files used for the Fast Marching Surface Tomography Package. Section A.2 gives the switch file used for compiling WAVEWATCH III and A.3 gives the input files used.
Appendix A

A.1 Input files for Fast Marching Surface Tomography Package

A.1.1 fm2dss.in

cccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccc
cccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccc
c INPUT PARAMETERS
cccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccc
sources.dat                    c: File containing source positions
receivers.dat                  c: File containing receiver positions
otimes.dat                     c: File containing source-receiver associations
gridi.vtx                      c: File containing velocity grid information
1     1                        c: Grid dicing in latitude and longitude
1                              c: Apply source grid refinement? (0=no,1=yes)
1    2                        c: Dicing level and extent of refined grid
6371.0                         c: Earth radius in km
1                              c: Use first-order(0) or mixed-order(1) scheme
0.5                            c: Narrow band size (0-1) as fraction of nrx*nrz
cccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccc
c OUTPUT FILES
cccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccc
1                              c: find source-receiver traveltimes (0=no,1=yes)
rtravel.out                    c: Name of file containing source-receiver traveltimes
1                              c: Calculate Frechet derivatives (0=no,1=yes)
frechet.out                    c: Name of file containing Frechet derivatives
2                              c: Write traveltime field to file? (0=no,>0=source id)
travelt.out                    c: Name of output file containing traveltime field
11                            c: Write out raypaths (<0=all,0=no,>0=source id)
raypath.out                    c: Name of file containing raypath geometry
A.1.2 tslicess.in

c Input file names
gridi.vtx c: True model velocity vertices
gridi.vtx c: Reference model velocity vertices
travelt.out.011 c: Traveltime field
raypath.out c: Raypaths

Bounding box file
bound.gmt c: GMT plotting bounds for depth slice

Plot sources and receivers
sources.dat c: Source input file
sources.dat c: Source output file
receivers.dat c: Receiver input file
receivers.dat c: Receiver output file

Velocity grid parameters
1 c: Plot velocity slice? (0=no,1=yes)
0 c: Absolute (0) or relative (1) velocity
1 1 c: Dicing in theta, phi
grid2dv.z c: GMT output file for velocity slice

Traveltime grid parameters
1 c: Plot traveltime slice? (0=no,1=yes)
grid2dt.z c: GMT output for traveltime slice

Ray path parameters
1 c: Plot raypaths? (0=no,1=yes)
rays.dat c: Output ray path file for depth slice
A.2 Switch file for WAVEWATCH III

A.2.1 Switch file

F90 DIST MPI LRB4 NOPA PR3 UQ FLX0 LN1 ST4 NL1 BT4 IC0 IS0
REF1 DB1 TR0 BS0 XX0 WNT1 WNX1 CRT1 CRX1 NOGRB O0 O1 O2 O3
O4 O5 O6 O7 MLIM NC4
A.2.2 Description of switch file

Selection of machine dependent code:

F90 FORTRAN-90 style data and time capturing and program abort.

Hardware model and message passing protocol:

DIST Distributed memory model.

MPI Message Passing Interface (MPI).

Word length used to determine record length in direct access files:

LRB4 4 byte words.

Compilation as a subroutine or stand-alone program:

NOPA Compilation as a stand-alone program.

Selection of propagation schemes and GSE alleviation method:

PR3 Higher-order schemes with Tolman (2002a) averaging technique.

UQ Third-order propagation scheme.

Selection of flux computation:

FLX0 No routine used; flux computation included in source terms.

Selection of linear input:

LN1 Cavaleri and Malanotte-Rizzoli with filter.

Selection of input and dissipation:

ST4 Arduhin et al. (2010) source term package.

Selection of nonlinear interactions:

NL1 Discrete interaction approximation (DIA).

Selection of bottom friction:

BT4 SHOWEX bottom friction formulation.

Selection of term for damping by sea ice:

IC0 No damping by sea ice.

Selection of term for scattering by sea ice:

IS0 No scattering by sea ice.

Selection of term for reflection:

REF1 Enables reflection of shorelines and icebergs.

Selection of depth-induced breaking of:

DB1 Battjes-Janssen.

Selection of triad interactions:

TR0 No triad interactions used.

Selection of bottom scattering:

BS0 No bottom scattering used.

Selection of supplemental source term:

XX0 No supplemental source term used.
Selection of method of wind interpolation (time):
  WNT1 Linear interpolation.

Selection of method of wind interpolation (space):
  WNX1 Approximately linear speed interpolation.

Selection of method of current interpolation (time):
  CRT1 Linear interpolation.

Selection of method of current interpolation (space)
  CRX1 Approximately linear speed interpolation.

Switch for user supplied GRIB package:
  NOGRB No package included.

Optional switches for output of WAVEWATCH III programs:
  O0 Output for namelists in grid preprocessor.
  O1 Output of boundary points in grid preprocessor.
  O2 Output of the grid point status map in grid preprocessor.
  O3 Additional output in loop over fields in field preprocessor.
  O4 Print plot of normalised one-dimensional energy spectrum
     in initial conditions program
  O5 Id. two-dimensional energy spectrum.
  O6 Id. spatial distribution of wave heights (not adapted for
     distributed memory).
  O7 Echo input data for homogeneous fields in generic shell.

Miscellaneous switches:
  MLIM Use Miche-style shallow water limiter.
  NC4 Activates the NetCDF-4 API in the NetCDF pre- and post-
     processing programs.
A.3 Input files for WAVEWATCH III

A commented input file is given for each input file type. The comments describe the input options and are from the example input files given in the WAVEWATCH III manual (WW3DG, 2016).
### A.3.1 ww3.grid.inp.glob: For grid preprocessor (Global 0.5 degree grid) [with comments]

```
$ WAVEWATCH III Grid preprocessor input file
$ Grid name (C*30, in quotes)
$ 'Global 0.5deg grid' 
$ Frequency increment factor and first frequency (Hz) ---------------- 
$ number of frequencies (wavenumbers) and directions, relative offset 
$ of first direction in terms of the directional increment [-0.5,0.5]. 
$ In versions 1.18 and 2.22 of the model this value was by definition 0, 
$ it is added to mitigate the GSE for a first order scheme. Note that 
$ this factor is IGNORED in the print plots in ww3_outp.
$ 1.10 0.83 32 36 0.
$ Set model flags ---------------------------------------------------- 
$ - FLDRY         Dry run (input/output only, no calculation).
$ - FLCX, FLCY    Activate X and Y component of propagation.
$ - FLCTH, FLCK   Activate direction and wavenumber shifts.
$ - FLSSO        Activate source terms.
$ F T T T T T
$ Set time steps ----------------------------------------------------- 
$ - Time step information (this information is always read)
$   maximum global time step, maximum CFL time step for x-y and 
$   k-theta, minimum source term time step (all in seconds).
$ 1800. 480. 900. 15.
$ Start of namelist input section ------------------------------------ 
$   Starting with WAVEWATCH III version 2.00, the tunable parameters
$   for source terms, propagation schemes, and numerics are read using
$   namelists. Any namelist found in the following sections up to the
$   end-of-section identifier string (see below) is temporarily written
$   to ww3_grid.scratch, and read from there if necessary. Namelists
$   not needed for the given switch settings will be skipped
$   automatically, and the order of the namelists is immaterial.
$   As an example, namelist input to change SWELLF and ZWND in the
$   Tolman and Chalikov input would be
$   &SIN2 SWELLF = 0.1, ZWND = 15. /
$ Define constants in source terms -----------------------------------
$ Stresses - - - - - - - - - - - - - - - - - - - - - - - - - - - - - -
$   TC 1996 with cap    : Namelist FLX3
$                       CDMAX  : Maximum allowed CD (cap)
$                       CTYPE  : Cap type :
$                       0: Discontinuous (default).
$                       1: Hyperbolic tangent.
$   Hwang 2011          : Namelist FLX4
$                       CDFAC  : re-scaling of drag
$ Linear input - - - - - - - - - - - - - - - - - - - - - - - - - - - -
$   Cavaleri and M-R    : Namelist SLN1
$                       CLIN   : Proportionality constant.
$                       RFPM   : Factor for fPM in filter.
$                       RFHF   : Factor for fh in filter.
$ Exponential input  - - - - - - - - - - - - - - - - - - - - - - - - - -
$   WAM-3               : Namelist SIN1
$                       CINP   : Proportionality constant.
$   Tolman and Chalikov : Namelist SIN2
$                       ZWDO  : Height of wind [m].
$                       SWELLF : swell factor in [n.nn].
$                       STABSH, STABOF, CNEG, CPOS, FNEG :
$                       c0, ST0, c1, c2 and f1 in . (n.nn)
$                       through (2.65) for definition of
$                       effective wind speed (!/STAB2).
$   WAM4 and variants  : Namelist SIN3
$                       ZWDO  : Height of wind [m].
$                       ALPHAW : minimum value of Charnock coefficient
$                       ZBMAKX : maximum value of air-side roughness zB
$                       BETAMAX : maximum value of wind-wave coupling
$                       SINTHMP : power of cosine in wind input
$                       ZALP  : wave age shift to account for gustiness
$                       TAUSHIELTER : sheltering of short waves to reduce u_star
$                       SWELLPMAR : choice of swell attenuation formulation
$                       SWELLF : swell attenuation factor
$                       Extra parameters for SWELLPMAR3 only
$                       SWELLF2, SWELLF3 : swell attenuation factors
$                       SWELLF4 : threshold Reynolds number for ACC2008
$                       SWELLF5 : Relative viscous decay below threshold
$                       ZDRAF : roughness for oscil. flow / mean flow
$ BYDRZ input         : Namelist SIN6
```
$SINA0$: factor for negative input

$SINU10$: wind speed scaling option

Nonlinear interactions

Discrete I.A.: Namelist SNL1

- $\Lambda$: Lambda in source term.
- NLPROP: C in source term. NOTE: default value depends on other source terms selected.
- KDCONV: Factor before $kd$ in Eq. (n.nn).
- KDMIN, SNLCS1, SNLCS2, SNLCS3: Minimum $kd$, and constants c1-3 in depth scaling function.

Exact interactions: Namelist SNL2

- IQTYPE: Type of depth treatment
  1: Deep water
  2: Deep water / WAM scaling
  3: Shallow water
- TAILNL: Parametric tail power.
- NDEPTH: Number of depths in for which integration space is established.

- Used for IQTYPE = 3 only

Gen. Multiple DIA: Namelist SNL3

- NQDEF: Number of quadruplets.
- MSC: Scaling constant 'm'.
- NSC: Scaling constant 'N'.
- KDFD: Deep water relative filter depth.
- KDFS: Shallow water relative filter depth.

- Namelist ANL2
  DEPTHS: Array with depths for NDEPTH = 3

- Namelist ANL3
  QPARMS: 5 x NQDEF parameters describing the quadruplets, repeating $\Lambda$, $\mu$, $DT_{12}$.

- Cdeep and Cshal. See examples below.

Traditional DIA setup (default):

- $SNL3$ NQDEF = 1, MSC = 0.00, NSC = -3.50 /
- $ANL3$ QPARMS = 0.250, 0.000, -1.0, 0.1000E+08, 0.0000E+00 /

GMD3 from 2010 report (G13d in later paper):

- $SNL3$ NQDEF = 3, MSC = 0.00, NSC = -3.50 /
- $ANL3$ QPARMS = 0.126, 0.000, -1.0, 0.4790E+08, 0.0000E+00 ,
  0.319, 0.000, -1.0, 0.1118E+08, 0.0000E+00 /

G35d from 2010 report:

- $SNL3$ NQDEF = 5, MSC = 0.00, NSC = -3.50 /
- $ANL3$ QPARMS = 0.066, 0.018, 21.4, 0.170E+09, 0.000E+00 ,
  0.127, 0.069, 19.6, 0.127E+09, 0.000E+00 ,
  0.228, 0.065, 2.6, 0.443E+08, 0.000E+00 ,
  0.265, 0.196, 40.5, 0.218E+09, 0.000E+00 ,
  0.369, 0.226, 11.5, 0.118E+08, 0.000E+00 /

Nonlinear filter based on DIA

Dissipation

- WAM-3: Namelist SDS1
  CDIS, APM: As in source term.

- Tolman and Chalikov: Namelist SDS2
  $SDSA$, $SDSA1$, $SDSA2$, $SDSB$, $SDSB1$, PHI_MIN :
  Constants $a_0$, $a_1$, $a_2$, $b_0$, $b_1$ and PHI_MIN.

- WAM4 and variants: Namelist SDS3
  $SDSC1$: WAM4 Cds coefficient
  $MMWAVE$, $MMWAVEPTAIL$: power of wavenumber
  for mean definitions in SDS and tail
  $SDSDelta1$, $SDSDelta2$: relative weights
  of k and k^2 parts of WAM4 dissipation
  $SDSFL$, $SDSHF$: coefficient for activation of WAM4 dissipation for unsaturated (SDSFL) and saturated (SDSHF) parts of the spectrum
  $SDSC2$: Saturation dissipation coefficient
  $SDSC4$: Value of $B=BB/Br$ for wich Sds is zero
  $SDSB1$: Threshold Br for saturation
  $SDSB2$: power of (BB/Br-B0) in Sds
  $SDSC5$: coefficient for turbulence dissipation
  $SDSC6$: Weight for the istropic part of Sds_SAT

Appendix A
SDSDTH: Angular half-width for integration of B

BYDRZ : Namelist SDS6
SDSET : Select threshold normalization spectra
SDS1, SDS2, SDS3, SDS4, SDS5 : Coefficients for dissipation terms T1 and T2
: Namelist SWL6
SWL6 : Coefficient for swell dissipation

Bottom friction

JONSWAP : Namelist SBT1
GAMMA : As it says.

Surf breaking

Battjes and Janssen : Namelist SDB1
BJALFA : Dissipation constant (default = 1)
BJGAM : Breaking threshold (default = 0.73)
BJFLAG : TRUE - Use Hmax/d ratio only (default)
FALSE - Use Hmax/d in Miche formulation

Triad nonlinear interactions

Lumped Triad Interaction (LTA) : Namelist STR1 (To be implemented)
PTRIAD1 : Proportionality coefficient (default 0.05)
PTRIAD2 : Multiple of Tm01 up to which interaction is computed (2.5)
PTRIAD3 : Ursell upper limit for computing interactions (not used, default 10.)
PTRIAD4 : Shape parameter for biphase computation (8.2)
PTRIAD5 : Ursell number threshold for computing interactions (0.01)

Shoreline reflections

ref. parameters : Namelist REF1
REFCOAST : Reflection coefficient at shoreline
REFFREQ : Activation of freq-dependent ref.
REFMAP : Scale factor for bottom slope map
REFMAX : maximum ref. coeff. (default 0.8)
REFREFPOL : power of frequency
REFICEBERG : Reflection coefficient for icebergs
REFSUBGRID : Reflection coefficient for islands
REFCOSP_STRAIGHT : power of cosine used for straight shoreline

Bound 2nd order spectrum and free IG

IG parameters : Namelist IGI
IGMETHOD : 1: Hasselmann, 2: Krasitskii-Janssen
IGADDSW : activation of bound wave correction
in ww3_outp / ww3_pump
IGSOURCE : 1: uses bound waves, 2: empirical
IGSTREFS : > 0 : no source term in IG band
IGMAXFREQ : maximum frequency of IG band
IGEMPIRICAL: constant in empirical free IG source
IGWELLMAX : activates free IG sources for all freq.

Propagation schemes

SMC grid propagation : Namelist PSMC
CFLTM : Maximum CFL number for propagation.
DTIME : Swell age (s) for diffusion term.
LATMIN : Maximum latitude (deg) for GCT.
RFMAXD : Maximum refraction turning (deg).
APSMC DTIME = 39600.0, LATMIN=85.0, RFMAXD = 36.0

UQ/UNO with diffusion : Namelist PRO2
CFLTM : Maximum CFL number for diffusion.
DTIME : Swell age (s) in garden sprinkler correction. If 0., all diffusion switched off. If small non-zero (DEFAULT !!!) only wave growth diffusion.
LATMIN : Maximum latitude used in calc. of strength of diffusion for prop.

UQ/UNO with averaging : Namelist PRO3
CFLTM : Maximum CFL number for refraction.
WDTCH : Tuning factor propag. direction.
WDTHTH : Tuning factor normal direction.

Note that UQ and UNO schemes have no tunable parameters.
All tuneable parameters are associated with the refraction limitation and the GSE alleviation.

Unstructured grids
$ UNST parameters : Namelist UNST
$ UGOBCAUTO : Automatic detection of OBC points
$ UGOBDEPTH : Threshold ( < 0) depth for OBC points
$ EXPFSN : Activation of N scheme
$ EXPFSPSI : Activation of PSI scheme
$ EXPFSFCT : Activation of FCT scheme
$ IMPFSNIMP : Activation of N implicit scheme
$ SMC grid propagation : Namelist PSMC
$ CFLTM : Maximum CFL number for propagation.
$ DTIME : Swell age (s) for diffusion term.
$ LATMIN : Maximum latitude (deg) for GCT.
$ RFMAXD : Maximum refraction turning (deg).
$ &PSMC DTIME = 39600.0, LATMIN=85.0, RFMAXD = 36.0 /
$ Miscellaneous ------------------------------------------------------ $
$ Misc. parameters : Namelist MISC
$ CICE0 : Ice concentration cut-off.
$ CICEN : Ice concentration cut-off.
$ PMOVE : Power p in GSE aleviation for moving grids in Eq. (D.4).
$ XSEED : Xseed in seeding alg. (1/SEED).
$ FLAGTR : Indicating presence and type of subgrid information :
  0 : No subgrid information.
  1 : Transparancies at cell boudaries between grid points.
  2 : Transp. at cell centers.
  3 : Like 1 with cont. ice.
  4 : Like 2 with cont. ice.
$ XP, XR, XFILT
  Xp, Xr and Xf for the dynamic integration scheme.
$ INMAX : Number of discrete levels in part.
$ HSPMIN : Minimum Hs in partitioning.
$ WSM : Wind speed multiplier in part.
$ WSC : Cut of wind sea fraction for identifying wind sea in part.
$ FLC : Flag for combining wind seas in partitioning.
$ NOSW : Number of partitioned swell fields in field output.
$ FMICHE : Constant in Miche limiter.
$ P2SF : ......
$ In the 'Out of the box' test setup we run with sub-grid obstacles
and with continuous ice treatment.
$ GMISC CICE0 = 0.25, CICEN = 0.75, FLAGTR = 4 /
$ &FLX3 CDMAX = 3.5E-3, CTYPE = 0 /
$ &SDB1 BJGAM = 1.26, BJFLAG = .FALSE. /
$ &REF1 REFCOAST=0.10, REFFREQ=1., REFMAP=0., REFICEBERG=0.4 /
$ &MISC CICE0 = 0.25, CICEN = 0.75, FLAGTR = 4 /
$ &OUTS P2SF  = 1, I1P2SF = 1, I2P2SF = 32,
  E3D   = 1, I1E3D  =  1, I2E3D  = 32 /
$ Mandatory string to identify end of namelist input section.
$ END OF NAMELISTS
$ Define grid -------------------------------------------------------- $
$ Five records containing :
$ 1 Type of grid, coordinate system and type of closure: GSTRG, FLAGLL,
$   CSTRG. Grid closure can only be applied in spherical coordinates.
$ GSTRG : String indicating the type of grid :
  'RECT' : rectilinear
  'CURV' : curvilinear
  'UNST' : unstructured (triangle-based)
$ FLAGLL : Flag to indicate coordinate system :
  T : Spherical (lon/lat in degrees)
  F : Cartesian (meters)
$ Five records containing :
$ 2 NX, NY. As the outer grid lines are always defined as land
$ 2 NX, NY. As the outer grid lines are always defined as land
points, the minimum size is 3x3.

Branch here based on grid type

IF ( RECTILINEAR GRID ) THEN

3 Grid increments SX, SY (degr. or m) and scaling (division) factor.
4 Coordinates of (1,1) (degr.) and scaling (division) factor.

ELSE IF ( CURVILINEAR GRID ) THEN

3 Unit number of file with x-coordinate.
Scale factor and add offset: x = scale_fac * x_read + add_offset.
IDLA, IDFM, format for formatted read, FROM and filename.

    IDLA : Layout indicator :
    1 : Read line-by-line bottom to top.
    2 : Like 1, single read statement.
    3 : Read line-by-line top to bottom.
    4 : Like 3, single read statement.

    IDFM : format indicator :
    1 : Free format.
    2 : Fixed format with above format descriptor.
    3 : Unformatted.

    FROM : file type parameter
    'UNIT' : open file by unit number only.
    'NAME' : open file by name and assign to unit.

If the above unit number equals 10, then the x-coord is read from this
file. The x-coord must follow the above record. No comment lines are
allowed within the x-coord input.

4 Unit number of file with y-coordinate.
Scale factor and add offset: y = scale_fac * y_read + add_offset.
IDLA, IDFM, format for formatted read, FROM and filename.

    IDLA : Layout indicator :
    1 : Read line-by-line bottom to top.
    2 : Like 1, single read statement.
    3 : Read line-by-line top to bottom.
    4 : Like 3, single read statement.

    IDFM : format indicator :
    1 : Free format.
    2 : Fixed format with above format descriptor.
    3 : Unformatted.

    FROM : file type parameter
    'UNIT' : open file by unit number only.
    'NAME' : open file by name and assign to unit.

If the above unit number equals 10, then the y-coord is read from this
file. The y-coord must follow the above record. No comment lines are
allowed within the y-coord input.

ELSE IF ( UNSTRUCTURED GRID ) THEN

Nothing to declare: all the data will be read from the GMESH file
END IF ( CURVILINEAR GRID )

5 Limiting bottom depth (m) to discriminate between land and sea
points, minimum water depth (m) as allowed in model, unit number
of file with bottom depths, scale factor for bottom depths (mult.),
IDLA, IDFM, format for formatted read, FROM and filename.

    IDLA : Layout indicator :
    1 : Read line-by-line bottom to top.
    2 : Like 1, single read statement.
    3 : Read line-by-line top to bottom.
    4 : Like 3, single read statement.

    IDFM : format indicator :
    1 : Free format.
    2 : Fixed format with above format descriptor.
    3 : Unformatted.

    FROM : file type parameter
    'UNIT' : open file by unit number only.
    'NAME' : open file by name and assign to unit.

If the above unit number equals 10, then the bottom depths are read from
this file. The depths must follow the above record. No comment lines are
allowed within the depth input. In the case of unstructured grids, the file
is expected to be a GMESH grid file containing node and element lists.

Example for rectilinear grid with spherical (lon/lat) coordinate system.
Note that for Cartesian coordinates the unit is meters (NOT km).

'RECT' T 'SMPL'
720  313  30
30  30  60.00
0 -78.00  1.00
-0.10  2.50  21  0.010000 1 1 '{....}' 'NAME' 'glob_30m.bot'
22  0.010000 1 1 '{....}' 'NAME' 'glob_30m.obst'
23  1 1 '{....}' 'NAME' 'glob_30m.mask'
Appendix A

Example for curvilinear grid with spherical (lon/lat) coordinate system.
Same spatial grid as preceding rectilinear example.
Note that for Cartesian coordinates the unit is meters (NOT km).

'CURV' T 'NONE'
12 12
10 0.25 -0.5 3 1 '(....)' 'NAME' 'x.inp'
1 2 3 4 5 6 7 8 9 10 11 12
2 2

-0.1 2.50 10 -10. 3 1 '(....)' 'NAME' 'bottom.inp'

SMC grid use the same spherical lat-lon grid parameters
'RECT' T 'NONE'
1824 784
plus 3 extra parameter:
Number of refined level NRLv, j-count shift, and boundary cell number.
Zero boundary cell number will make the model to skip all boundary update lines.
Non-zero boundary cell number will require an extra boundary cell list input file.

3 1344 0
SMC grid base level resolution dlon dlalt and start lon lat
0.35156250 0.23437500 1.
0.17578125 -78.6328125 1.
And the usual depth, subgrid-obstruction, and mask as in spherical lat-lon grid
plus SMC cell and face arrays:
-0.1 10.0 38 -1. 1 1 '(....)' 'NAME' 'SMC25Depth.dat'
31 1 0 1 1 '(....)' 'NAME' 'SMC25Subtr.dat'
32 1 1 '(....)' 'S6125MCels.dat'
33 1 1 '(....)' 'S6125ISide.dat'
34 1 1 '(....)' 'S6125JSide.dat'
35 1 1 '(....)' 'S6125Bundy.dat'
Boundary cell id list file (unit 35) is only required if boundary cell number entered above is non-zero. The cell id number should be the sequential number in the cell array (unit 32) S625MCels.dat.

If sub-grid information is available as indicated by FLAGTR above, additional input to define this is needed below. In such cases a field of fractional obstructions at or between grid points needs to be supplied. First the location and format of the data is defined by (as above):
- Unit number of file (can be 10, and/or identical to bottom depth unit), scale factor for fractional obstruction, IDLA, IDFM, format for formatted read, FROM and filename

10 0.2 3 1 '([...])' 'NAME' 'obstr.inp'

*** NOTE if this unit number is the same as the previous bottom depth unit number, it is assumed that this is the same file without further checks. ***

If the above unit number equals 10, the bottom data is read from this file and follows below (no intermediate comment lines allowed, except between the two fields).

0 0 0 0 0 0 0 0 0 0 0 0
0 0 0 0 0 0 0 0 0 0 0 0
0 0 0 0 0 0 0 0 0 0 0 0
0 0 0 0 0 0 0 0 0 0 0 0
0 0 0 0 0 0 0 0 0 0 0 0
0 0 0 0 0 0 0 0 0 0 0 0
0 0 0 0 0 0 0 0 0 0 0 0
0 0 0 0 0 0 0 0 0 0 0 0
0 0 0 0 0 0 0 0 0 0 0 0
0 0 0 0 0 0 0 0 0 0 0 0

*** NOTE size of fields is always NX * NY ***

Input boundary points and excluded points --------------------------
The first line identifies where to get the map data, by unit number IDLA and IDFM, format for formatted read, FROM and filename if FROM = 'PART', then segmented data is read from below, else the data is read from file as with the other inputs (as INTEGER)

10 3 1 '([...])' 'PART' 'mapsta.inp'

Read the status map from file ( FROM != PART ) ---------------------

3 3 3 3 3 3 3 3 3 3 3 3
3 2 1 1 1 0 1 1 1 1 1 3
3 2 1 1 1 0 1 1 1 1 1 3
3 2 1 1 1 0 1 1 1 1 1 3
3 2 1 1 1 0 1 1 1 1 1 3
3 2 1 1 1 0 1 1 1 1 1 3
3 2 1 1 1 0 1 1 1 1 1 3
3 2 1 1 1 0 1 1 1 1 1 3
3 2 1 1 1 0 1 1 1 1 1 3
3 3 3 3 3 3 3 3 3 3 3 3

The legend for the input map is:
0 : Land point.
1 : Regular sea point.
2 : Active boundary point.
3 : Point excluded from grid.

Input boundary points from segment data ( FROM = PART ) ------------------
An unlimited number of lines identifying points at which input boundary conditions are to be defined. If the actual input data is not defined in the actual wave model run, the initial conditions will be applied as constant boundary conditions. Each line contains: discrete grid counters (IX,IY) of the active point and a connect flag. If this flag is true, and the present and previous
point are on a grid line or diagonal, all intermediate points
are also defined as boundary points.

2 2 F
2 11 T

Close list by defining point (0,0) (mandatory)

0 0 F

Excluded grid points from segment data ( FROM != PART )
First defined as lines, identical to the definition of the input
boundary points, and closed the same way.

0 0 F

Second, define a point in a closed body of sea points to remove
the entire body of sea points. Also close by point (0,0)

0 0

Sedimentary bottom map if namelist &SBT4 SEDMAPO50 = T

22 1.1 'NAME' 'SED.txt'

Output boundary points ---------------------------------------------
Output boundary points are defined as a number of straight lines,
devided by its starting point (X0,Y0), increments (DX,DY) and number
of points. A negative number of points starts a new output file.
Note that this data is only generated if requested by the actual
program. Example again for spherical grid in degrees. Note, these do
not need to be defined for data transfer between grids in the multi
grid driver.

1.75 1.50 0.25 -0.10 3
2.25 1.50 -0.10 0.00 -6
0.10 0.10 0.10 0.00 -10

Close list by defining line with 0 points (mandatory)

0. 0. 0. 0. 0

---------------------------------------------
End of input file
---------------------------------------------
A.3.2  

ww3.grid.inp.casc_6m: For grid preprocessor (Cascadia 0.1 degree grid)
A.3.3  

ww3_grid.inp.points: For grid preprocessor (Point output)

$--------------------------------------------------------------------$
$WAVEWATCH III Grid preprocessor input file                           $
$--------------------------------------------------------------------$

'Spectral resolution for points        '$

 1.10  0.03  32  36  0

F T T T T T

3600. 480. 1800. 30.

&REF1 REFCOAST=0.10, REFFREQ=1., REFRMAX=0.8, REFSUBGRID=0.20, 
  REFICEBERG=0.4 /
&REF2 CICEB = 0.25, CICEM = 0.75, FLAGTR = 4 /
&OUTS P5SF = 1, I1P5SF = 1, I2P5SF = 32, 
  E3D = 1, I1E3D = 1, I2E3D = 32 /
END OF NAMELISTS

'RECT’ T ’SMPL’

  720  313
  30  30  60.00
  0  -78.0000  1.00
  -0.10  2.50  21  0.001000 1 1 ’(....)’ ’NAME’ ’glob_30m.bot’
  22  0.01000  1 1 ’(....)’ ’NAME’ ’glob_30m.obst’
  23  1  1 ’(....)’ ’NAME’ ’glob_30m.mask’

$ 0.  0.  0.  0.  0$
$--------------------------------------------------------------------$
$End of input file                                                   $
$--------------------------------------------------------------------$
A.3.4  ww3_strt.inp: For initial conditions program [with comments]

```
$ WAVEWATCH III Initial conditions input file
$ type of initial field ITYPE .
$ 5
$ ITYPE = 1 ----------------------------------------------------------
$ Gaussian in frequency and space, cos type in direction.
$  fp and spread (Hz), mean direction (degr., oceanographic
$  convention) and cosine power, Xm and spread (degr. or m) Ym and
$  spread (degr. or m), Hmax (m) (Example for lon-lat grid in degr.).
$  0.10  0.01  270. 2  1. 0.5 1. 0.5 2.5
$  0.10  0.01  270. 2  0. 1000. 1. 1000. 0.01
$  0.10  0.01  270. 2  0. 1000. 1. 1000. 0.
$ ITYPE = 2 ----------------------------------------------------------
$ JONSWAP spectrum with Hasselmann et al. (1980) direct. distribution.
$  alfa, peak freq. (Hz), mean direction (degr., oceanographical
$  convention), gamma, sigA, sigB, Xm and spread (degr. or m) Ym and
$  spread (degr. or m) (Example for lon-lat grid in degr.).
$  alfa, sigA, sigB give default values if less than or equal to 0.
$  0.0081  0.1  270. 1.0 0. 0. 1. 100. 1. 100.
$ ITYPE = 3 ----------------------------------------------------------
$ Fetch-limited JONSWAP
$ - No additional data, the local spectrum is calculated using the
$  local wind speed and direction, using the spatial grid size as
$  fetch, and assuring that the spectrum is within the discrete
$  frequency range.
$ ITYPE = 4 ----------------------------------------------------------
$ User-defined spectrum
$ - Scale factor., defaults to 1 if less than or equal 0.
$ - Spectrum F(f,theta) (single read statement)
$ -0.1
$  0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0
$  0 1 1 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0
$  0 1 4 2 1 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0
$  0 0 1 1 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0
$  0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0
$  0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0
$  0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0
$  0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0
$  0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0
$  0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0
$  0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0
$  0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0
$  0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0
$  0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0
$  0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0
$  0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0
$  0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0
$ ITYPE = 5 ----------------------------------------------------------
$ Starting from calm conditions.
$ - No additional data.
$ End of input file
```
A.3.5  ww3_prnc.inp.wind: For NetCDF input field preprocessor (Wind) [with comments]

$ WAVEWATCH III Field preprocessor input file $
$ ------------------------------------------------------------- $
$ WAVEWATCH III Field preprocessor input file $
$ ------------------------------------------------------------- $

$ Major types of field and time flag $
$ Field types  :  ICE   Ice concentrations. $  
$ ISI Icebergs and sea ice. $  
$ LEV Water levels. $  
$ WND Winds. $  
$ WNS Winds (including air-sea temp. dif.) $  
$ CUR Currents. $  
$ DAT Data for assimilation. $  
$ Format types :  AI  Transfer field 'as is'. (ITYPE 1) $  
$ LL Field defined on regular longitude-latitude $  
$ or Cartesian grid. (ITYPE 2) $  
$ Format types :  AT  Transfer field 'as is', performs tidal $  
$ analysis on the time series (ITYPE 6) $  
$ Time flag    : If true, time is included in file. $  
$ Header flag  : If true, header is added to file. $  
$ (necessary for reading, FALSE is used only for $  
$ incremental generation of a data file.) $  
$ 'WND' 'LL' T T $  
$ Name of spatial dimensions------------------------------------------ $  
$ NB: time dimension is expected to be called 'time' $  
$ longitude latitude $  
$ Variables to use --------------------------------------------------- $  
$ u10 v10 $  
$ Additional time input ---------------------------------------------- $  
$ If time flag is .FALSE., give time of field in yyyyddd hhmmss format. $  
$ 19680606 053000 $  
$ Define data files -------------------------------------------------- $  
$ The input line identifies the filename using for the forcing field. $  
'wind.nc' $  
$ End of input file $
Appendix A

A.3.6  ww3_prnc.inp.ice: For NetCDF input field preprocessor (Ice)

$ WAVEWATCH III Field preprocessor input file
$ -------------------------------------------
 'ICE' 'LL' T T
 longitude latitude
 ci
 'ice.nc'
$
$ End of input file
A.3.7  ww3_multi.inp: For the multi-grid shell [with comments]

The first input line sets up the general multi-grid model definition
by defining the following six parameters:

1) Number of wave model grids. (NRGRD)
2) Number of grids defining input fields. (NRINP)
3) Flag for using unified point output file. (UNIPTS)
4) Output server type as in ww3_shel.inp
5) Flag for dedicated process for unified point output.
6) Flag for grids sharing dedicated output processes.

Now each actual wave model grid is defined using 13 parameters to be
read for a single line in the file. Each line contains the following parameters:

1) Define the grid with the extension of the mod_def file.
2-8) Define the inputs used by the grids with 8 keywords,
corresponding to the 8 flags defining the input in the
input files. Valid keywords are:
   'native' : This grid has its own input files, e.g. grid
   'no' : This input is not used.
   'MODID' : Take input from the grid identified by
   'MODID'. In the example below, all grids get
their wind from wind.input (mod_def.input).
9) Rank number of grid (internally sorted and reassigned).
10) Group number (internally reassigned so that different
ranks result in different group numbers).
11-12) Define fraction of communicator (processes) used for this
grid.
13) Flag identifying dumping of boundary data used by this
grid. If true, the file nest.MODID is generated.

In this example we need the file mod_def.input to define the
grid and the file wind.input to provide the corresponding wind data.

In this example, we need the file mod_def.input to define the
grid. All grids get their winds from wind.input. If there are input data grids defined (NRINP > 0 ), then these
grids are defined first. These grids are defined as if they are wave
model grids using the file mod_def.MODID. Each grid is defined on
a separate input line with MODID, and eight input flags identifying
the presence of 1) water levels 2) currents 3) winds 4) ice and
5-7) assimilation data as in the file ww3_shel.inp.

In this example three grids are used requiring the files

<table>
<thead>
<tr>
<th>Grd</th>
<th>native</th>
<th>MODID</th>
<th>wind.input</th>
<th>ice.input</th>
<th>wind.input</th>
<th>ice.input</th>
<th>wind.input</th>
<th>ice.input</th>
<th>wind.input</th>
</tr>
</thead>
<tbody>
<tr>
<td>Grid1</td>
<td>no</td>
<td>1</td>
<td>no</td>
<td>no</td>
<td>no</td>
<td>no</td>
<td>no</td>
<td>no</td>
<td>no</td>
</tr>
<tr>
<td>Grid2</td>
<td>no</td>
<td>no</td>
<td>no</td>
<td>yes</td>
<td>yes</td>
<td>no</td>
<td>no</td>
<td>no</td>
<td>no</td>
</tr>
<tr>
<td>Grid3</td>
<td>no</td>
<td>no</td>
<td>no</td>
<td>yes</td>
<td>yes</td>
<td>yes</td>
<td>no</td>
<td>no</td>
<td>no</td>
</tr>
</tbody>
</table>

In this example three grids are used requiring the files
mod_def.MODID. All files get their winds from the grid input
defined by mod_def.input, and no other inputs are used. In the times
that are commented out, each grid runs on a part of the pool of
processes assigned to the computation.

Starting and ending times for the entire model run:

20140901 000000 20140902 000000

Specific multi-scale model settings (single line).
Flag for masking computation in two-way nesting (except at
output times).
Flag for masking at printout time.

Conventional output requests as in ww3_shel.inp. Will be applied
+ to all grids.

Output request flags identifying fields as in ww3_shel.inp. See that
file for a full documentation of field output options. Namelist type
selection is used here for alternative F/T flags, see ww3_shel.inp.
### Global output point data file for multi-grid wave model

#### Key to data in file:

- **LON**: Longitude, east positive
- **LAT**: Latitude
- **NAME**: Output point name C*10, no blanks in name allowed
- **AH**: Anemometer height, dummy value for none-data points
- **TYPE**: Buoy type indicator, used for plotting and postprocessing
  - **DAT**: Data point
  - **XDT**: Former data point
  - **BPT**: Boundary data for external models.
  - **VBY**: 'Virtual buoy'
- **SOURCE**: Source of data point
  - **ENCAN**: Environment Canada
  - **GOMOOS**: Gulf of Maine OOS
  - **IDT**: Irish Department of Transportation
  - **METFR**: Meteo France
  - **NCEP**: Boundary and other data points
  - **NDBC**: National Data Buoy Center
  - **PRIV**: Private and incidental data sources
  - **SCRIPPS**: Scripps
  - **UKMO**: UK Met Office
- **SCALE**: Scale indicator for plotting of locations on map

#### Notes:

- The `$` at the first position identifies comments for WAVEWATCH III input.
- The first three data columns are used by the forecasts code, the other are used by postprocessing scripts.

### NE Pacific deep ocean

<table>
<thead>
<tr>
<th>LON</th>
<th>LAT</th>
<th>NAME</th>
<th>AH</th>
<th>TYPE</th>
<th>SOURCE</th>
<th>SCALE</th>
</tr>
</thead>
<tbody>
<tr>
<td>-130.47</td>
<td>42.60</td>
<td>'46002'</td>
<td>5.0</td>
<td>DAT</td>
<td>NDBC</td>
<td>1</td>
</tr>
<tr>
<td>-136.10</td>
<td>58.93</td>
<td>'46004'</td>
<td>5.0</td>
<td>DAT</td>
<td>NDBC</td>
<td>1</td>
</tr>
<tr>
<td>-131.00</td>
<td>45.96</td>
<td>'46005'</td>
<td>5.0</td>
<td>DAT</td>
<td>NDBC</td>
<td>1</td>
</tr>
<tr>
<td>-137.48</td>
<td>40.00</td>
<td>'46006'</td>
<td>5.0</td>
<td>DAT</td>
<td>NDBC</td>
<td>1</td>
</tr>
<tr>
<td>-177.58</td>
<td>57.05</td>
<td>'46035'</td>
<td>10.0</td>
<td>DAT</td>
<td>NDBC</td>
<td>1</td>
</tr>
<tr>
<td>-133.94</td>
<td>48.35</td>
<td>'46036'</td>
<td>5.0</td>
<td>DAT</td>
<td>NDBC</td>
<td>1</td>
</tr>
<tr>
<td>-177.03</td>
<td>57.05</td>
<td>'46033'</td>
<td>5.0</td>
<td>DAT</td>
<td>NDBC</td>
<td>1</td>
</tr>
<tr>
<td>-172.03</td>
<td>54.94</td>
<td>'46073'</td>
<td>10.0</td>
<td>DAT</td>
<td>NDBC</td>
<td>1</td>
</tr>
<tr>
<td>-138.50</td>
<td>53.91</td>
<td>'46184'</td>
<td>5.0</td>
<td>DAT</td>
<td>OCEAN</td>
<td>1</td>
</tr>
</tbody>
</table>

### NE Pacific coastal

<table>
<thead>
<tr>
<th>LON</th>
<th>LAT</th>
<th>NAME</th>
<th>AH</th>
<th>TYPE</th>
<th>SOURCE</th>
<th>SCALE</th>
</tr>
</thead>
<tbody>
<tr>
<td>-146.83</td>
<td>68.22</td>
<td>'46061'</td>
<td>5.0</td>
<td>DAT</td>
<td>NDBC</td>
<td>1</td>
</tr>
<tr>
<td>179.05</td>
<td>51.16</td>
<td>'46071'</td>
<td>5.0</td>
<td>DAT</td>
<td>NDBC</td>
<td>1</td>
</tr>
<tr>
<td>171.73</td>
<td>52.25</td>
<td>'46072'</td>
<td>5.0</td>
<td>DAT</td>
<td>NDBC</td>
<td>1</td>
</tr>
<tr>
<td>-168.81</td>
<td>53.93</td>
<td>'46075'</td>
<td>5.0</td>
<td>DAT</td>
<td>NDBC</td>
<td>1</td>
</tr>
<tr>
<td>-148.00</td>
<td>55.58</td>
<td>'46076'</td>
<td>5.0</td>
<td>DAT</td>
<td>NDBC</td>
<td>1</td>
</tr>
<tr>
<td>-152.45</td>
<td>56.05</td>
<td>'46078'</td>
<td>5.0</td>
<td>DAT</td>
<td>NDBC</td>
<td>1</td>
</tr>
<tr>
<td>-158.00</td>
<td>58.00</td>
<td>'46080'</td>
<td>5.0</td>
<td>DAT</td>
<td>NDBC</td>
<td>1</td>
</tr>
<tr>
<td>-143.42</td>
<td>59.69</td>
<td>'46082'</td>
<td>5.0</td>
<td>DAT</td>
<td>NDBC</td>
<td>1</td>
</tr>
<tr>
<td>-138.00</td>
<td>58.25</td>
<td>'46083'</td>
<td>5.0</td>
<td>DAT</td>
<td>NDBC</td>
<td>1</td>
</tr>
<tr>
<td>-136.16</td>
<td>56.59</td>
<td>'46084'</td>
<td>5.0</td>
<td>DAT</td>
<td>NDBC</td>
<td>1</td>
</tr>
</tbody>
</table>
Appendix A

$ Canada$

-127.93 49.74 `46132` 5.0 DAT ENCAN 1
-132.45 54.38 `46145` 5.0 DAT ENCAN 1
-131.22 51.63 `46147` 5.0 DAT ENCAN 1
-131.10 53.62 `46183` 5.0 DAT ENCAN 1
-129.81 53.42 `46185` 5.0 DAT ENCAN 1
-126.75 51.37 `46264` 5.0 DAT ENCAN 1
-134.28 54.16 `46285` 5.0 DAT ENCAN 1
-126.80 48.84 `46286` 5.0 DAT ENCAN 1
-129.92 58.87 `46287` 5.0 DAT ENCAN 1
-132.68 52.52 `46288` 5.0 DAT ENCAN 1

$ USA$

-120.87 34.88 `46011` 5.0 DAT NDBC 1
-122.88 37.36 `46012` 5.0 DAT NDBC 1
-123.32 38.23 `46013` 5.0 DAT NDBC 1
-123.97 39.22 `46014` 5.0 DAT NDBC 1
-124.85 42.75 `46015` 5.0 DAT NDBC 1
-124.54 48.78 `46022` 5.0 DAT NDBC 1
-128.97 34.71 `46033` 10.0 DAT NDBC 1
-119.88 33.75 `46035` 5.0 DAT NDBC 1
-122.82 37.75 `46036` 5.0 DAT NDBC 1
-124.38 41.85 `46077` 5.0 DAT NDBC 1
-121.89 35.74 `46082` 5.0 DAT NDBC 1
-124.51 46.12 `46089` 5.0 DAT NDBC 1
-124.53 48.42 `46088` 5.0 DAT NDBC 1
-124.75 47.34 `46041` 5.0 DAT NDBC 1
-122.42 36.75 `46042` 5.0 DAT NDBC 1
-119.53 32.43 `46043` 5.0 DAT NDBC 1
-124.53 44.62 `46058` 5.0 DAT NDBC 1
-119.85 34.24 `46053` 5.0 DAT NDBC 1
-128.45 34.27 `46054` 10.0 DAT NDBC 1
-121.81 35.18 `46062` 5.0 DAT NDBC 1
-126.78 34.27 `46063` 5.0 DAT NDBC 1
-128.28 33.65 `46069` 5.0 DAT NDBC 1
-118.08 32.58 `46086` 5.0 DAT NDBC 1
-124.73 48.49 `46087` 5.0 DAT NDBC 1
-125.77 48.88 `46089` 5.0 DAT NDBC 1
-124.24 46.06 `46211` 999. DAT SCRIPPS 1
-124.74 46.29 `46213` 999. DAT SCRIPPS 1
-122.47 37.95 `46214` 999. DAT SCRIPPS 1
-119.88 34.33 `46216` 999. DAT SCRIPPS 1
-119.43 34.17 `46217` 999. DAT SCRIPPS 1
-120.78 34.45 `46218` 999. DAT SCRIPPS 1
-118.63 33.85 `46221` 999. DAT SCRIPPS 1
-118.32 33.62 `46222` 999. DAT SCRIPPS 1
-117.77 33.46 `46223` 999. DAT SCRIPPS 1
-117.47 33.18 `46224` 999. DAT SCRIPPS 1
-117.39 32.93 `46225` 999. DAT SCRIPPS 1
-117.44 32.63 `46227` 999. DAT SCRIPPS 1
-124.55 43.77 `46229` 999. DAT SCRIPPS 1
-117.37 32.75 `46231` 999. DAT SCRIPPS 1
-117.33 32.43 `46232` 999. DAT SCRIPPS 1

$ USA$

-117.75 32.64 `45001` 999. VBY NCEP 1

$ NE Pacific boundary data$

-124.25 48.00 `EKA51` 999. BPT NCEP 1
-125.25 48.00 `EKA52` 999. BPT NCEP 1
-127.25 48.00 `EKA53` 999. BPT NCEP 1
-127.25 41.25 `EKA54` 999. BPT NCEP 1
-127.25 42.50 `EKA55` 999. BPT NCEP 1
-127.25 41.75 `EKA56` 999. BPT NCEP 1
-125.58 43.75 `EKA57` 999. BPT NCEP 1
-124.75 43.75 `EKA58` 999. BPT NCEP 1
-127.25 39.00 `EKA59` 999. BPT NCEP 1
-127.25 38.00 `EKA60` 999. BPT NCEP 1
-127.25 31.00 `EKA61` 999. BPT NCEP 1
-122.25 37.00 `EKA62` 999. BPT NCEP 1
-123.25 37.00 `EKA63` 999. BPT NCEP 1
-122.34 37.00 `EKA64` 999. BPT NCEP 1
-127.25 44.75 `EKA65` 999. BPT NCEP 1
-127.25 46.00 `EKA66` 999. BPT NCEP 1
-125.25 46.00 `EKA67` 999. BPT NCEP 1
-124.25 46.00 `EKA68` 999. BPT NCEP 1
-125.58 42.75 `MOR51` 999. BPT NCEP 1
-126.58 42.75 `MOR52` 999. BPT NCEP 1
-126.58 43.75 `MOR53` 999. BPT NCEP 1
-126.58 44.75 `MOR54` 999. BPT NCEP 1
-126.58 45.75 `MOR55` 999. BPT NCEP 1
-126.58 46.75 `MOR56` 999. BPT NCEP 1
-126.58 47.75 `MOR57` 999. BPT NCEP 1
-126.58 48.50 `MOR58` 999. BPT NCEP 1
-125.58 48.50 `MOR59` 999. BPT NCEP 1
-125.00 48.00 `MFR51` 999. BPT NCEP 1
-126.00 48.00 `MFR52` 999. BPT NCEP 1
-127.00 48.00 `MFR53` 999. BPT NCEP 1
-128.00 48.00 `MFR54` 999. BPT NCEP 1
-128.00 41.00 `MFR55` 999. BPT NCEP 1
-128.00 42.00 `MFR56` 999. BPT NCEP 1
-128.00 43.00 `MFR57` 999. BPT NCEP 1
-128.00 44.00 `MFR58` 999. BPT NCEP 1
-128.00 45.00 `MFR59` 999. BPT NCEP 1
-128.00 46.00 `MFR60` 999. BPT NCEP 1
-127.00 46.00 `MFR61` 999. BPT NCEP 1
-126.00 46.00 `MFR62` 999. BPT NCEP 1
-125.00 46.00 `MFR63` 999. BPT NCEP 1
-121.00 34.50 `SDX51` 999. BPT NCEP 1
-121.00 33.50 `SDX52` 999. BPT NCEP 1
<table>
<thead>
<tr>
<th>LON</th>
<th>LAT</th>
<th>NAME</th>
<th>AH</th>
<th>TYPE</th>
<th>SOURCE</th>
<th>SCALE</th>
</tr>
</thead>
<tbody>
<tr>
<td>136.00</td>
<td>28.80</td>
<td>'23NN4'</td>
<td>999</td>
<td>VBY</td>
<td>NCEP</td>
<td>1</td>
</tr>
<tr>
<td>126.00</td>
<td>28.10</td>
<td>'22NN1'</td>
<td>999</td>
<td>VBY</td>
<td>NCEP</td>
<td>1</td>
</tr>
</tbody>
</table>

$\text{Japanese buoys}$

<table>
<thead>
<tr>
<th>LON</th>
<th>LAT</th>
<th>NAME</th>
<th>AH</th>
<th>TYPE</th>
<th>SOURCE</th>
<th>SCALE</th>
</tr>
</thead>
<tbody>
<tr>
<td>160.00</td>
<td>-5.00</td>
<td>'SOLOMON_N'</td>
<td>999</td>
<td>VBY</td>
<td>NCEP</td>
<td>1</td>
</tr>
<tr>
<td>165.00</td>
<td>-12.00</td>
<td>'SOLOMON_SE'</td>
<td>999</td>
<td>VBY</td>
<td>NCEP</td>
<td>1</td>
</tr>
<tr>
<td>160.00</td>
<td>-12.00</td>
<td>'SOLOMON_SW'</td>
<td>999</td>
<td>VBY</td>
<td>NCEP</td>
<td>1</td>
</tr>
<tr>
<td>171.90</td>
<td>-9.20</td>
<td>'NUKUNONO'</td>
<td>999</td>
<td>VBY</td>
<td>NCEP</td>
<td>1</td>
</tr>
<tr>
<td>167.00</td>
<td>-0.50</td>
<td>'NAURU'</td>
<td>999</td>
<td>VBY</td>
<td>NCEP</td>
<td>1</td>
</tr>
<tr>
<td>166.30</td>
<td>23.90</td>
<td>'FF_SHOALS'</td>
<td>999</td>
<td>VBY</td>
<td>NCEP</td>
<td>1</td>
</tr>
<tr>
<td>169.90</td>
<td>-19.10</td>
<td>'NIUE'</td>
<td>999</td>
<td>VBY</td>
<td>NCEP</td>
<td>1</td>
</tr>
<tr>
<td>174.00</td>
<td>1.00</td>
<td>'TARAWA'</td>
<td>999</td>
<td>VBY</td>
<td>NCEP</td>
<td>1</td>
</tr>
<tr>
<td>149.60</td>
<td>-19.00</td>
<td>'PAPEETE'</td>
<td>999</td>
<td>VBY</td>
<td>NCEP</td>
<td>1</td>
</tr>
<tr>
<td>174.00</td>
<td>1.00</td>
<td>'TARAWA'</td>
<td>999</td>
<td>VBY</td>
<td>NCEP</td>
<td>1</td>
</tr>
<tr>
<td>165.00</td>
<td>-12.00</td>
<td>'SOLOMON_SE'</td>
<td>999</td>
<td>VBY</td>
<td>NCEP</td>
<td>1</td>
</tr>
<tr>
<td>168.00</td>
<td>-5.00</td>
<td>'SOLOMON_N'</td>
<td>999</td>
<td>VBY</td>
<td>NCEP</td>
<td>1</td>
</tr>
</tbody>
</table>

$\text{Pacific training points}$

<table>
<thead>
<tr>
<th>LON</th>
<th>LAT</th>
<th>NAME</th>
<th>AH</th>
<th>TYPE</th>
<th>SOURCE</th>
<th>SCALE</th>
</tr>
</thead>
<tbody>
<tr>
<td>28.20</td>
<td>13.54</td>
<td>'5220'</td>
<td>999</td>
<td>VBY</td>
<td>NCEP</td>
<td>1</td>
</tr>
</tbody>
</table>

$\text{Other deep Pacific}$

<table>
<thead>
<tr>
<th>LON</th>
<th>LAT</th>
<th>NAME</th>
<th>AH</th>
<th>TYPE</th>
<th>SOURCE</th>
<th>SCALE</th>
</tr>
</thead>
<tbody>
<tr>
<td>138.75</td>
<td>3.00</td>
<td>'51B08'</td>
<td>999</td>
<td>VBY</td>
<td>NCEP</td>
<td>1</td>
</tr>
<tr>
<td>144.70</td>
<td>13.54</td>
<td>'5220'</td>
<td>999</td>
<td>VBY</td>
<td>NCEP</td>
<td>1</td>
</tr>
</tbody>
</table>

$\text{Gulf of Mexico and Caribbean}$

<table>
<thead>
<tr>
<th>LON</th>
<th>LAT</th>
<th>NAME</th>
<th>AH</th>
<th>TYPE</th>
<th>SOURCE</th>
<th>SCALE</th>
</tr>
</thead>
<tbody>
<tr>
<td>121.00</td>
<td>32.50</td>
<td>'SDX53'</td>
<td>999</td>
<td>VBY</td>
<td>NCEP</td>
<td>1</td>
</tr>
<tr>
<td>121.00</td>
<td>31.50</td>
<td>'SDX54'</td>
<td>999</td>
<td>VBY</td>
<td>NCEP</td>
<td>1</td>
</tr>
<tr>
<td>120.00</td>
<td>31.50</td>
<td>'SDX55'</td>
<td>999</td>
<td>VBY</td>
<td>NCEP</td>
<td>1</td>
</tr>
<tr>
<td>119.00</td>
<td>31.50</td>
<td>'SDX56'</td>
<td>999</td>
<td>VBY</td>
<td>NCEP</td>
<td>1</td>
</tr>
<tr>
<td>118.00</td>
<td>31.50</td>
<td>'SDX57'</td>
<td>999</td>
<td>VBY</td>
<td>NCEP</td>
<td>1</td>
</tr>
<tr>
<td>117.00</td>
<td>31.50</td>
<td>'SDX58'</td>
<td>999</td>
<td>VBY</td>
<td>NCEP</td>
<td>1</td>
</tr>
<tr>
<td>112.00</td>
<td>36.00</td>
<td>'LDX51'</td>
<td>999</td>
<td>VBY</td>
<td>NCEP</td>
<td>1</td>
</tr>
<tr>
<td>122.00</td>
<td>36.00</td>
<td>'LDX52'</td>
<td>999</td>
<td>VBY</td>
<td>NCEP</td>
<td>1</td>
</tr>
</tbody>
</table>

$\text{NWS forecast points}$

<table>
<thead>
<tr>
<th>LON</th>
<th>LAT</th>
<th>NAME</th>
<th>AH</th>
<th>TYPE</th>
<th>SOURCE</th>
<th>SCALE</th>
</tr>
</thead>
<tbody>
<tr>
<td>144.80</td>
<td>13.54</td>
<td>'5220'</td>
<td>999</td>
<td>VBY</td>
<td>NCEP</td>
<td>1</td>
</tr>
<tr>
<td>160.29</td>
<td>19.02</td>
<td>'51B03'</td>
<td>999</td>
<td>VBY</td>
<td>NCEP</td>
<td>1</td>
</tr>
</tbody>
</table>

$\text{Appendix A}$

$\text{Hawaii}$

<table>
<thead>
<tr>
<th>LON</th>
<th>LAT</th>
<th>NAME</th>
<th>AH</th>
<th>TYPE</th>
<th>SOURCE</th>
<th>SCALE</th>
</tr>
</thead>
<tbody>
<tr>
<td>167.50</td>
<td>-18.00</td>
<td>'PORT_VILA'</td>
<td>999</td>
<td>VBY</td>
<td>NCEP</td>
<td>1</td>
</tr>
<tr>
<td>167.50</td>
<td>-24.00</td>
<td>'NOUMEA'</td>
<td>999</td>
<td>VBY</td>
<td>NCEP</td>
<td>1</td>
</tr>
<tr>
<td>159.80</td>
<td>-21.20</td>
<td>'RAROTONGA'</td>
<td>999</td>
<td>VBY</td>
<td>NCEP</td>
<td>1</td>
</tr>
<tr>
<td>175.00</td>
<td>-22.00</td>
<td>'TONGATAPU'</td>
<td>999</td>
<td>VBY</td>
<td>NCEP</td>
<td>1</td>
</tr>
<tr>
<td>176.25</td>
<td>-18.00</td>
<td>'NADI'</td>
<td>999</td>
<td>VBY</td>
<td>NCEP</td>
<td>1</td>
</tr>
<tr>
<td>172.50</td>
<td>-16.70</td>
<td>'JOHNSTON'</td>
<td>999</td>
<td>VBY</td>
<td>NCEP</td>
<td>1</td>
</tr>
<tr>
<td>177.40</td>
<td>28.20</td>
<td>'MIDWAY'</td>
<td>999</td>
<td>VBY</td>
<td>NCEP</td>
<td>1</td>
</tr>
</tbody>
</table>

$\text{Pacific training points}$

<table>
<thead>
<tr>
<th>LON</th>
<th>LAT</th>
<th>NAME</th>
<th>AH</th>
<th>TYPE</th>
<th>SOURCE</th>
<th>SCALE</th>
</tr>
</thead>
<tbody>
<tr>
<td>177.40</td>
<td>28.20</td>
<td>'MIKAWA'</td>
<td>999</td>
<td>VBY</td>
<td>NCEP</td>
<td>1</td>
</tr>
<tr>
<td>169.50</td>
<td>16.70</td>
<td>'JOHNSTON'</td>
<td>999</td>
<td>VBY</td>
<td>NCEP</td>
<td>1</td>
</tr>
<tr>
<td>176.25</td>
<td>19.02</td>
<td>'NOUMEA'</td>
<td>999</td>
<td>VBY</td>
<td>NCEP</td>
<td>1</td>
</tr>
<tr>
<td>175.00</td>
<td>19.02</td>
<td>'TONGATAPU'</td>
<td>999</td>
<td>VBY</td>
<td>NCEP</td>
<td>1</td>
</tr>
<tr>
<td>166.50</td>
<td>23.90</td>
<td>'FF_SHOALS'</td>
<td>999</td>
<td>VBY</td>
<td>NCEP</td>
<td>1</td>
</tr>
<tr>
<td>167.00</td>
<td>19.02</td>
<td>'TONGATAPU'</td>
<td>999</td>
<td>VBY</td>
<td>NCEP</td>
<td>1</td>
</tr>
<tr>
<td>171.90</td>
<td>19.02</td>
<td>'MIDWAY'</td>
<td>999</td>
<td>VBY</td>
<td>NCEP</td>
<td>1</td>
</tr>
<tr>
<td>165.00</td>
<td>12.00</td>
<td>'SOLOMON_SE'</td>
<td>999</td>
<td>VBY</td>
<td>NCEP</td>
<td>1</td>
</tr>
<tr>
<td>168.00</td>
<td>5.00</td>
<td>'SOLOMON_N'</td>
<td>999</td>
<td>VBY</td>
<td>NCEP</td>
<td>1</td>
</tr>
</tbody>
</table>
### Appendix A

#### NE Atlantic boundary data

<table>
<thead>
<tr>
<th>LON</th>
<th>LAT</th>
<th>NAME</th>
<th>AH</th>
<th>TYPE</th>
<th>SOURCE</th>
<th>SCALE</th>
</tr>
</thead>
<tbody>
<tr>
<td>89.67</td>
<td>25.90</td>
<td>'42001'</td>
<td></td>
<td>DAT</td>
<td>NDBC</td>
<td>1</td>
</tr>
<tr>
<td>94.42</td>
<td>25.17</td>
<td>'42002'</td>
<td></td>
<td>DAT</td>
<td>NDBC</td>
<td>1</td>
</tr>
<tr>
<td>85.94</td>
<td>26.07</td>
<td>'42003'</td>
<td></td>
<td>DAT</td>
<td>NDBC</td>
<td>1</td>
</tr>
<tr>
<td>88.77</td>
<td>20.89</td>
<td>'42007'</td>
<td></td>
<td>DAT</td>
<td>NDBC</td>
<td>1</td>
</tr>
<tr>
<td>95.36</td>
<td>27.91</td>
<td>'42019'</td>
<td></td>
<td>DAT</td>
<td>NDBC</td>
<td>1</td>
</tr>
<tr>
<td>96.78</td>
<td>26.04</td>
<td>'42026'</td>
<td></td>
<td>DAT</td>
<td>NDBC</td>
<td>1</td>
</tr>
<tr>
<td>94.40</td>
<td>29.92</td>
<td>'42035'</td>
<td></td>
<td>DAT</td>
<td>NDBC</td>
<td>1</td>
</tr>
<tr>
<td>84.52</td>
<td>28.58</td>
<td>'42036'</td>
<td></td>
<td>DAT</td>
<td>NDBC</td>
<td>1</td>
</tr>
<tr>
<td>92.55</td>
<td>27.42</td>
<td>'42038'</td>
<td></td>
<td>DAT</td>
<td>NDBC</td>
<td>1</td>
</tr>
<tr>
<td>86.62</td>
<td>28.79</td>
<td>'42039'</td>
<td></td>
<td>DAT</td>
<td>NDBC</td>
<td>1</td>
</tr>
<tr>
<td>88.71</td>
<td>29.18</td>
<td>'42040'</td>
<td></td>
<td>DAT</td>
<td>NDBC</td>
<td>1</td>
</tr>
<tr>
<td>90.46</td>
<td>27.50</td>
<td>'42041'</td>
<td></td>
<td>DAT</td>
<td>NDBC</td>
<td>1</td>
</tr>
</tbody>
</table>

#### NW Atlantic coastal

<table>
<thead>
<tr>
<th>LON</th>
<th>LAT</th>
<th>NAME</th>
<th>AH</th>
<th>TYPE</th>
<th>SOURCE</th>
<th>SCALE</th>
</tr>
</thead>
<tbody>
<tr>
<td>-72.66</td>
<td>34.68</td>
<td>'41001'</td>
<td></td>
<td>DAT</td>
<td>NDBC</td>
<td>1</td>
</tr>
<tr>
<td>-75.36</td>
<td>32.32</td>
<td>'41002'</td>
<td></td>
<td>DAT</td>
<td>NDBC</td>
<td>1</td>
</tr>
<tr>
<td>-79.09</td>
<td>33.50</td>
<td>'41004'</td>
<td></td>
<td>DAT</td>
<td>NDBC</td>
<td>1</td>
</tr>
<tr>
<td>-80.87</td>
<td>31.40</td>
<td>'41008'</td>
<td></td>
<td>DAT</td>
<td>NDBC</td>
<td>1</td>
</tr>
<tr>
<td>-66.50</td>
<td>41.11</td>
<td>'41011'</td>
<td></td>
<td>DAT</td>
<td>NDBC</td>
<td>1</td>
</tr>
<tr>
<td>-62.88</td>
<td>42.26</td>
<td>'41137'</td>
<td></td>
<td>DAT</td>
<td>ENCA</td>
<td>1</td>
</tr>
<tr>
<td>-53.62</td>
<td>44.26</td>
<td>'41239'</td>
<td></td>
<td>DAT</td>
<td>ENCA</td>
<td>1</td>
</tr>
<tr>
<td>-57.88</td>
<td>44.26</td>
<td>'41359'</td>
<td></td>
<td>DAT</td>
<td>ENCA</td>
<td>1</td>
</tr>
<tr>
<td>-53.74</td>
<td>43.75</td>
<td>'41400'</td>
<td></td>
<td>DAT</td>
<td>ENCA</td>
<td>1</td>
</tr>
<tr>
<td>-58.88</td>
<td>43.80</td>
<td>'41413'</td>
<td></td>
<td>DAT</td>
<td>ENCA</td>
<td>1</td>
</tr>
<tr>
<td>-64.62</td>
<td>42.50</td>
<td>'41442'</td>
<td></td>
<td>DAT</td>
<td>ENCA</td>
<td>1</td>
</tr>
<tr>
<td>-64.81</td>
<td>42.50</td>
<td>'41450'</td>
<td></td>
<td>DAT</td>
<td>ENCA</td>
<td>1</td>
</tr>
<tr>
<td>-53.39</td>
<td>44.44</td>
<td>'41251'</td>
<td></td>
<td>DAT</td>
<td>ENCA</td>
<td>1</td>
</tr>
<tr>
<td>-57.35</td>
<td>47.28</td>
<td>'41255'</td>
<td></td>
<td>DAT</td>
<td>ENCA</td>
<td>1</td>
</tr>
<tr>
<td>-48.81</td>
<td>46.77</td>
<td>'W0807'</td>
<td></td>
<td>DAT</td>
<td>OTHER</td>
<td>1</td>
</tr>
</tbody>
</table>

#### NW Atlantic deep ocean

<table>
<thead>
<tr>
<th>LON</th>
<th>LAT</th>
<th>NAME</th>
<th>AH</th>
<th>TYPE</th>
<th>SOURCE</th>
<th>SCALE</th>
</tr>
</thead>
<tbody>
<tr>
<td>-56.20</td>
<td>14.60</td>
<td>'41101'</td>
<td></td>
<td>DAT</td>
<td>ENCA</td>
<td>1</td>
</tr>
<tr>
<td>-81.75</td>
<td>24.00</td>
<td>'4W41'</td>
<td>999</td>
<td>VBY</td>
<td>NCEP</td>
<td>1</td>
</tr>
<tr>
<td>-82.25</td>
<td>25.00</td>
<td>'4W52'</td>
<td>999</td>
<td>VBY</td>
<td>NCEP</td>
<td>1</td>
</tr>
<tr>
<td>-66.58</td>
<td>19.00</td>
<td>'PUERTO_R'</td>
<td>999</td>
<td>VBY</td>
<td>NCEP</td>
<td>1</td>
</tr>
<tr>
<td>-66.58</td>
<td>17.50</td>
<td>'PUERTO_R_S'</td>
<td>999</td>
<td>VBY</td>
<td>NCEP</td>
<td>1</td>
</tr>
</tbody>
</table>

#### NW Atlantic boundary data

<table>
<thead>
<tr>
<th>LON</th>
<th>LAT</th>
<th>NAME</th>
<th>AH</th>
<th>TYPE</th>
<th>SOURCE</th>
<th>SCALE</th>
</tr>
</thead>
<tbody>
<tr>
<td>-72.66</td>
<td>34.68</td>
<td>'41001'</td>
<td></td>
<td>DAT</td>
<td>NDBC</td>
<td>1</td>
</tr>
<tr>
<td>-69.16</td>
<td>43.19</td>
<td>'41005'</td>
<td></td>
<td>DAT</td>
<td>NDBC</td>
<td>1</td>
</tr>
<tr>
<td>-56.00</td>
<td>43.01</td>
<td>'41012'</td>
<td></td>
<td>DAT</td>
<td>NDBC</td>
<td>1</td>
</tr>
<tr>
<td>-51.74</td>
<td>43.75</td>
<td>'41400'</td>
<td></td>
<td>DAT</td>
<td>ENCA</td>
<td>1</td>
</tr>
<tr>
<td>-58.88</td>
<td>43.80</td>
<td>'41413'</td>
<td></td>
<td>DAT</td>
<td>ENCA</td>
<td>1</td>
</tr>
<tr>
<td>-64.62</td>
<td>42.50</td>
<td>'41442'</td>
<td></td>
<td>DAT</td>
<td>ENCA</td>
<td>1</td>
</tr>
<tr>
<td>-64.81</td>
<td>42.50</td>
<td>'41450'</td>
<td></td>
<td>DAT</td>
<td>ENCA</td>
<td>1</td>
</tr>
<tr>
<td>-53.39</td>
<td>44.44</td>
<td>'41251'</td>
<td></td>
<td>DAT</td>
<td>ENCA</td>
<td>1</td>
</tr>
<tr>
<td>-57.35</td>
<td>47.28</td>
<td>'41255'</td>
<td></td>
<td>DAT</td>
<td>ENCA</td>
<td>1</td>
</tr>
<tr>
<td>-48.81</td>
<td>46.77</td>
<td>'W0807'</td>
<td></td>
<td>DAT</td>
<td>OTHER</td>
<td>1</td>
</tr>
</tbody>
</table>
### Appendix A

#### Bermuda boundary data

<table>
<thead>
<tr>
<th>LON</th>
<th>LAT</th>
<th>NAME</th>
<th>AH</th>
<th>TYPE</th>
<th>SOURCE</th>
<th>SCALE</th>
</tr>
</thead>
<tbody>
<tr>
<td>-65.50</td>
<td>31.50</td>
<td>BER75</td>
<td>999</td>
<td>BPT</td>
<td>NCEP</td>
<td>1</td>
</tr>
<tr>
<td>-65.50</td>
<td>31.80</td>
<td>BER74</td>
<td>999</td>
<td>BPT</td>
<td>NCEP</td>
<td>1</td>
</tr>
<tr>
<td>-65.50</td>
<td>32.00</td>
<td>BER73</td>
<td>999</td>
<td>BPT</td>
<td>NCEP</td>
<td>1</td>
</tr>
<tr>
<td>-65.50</td>
<td>32.50</td>
<td>BER72</td>
<td>999</td>
<td>BPT</td>
<td>NCEP</td>
<td>1</td>
</tr>
<tr>
<td>-65.50</td>
<td>33.00</td>
<td>BER71</td>
<td>999</td>
<td>BPT</td>
<td>NCEP</td>
<td>1</td>
</tr>
<tr>
<td>-65.50</td>
<td>33.50</td>
<td>BER70</td>
<td>999</td>
<td>BPT</td>
<td>NCEP</td>
<td>1</td>
</tr>
<tr>
<td>-65.50</td>
<td>34.00</td>
<td>BER69</td>
<td>999</td>
<td>BPT</td>
<td>NCEP</td>
<td>1</td>
</tr>
<tr>
<td>-65.50</td>
<td>34.50</td>
<td>BER68</td>
<td>999</td>
<td>BPT</td>
<td>NCEP</td>
<td>1</td>
</tr>
<tr>
<td>-65.50</td>
<td>35.00</td>
<td>BER67</td>
<td>999</td>
<td>BPT</td>
<td>NCEP</td>
<td>1</td>
</tr>
<tr>
<td>-65.50</td>
<td>35.50</td>
<td>BER66</td>
<td>999</td>
<td>BPT</td>
<td>NCEP</td>
<td>1</td>
</tr>
<tr>
<td>-65.50</td>
<td>36.00</td>
<td>BER65</td>
<td>999</td>
<td>BPT</td>
<td>NCEP</td>
<td>1</td>
</tr>
<tr>
<td>-65.50</td>
<td>36.50</td>
<td>BER64</td>
<td>999</td>
<td>BPT</td>
<td>NCEP</td>
<td>1</td>
</tr>
<tr>
<td>-65.50</td>
<td>37.00</td>
<td>BER63</td>
<td>999</td>
<td>BPT</td>
<td>NCEP</td>
<td>1</td>
</tr>
<tr>
<td>-65.50</td>
<td>37.50</td>
<td>BER62</td>
<td>999</td>
<td>BPT</td>
<td>NCEP</td>
<td>1</td>
</tr>
<tr>
<td>-65.50</td>
<td>38.00</td>
<td>BER61</td>
<td>999</td>
<td>BPT</td>
<td>NCEP</td>
<td>1</td>
</tr>
<tr>
<td>-65.50</td>
<td>38.50</td>
<td>BER60</td>
<td>999</td>
<td>BPT</td>
<td>NCEP</td>
<td>1</td>
</tr>
</tbody>
</table>

#### NW Atlantic

<table>
<thead>
<tr>
<th>LON</th>
<th>LAT</th>
<th>NAME</th>
<th>AH</th>
<th>TYPE</th>
<th>SOURCE</th>
<th>SCALE</th>
</tr>
</thead>
<tbody>
<tr>
<td>-5.00</td>
<td>45.20</td>
<td>62981</td>
<td>3.0</td>
<td>DAT</td>
<td>UKMO</td>
<td>1</td>
</tr>
<tr>
<td>-28.00</td>
<td>41.60</td>
<td>62982</td>
<td>999</td>
<td>DAT</td>
<td>UNKNOWN</td>
<td>1</td>
</tr>
<tr>
<td>-12.40</td>
<td>46.70</td>
<td>62983</td>
<td>3.0</td>
<td>DAT</td>
<td>UKMO</td>
<td>1</td>
</tr>
<tr>
<td>-7.00</td>
<td>51.40</td>
<td>62984</td>
<td>999</td>
<td>DAT</td>
<td>UNKNOWN</td>
<td>1</td>
</tr>
<tr>
<td>1.20</td>
<td>53.30</td>
<td>62985</td>
<td>999</td>
<td>DAT</td>
<td>UNKNOWN</td>
<td>1</td>
</tr>
<tr>
<td>-5.60</td>
<td>48.50</td>
<td>62986</td>
<td>999</td>
<td>DAT</td>
<td>METFR</td>
<td>1</td>
</tr>
<tr>
<td>-13.00</td>
<td>51.00</td>
<td>62987</td>
<td>3.0</td>
<td>DAT</td>
<td>UKMO</td>
<td>1</td>
</tr>
<tr>
<td>-11.20</td>
<td>53.13</td>
<td>62988</td>
<td>4.5</td>
<td>DAT</td>
<td>IDT</td>
<td>1</td>
</tr>
<tr>
<td>-5.42</td>
<td>53.47</td>
<td>62989</td>
<td>4.5</td>
<td>DAT</td>
<td>IDT</td>
<td>1</td>
</tr>
<tr>
<td>-18.65</td>
<td>51.67</td>
<td>62990</td>
<td>4.5</td>
<td>DAT</td>
<td>IDT</td>
<td>1</td>
</tr>
<tr>
<td>-9.07</td>
<td>54.67</td>
<td>62991</td>
<td>4.5</td>
<td>DAT</td>
<td>IDT</td>
<td>1</td>
</tr>
<tr>
<td>-6.70</td>
<td>51.69</td>
<td>62992</td>
<td>4.5</td>
<td>DAT</td>
<td>IDT</td>
<td>1</td>
</tr>
<tr>
<td>-15.92</td>
<td>53.06</td>
<td>62993</td>
<td>4.5</td>
<td>DAT</td>
<td>IDT</td>
<td>1</td>
</tr>
<tr>
<td>-2.00</td>
<td>49.00</td>
<td>62994</td>
<td>3.0</td>
<td>DAT</td>
<td>UKMO</td>
<td>1</td>
</tr>
<tr>
<td>-12.30</td>
<td>54.54</td>
<td>62995</td>
<td>3.0</td>
<td>DAT</td>
<td>UKMO</td>
<td>1</td>
</tr>
<tr>
<td>-9.00</td>
<td>57.00</td>
<td>62996</td>
<td>4.5</td>
<td>DAT</td>
<td>IDT</td>
<td>1</td>
</tr>
<tr>
<td>-6.10</td>
<td>58.10</td>
<td>62997</td>
<td>4.5</td>
<td>DAT</td>
<td>IDT</td>
<td>1</td>
</tr>
<tr>
<td>-19.50</td>
<td>53.50</td>
<td>62998</td>
<td>3.0</td>
<td>DAT</td>
<td>UKMO</td>
<td>1</td>
</tr>
<tr>
<td>0.00</td>
<td>57.00</td>
<td>62999</td>
<td>999</td>
<td>DAT</td>
<td>UNKNOWN</td>
<td>1</td>
</tr>
<tr>
<td>8.40</td>
<td>58.10</td>
<td>63110</td>
<td>999</td>
<td>DAT</td>
<td>UNKNOWN</td>
<td>1</td>
</tr>
<tr>
<td>1.00</td>
<td>58.78</td>
<td>63111</td>
<td>999</td>
<td>DAT</td>
<td>UNKNOWN</td>
<td>1</td>
</tr>
<tr>
<td>1.40</td>
<td>57.78</td>
<td>63112</td>
<td>999</td>
<td>DAT</td>
<td>UNKNOWN</td>
<td>1</td>
</tr>
<tr>
<td>0.00</td>
<td>57.00</td>
<td>63113</td>
<td>999</td>
<td>DAT</td>
<td>UNKNOWN</td>
<td>1</td>
</tr>
<tr>
<td>2.00</td>
<td>57.80</td>
<td>63114</td>
<td>999</td>
<td>DAT</td>
<td>UNKNOWN</td>
<td>1</td>
</tr>
<tr>
<td>2.00</td>
<td>56.40</td>
<td>63115</td>
<td>999</td>
<td>DAT</td>
<td>UNKNOWN</td>
<td>1</td>
</tr>
<tr>
<td>1.00</td>
<td>57.00</td>
<td>63116</td>
<td>999</td>
<td>DAT</td>
<td>UNKNOWN</td>
<td>1</td>
</tr>
<tr>
<td>2.00</td>
<td>53.00</td>
<td>63117</td>
<td>999</td>
<td>DAT</td>
<td>PRIVY</td>
<td>1</td>
</tr>
<tr>
<td>1.70</td>
<td>53.40</td>
<td>63118</td>
<td>999</td>
<td>DAT</td>
<td>PRIVY</td>
<td>1</td>
</tr>
<tr>
<td>2.00</td>
<td>53.10</td>
<td>63119</td>
<td>999</td>
<td>DAT</td>
<td>PRIVY</td>
<td>1</td>
</tr>
<tr>
<td>0.50</td>
<td>57.40</td>
<td>63120</td>
<td>999</td>
<td>DAT</td>
<td>PRIVY</td>
<td>1</td>
</tr>
<tr>
<td>0.50</td>
<td>57.20</td>
<td>63121</td>
<td>999</td>
<td>DAT</td>
<td>PRIVY</td>
<td>1</td>
</tr>
<tr>
<td>-4.70</td>
<td>52.30</td>
<td>63122</td>
<td>3.0</td>
<td>DAT</td>
<td>UKMO</td>
<td>1</td>
</tr>
<tr>
<td>-5.10</td>
<td>51.60</td>
<td>63123</td>
<td>3.0</td>
<td>DAT</td>
<td>UKMO</td>
<td>1</td>
</tr>
<tr>
<td>1.00</td>
<td>51.10</td>
<td>63124</td>
<td>3.0</td>
<td>DAT</td>
<td>UKMO</td>
<td>1</td>
</tr>
<tr>
<td>0.80</td>
<td>58.40</td>
<td>63125</td>
<td>3.0</td>
<td>DAT</td>
<td>UKMO</td>
<td>1</td>
</tr>
<tr>
<td>1.70</td>
<td>68.60</td>
<td>63126</td>
<td>999</td>
<td>DAT</td>
<td>UNKNOWN</td>
<td>1</td>
</tr>
<tr>
<td>1.50</td>
<td>59.20</td>
<td>63127</td>
<td>999</td>
<td>DAT</td>
<td>UNKNOWN</td>
<td>1</td>
</tr>
<tr>
<td>1.00</td>
<td>61.20</td>
<td>63128</td>
<td>999</td>
<td>DAT</td>
<td>UNKNOWN</td>
<td>1</td>
</tr>
<tr>
<td>1.00</td>
<td>67.80</td>
<td>63129</td>
<td>999</td>
<td>DAT</td>
<td>UNKNOWN</td>
<td>1</td>
</tr>
<tr>
<td>1.00</td>
<td>59.50</td>
<td>63130</td>
<td>999</td>
<td>DAT</td>
<td>PRIVY</td>
<td>1</td>
</tr>
<tr>
<td>1.00</td>
<td>61.60</td>
<td>63131</td>
<td>999</td>
<td>DAT</td>
<td>PRIVY</td>
<td>1</td>
</tr>
<tr>
<td>-11.80</td>
<td>59.10</td>
<td>64848</td>
<td>3.0</td>
<td>DAT</td>
<td>UKMO</td>
<td>1</td>
</tr>
<tr>
<td>-4.50</td>
<td>68.70</td>
<td>64846</td>
<td>3.0</td>
<td>DAT</td>
<td>UKMO</td>
<td>1</td>
</tr>
</tbody>
</table>

#### Norwegian Sea

<table>
<thead>
<tr>
<th>LON</th>
<th>LAT</th>
<th>NAME</th>
<th>AH</th>
<th>TYPE</th>
<th>SOURCE</th>
<th>SCALE</th>
</tr>
</thead>
<tbody>
<tr>
<td>3.70</td>
<td>68.60</td>
<td>LF48</td>
<td>999</td>
<td>DAT</td>
<td>UNKNOWN</td>
<td>1</td>
</tr>
<tr>
<td>2.00</td>
<td>59.60</td>
<td>LF49</td>
<td>999</td>
<td>DAT</td>
<td>UNKNOWN</td>
<td>1</td>
</tr>
<tr>
<td>7.80</td>
<td>64.30</td>
<td>LF50</td>
<td>999</td>
<td>DAT</td>
<td>UNKNOWN</td>
<td>1</td>
</tr>
<tr>
<td>2.00</td>
<td>61.20</td>
<td>LF51</td>
<td>999</td>
<td>DAT</td>
<td>UNKNOWN</td>
<td>1</td>
</tr>
<tr>
<td>7.80</td>
<td>65.30</td>
<td>LF52</td>
<td>999</td>
<td>DAT</td>
<td>UNKNOWN</td>
<td>1</td>
</tr>
<tr>
<td>1.00</td>
<td>58.40</td>
<td>LF53</td>
<td>999</td>
<td>DAT</td>
<td>UNKNOWN</td>
<td>1</td>
</tr>
<tr>
<td>8.10</td>
<td>66.00</td>
<td>LF54</td>
<td>999</td>
<td>DAT</td>
<td>UNKNOWN</td>
<td>1</td>
</tr>
<tr>
<td>3.20</td>
<td>56.50</td>
<td>LF55</td>
<td>999</td>
<td>DAT</td>
<td>UNKNOWN</td>
<td>1</td>
</tr>
<tr>
<td>LON</td>
<td>LAT</td>
<td>NAME</td>
<td>AH</td>
<td>TYPE</td>
<td>SOURCE</td>
<td>SCALE</td>
</tr>
<tr>
<td>------</td>
<td>------</td>
<td>--------------</td>
<td>-----</td>
<td>-------</td>
<td>--------</td>
<td>-------</td>
</tr>
<tr>
<td>-135</td>
<td>20</td>
<td>TPC51</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>-118</td>
<td>30</td>
<td>TPC50</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>-80</td>
<td>15</td>
<td>TPC23</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>-77</td>
<td>12</td>
<td>TPC22</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>-80</td>
<td>24</td>
<td>TPC25</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>-86</td>
<td>21</td>
<td>TPC26</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>-80</td>
<td>15</td>
<td>TPC20</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>-73</td>
<td>15</td>
<td>TPC19</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>-63</td>
<td>15</td>
<td>TPC18</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>-77</td>
<td>12</td>
<td>TPC17</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>-80</td>
<td>15</td>
<td>TPC16</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>-76</td>
<td>22</td>
<td>TPC15</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>-80</td>
<td>24</td>
<td>TPC14</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>-86</td>
<td>21</td>
<td>TPC13</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>-118</td>
<td>30</td>
<td>TPC58</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>-135</td>
<td>20</td>
<td>TPC51</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>-117</td>
<td>20</td>
<td>TPC52</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>-120</td>
<td>0</td>
<td>TPC55</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>
$ WAM AFOS locations
$  -72.5  35.0   'ZNT41     '  999.   VBY  NCEP      1
$  -70.0  37.5   'ZNT42     '  999.   VBY  NCEP      1
$  -67.5  40.0   'ZNT43     '  999.   VBY  NCEP      1
$  -85.0  25.0   'ZNT45     '  999.   VBY  NCEP      1
$ -125.0  47.5   'ZPZ41     '  999.   VBY  NCEP      1
$ -125.0  45.0   'ZPZ42     '  999.   VBY  NCEP      1
$ -130.0  42.5   'ZPZ43     '  999.   VBY  NCEP      1
$ -122.5  35.0   'ZPZ44     '  999.   VBY  NCEP      1
$ -120.0  32.5   'ZPZ45     '  999.   VBY  NCEP      1
$ -122.5  27.5   'ZPZ46     '  999.   VBY  NCEP      1
$ End of list
$ 4.00  0.00  'STOPSTRING'  999.   XXX  NCEP     99
$ Four additional output types: see ww3_shel.inp for documentation.
$ track output
$ 19680606 000000      0  19680608 000000
$ restart files
$ 19680606 000000      0  19680608 000000
$ boundary output
$ 19680606 000000      0  19680608 000000
$ separated wave field data
$ 19680606 000000      0  19680608 000000
$ 20140901 000000     0  20140902 000000
$ 20140902 000000     1  20140902 000000
$ 20140901 000000     0  20140902 000000
$ 20140901 000000     0  20140902 000000
$ Output requests per grid and type to overwrite general setup as defined above. First record per set is the grid name MODID and the output type number. Then follows the standard time string, and conventional data as per output type. In mww3_test_05 this is not used. Below, one example generating partitioning output for the inner grid is included but commented out.
$ 'grd3'  6
$ 19680606 000000    900  19680608 000000
$     0 999 1 0 999 1 T
$ Mandatory end of output requests per grid, identified by output type set to 0.
$ 'the_end' 0
$ Moving grid data as in ww3_shel.inp. All grids will use same data.
$ 'MOV' 19680606 000000   5.0  90.
$ 'STP'
$ End of input file
A.3.8  ww3_ounf_glob_30m.inp: For the gridded NetCDF output post-processor (Global 0.5 degree grid) [with comments]

```plaintext
$ WAVEWATCH III Grid output post-processing $  
$ First output time (yyyyymmdd hhmmss), increment of output (s), and number of output times. $  
$ 20140901 000000 3600 25 $  
$ Fields requested ------------------------------- $  
$ Output request flags identifying fields as in ww3_shel.inp. See that file for a full documentation of field output options. Namelist type selection is used here (for alternative F/T flags, see ww3_shel.inp). $  
N  
DPT WND ICE HS FP T01 T02 T0M1 DP DIR SPR EF P2S P2L  
$ netCDF version [3,4] $  
$ and variable type 4 [2 = SHORT, 3 = it depends , 4 = REAL] $  
$ swell partitions [0 1 2 3 4 5] $  
$ variables in same file [T] or not [F] $  
$ 4 2 $  
$ 0 1 2 3 4 5 $  
$ T $  
$ File prefix $  
$ number of characters in date [4(yearly),6(monthly),8(daily),10(hourly)] $  
$ IX and IY ranges [regular:IX NX IY NY, unstructured:IP NP 1 1] $  
$ WW3-glob_30m $  
$ 1 720 1 313 $  
$ For each field and time a new file is generated with the file name ww3.date_xxx.nc, where date is a conventional time indicator with 53 characters, and xxx is a field identifier. $  
$ $  
$ End of input file $  
$ $ 
```
A.3.9  ww3_ounf_casc_6m.inp: For the gridded NetCDF output post-processor (Cascadia 0.1 degree grid)

```plaintext
$ WAVEWATCH III Grid output post-processing $  
$ 20140901 000000 3600 25 $  
N DPT WND ICE HS FP T01 T02 T0M1 DP DIR SPR EF P2S P2L $  
4 2 0 1 2 3 4 5 $  
$ WW3-casc_6m $  
8 $  
1 111 1 101 $  
$ $  
End of input file $  
$ $  
$ $  
```
Appendix A

A.3.10  ww3_oup.inp: For the point output NetCDF post-processor
[with comments]

```
$ WAVEWATCH III NETCDF Point output post-processing
$ First output time (yyyymmdd hhmmss), increment of output (s),
$ and number of output times.
$ 20140901 000000  3600.  25

$ Points requested
$ Define points index for which output is to be generated.
$ If no one defined, all points are selected
$ One index number per line, negative number identifies end of list.
 1 2
$ mandatory end of list

$ file prefix
$ number of characters in date [4(yearly),6(monthly),8(daily),10(hourly)]
$ netCDF version [3,4]
$ points in same file [T] or not [F]
$ and max number of points to be processed in one pass
$ output type ITYPE [0,1,2,3]
$ flag for global attributes WW3 [0] or variable version [1-2-3-4]
$ flag for dimensions order time,station [T] or station,time [F]

$ WW3- 8
$ T 500
$ 1 0
$ T 5

$ ITYPE = 0, inventory of file.
$ No additional input, the above time range is ignored.

$ ITYPE = 1, netCDF Spectra.
$ - Sub-type OTYPE :  1 : Print plots.
$  2 : Table of 1-D spectra
$  3 : Transfer file.
$  4 : Spectral partitioning.
$  - Scaling factors for 1-D and 2-D spectra Negative factor
$    disables output, factor = 0. gives normalized spectrum.
$  3  1  0
$ The transfer file contains records with the following contents.
$ - File ID in quotes, number of frequencies, directions and points.
$  grid name in quotes (for unformatted file C*21,3I,C*30).
$ - Bin frequencies in Hz for all bins.
$ - Bin directions in radians for all bins (Oceanographic conv.).
$ - Time in yyyymmdd hhmmss format                         | loop
$  -+           | loop over
$ - Point name (C*10), lat, lon, d, U10 and    |  loop     | over
$   direction, current speed and direction     |   over    | points times
$ - E(f,theta)                                 |  points   | times
$                                             -+          -+

$ ITYPE = 2, netCDF Tables of (mean) parameter
$ - Sub-type OTYPE :  1 : Depth, current, wind
$  2 : Mean wave pars.
$  3 : Nondimensional pars. (Um)
$  4 : Nondimensional pars. (U10)
$  5 : 'Validation table'
$  6 : WMO standard output
$  4

$ ITYPE = 3, netCDF Source terms
$ - Sub-type OTYPE :  1 : Print plots.
$  2 : Table of 1-D S(f),
$  3 : Table of 1-D inverse time scales
$ (1/T = S/F).
$  4 : Transfer file
$  - Scaling factors for 1-D and 2-D source terms. Negative factor
$    disables print plots, factor = 0. gives normalized print plots.
$  - Flags for spectrum, input, interactions, dissipation,
$    bottom and total source term.
$  - scale ISCALE for OTYPE=2,3
$  0 : Dimensional.
$  1 : Nondimensional in terms of U10
$  2 : Nondimensional in terms of U*
$  3-5: like 0-2 with f normalized with fp.
$  4 0 0 0 0 0 0 0
```
The transfer file contains records with the following contents.

- File ID in quotes, number of frequencies, directions and points, flags for spectrum and source terms (C*21, 3I, 6L)
- Bin frequencies in Hz for all bins.
- Bin directions in radians for all bins (Oceanographic conv.).
- Time in yyyymmdd hhmmss format
- Point name (C*10), depth, wind speed and direction, current speed and direction
- E(f,theta) if requested
- Sin(f,theta) if requested
- Snl(f,theta) if requested
- Sds(f,theta) if requested
- Sbt(f,theta) if requested
- Stot(f,theta) if requested

End of input file

--------------------------------------------------------------------

$ The transfer file contains records with the following contents.$$ - File ID in quotes, number of frequencies, directions and points, flags for spectrum and source terms (C*21, 3I, 6L)$$ - Bin frequencies in Hz for all bins.$$ - Bin directions in radians for all bins (Oceanographic conv.).$$ - Time in yyyymmdd hhmmss format$$ - Point name (C*10), depth, wind speed and direction, current speed and direction$$ - E(f,theta) if requested$$ - Sin(f,theta) if requested$$ - Snl(f,theta) if requested$$ - Sds(f,theta) if requested$$ - Sbt(f,theta) if requested$$ - Stot(f,theta) if requested$$ $ End of input file$ $ $ $
Appendix B

Section B.1 of this Appendix contains additional material for Chapter 2 which were given as Supporting Information in the published paper. Section B.2 contains the formatted paper published in the Journal of Geophysical Research: Oceans in September 2015.
B.1 Additional figures and tables for Chapter 2

Figure B.1: Station map with lines connecting station pairs that have separation distances greater than 30 km and less than 120 km.
Figure B.2: Weighted backprojection of September 2012-May 2013 stack. First, second and third row for 100 s, 150 s, and 200 s. Column 1: Backprojection of stack. Column 2: Response of array to isotropic wavefield. Column 3: Weighted stack.
Figure B.3: Beamforming of September 2012-May 2013 stack. First, second and third row for 100 s, 146 s, and 200 s. Column 1: Beamforming of stack. Radial axis is slowness in s/km, angular axis is azimuth of wave arrival. Column 2: Response of array to isotropic wavefield. Column 3: Beamform output at slowness of 12.3704 s/km (for 100 s), 8.22 s/km (for 146 s), 7.7037 s/km (for 220 s) projected onto map of Pacific.
Figure B.4: Backprojection for the days with westerly arrivals. (a) 18 January 2013 (b) 25 October 2012 (c) 4 April 2013 (d) 17 January 2013 (e) 11 May 2013. Columns left to right are periods of 100 s, 125 s, 150 s, 175 s, 200 s.
Figure B.5: Examples of unclear cross-correlations.
Figure B.6: Same as Figure 2.9 but only station pairs which showed clear peaks at $\pm t_1$ are plotted and used in the calculations of the mean.
Figure B.7: Wave spectra from buoy 46015 on a) 25 October 2012 b) 4 April 2013 c) 17 January 2013 d) 18 January 2013 e) 11 May 2013. The shaded regions show the range of the hourly spectra over the day.
Figure B.8: Cross-correlation function from Harmon et al. (2012). a) Stacked cross-correlation of stations 42-45 from Harmon et al. (2012) band-pass-filtered between 60 and 500 s. b) Projection of the cross-correlation onto $t_1 + 2t_2$ travel time grid (IR peak). For the projection, the positive and negative lags were stacked. c) Projection of (b) zoomed in to show relation of projection to bathymetry. The red line is on an azimuth that connects Station 1 and Station 2.
Table B.1: Stations of the Cascadia Array used in the study.

<table>
<thead>
<tr>
<th>Stations Used</th>
<th>Latitude</th>
<th>Longitude</th>
<th>Depth (m)</th>
</tr>
</thead>
<tbody>
<tr>
<td>FS05B</td>
<td>40.3866</td>
<td>-124.8996</td>
<td>2316</td>
</tr>
<tr>
<td>FS06B</td>
<td>40.3812</td>
<td>-124.7853</td>
<td>2198</td>
</tr>
<tr>
<td>FS09B</td>
<td>40.4387</td>
<td>-124.8085</td>
<td>2161</td>
</tr>
<tr>
<td>G02B</td>
<td>40.0486</td>
<td>-125.2969</td>
<td>1920</td>
</tr>
<tr>
<td>G04B</td>
<td>40.0585</td>
<td>-126.9285</td>
<td>4368</td>
</tr>
<tr>
<td>G05B</td>
<td>40.0706</td>
<td>-127.7478</td>
<td>4462</td>
</tr>
<tr>
<td>G10B</td>
<td>40.6779</td>
<td>-125.5533</td>
<td>2936</td>
</tr>
<tr>
<td>G11B</td>
<td>40.6875</td>
<td>-126.3764</td>
<td>3123</td>
</tr>
<tr>
<td>G13B</td>
<td>40.6840</td>
<td>-128.0283</td>
<td>3215</td>
</tr>
<tr>
<td>G20B</td>
<td>41.2987</td>
<td>-126.6133</td>
<td>3141</td>
</tr>
<tr>
<td>G21B</td>
<td>41.3160</td>
<td>-127.4537</td>
<td>3156</td>
</tr>
<tr>
<td>G22B</td>
<td>41.3092</td>
<td>-128.2740</td>
<td>3038</td>
</tr>
<tr>
<td>G28B</td>
<td>41.9428</td>
<td>-126.7339</td>
<td>3327</td>
</tr>
<tr>
<td>G29B</td>
<td>41.9772</td>
<td>-127.4834</td>
<td>3197</td>
</tr>
<tr>
<td>G30B</td>
<td>41.9565</td>
<td>-128.3198</td>
<td>3119</td>
</tr>
<tr>
<td>G35B</td>
<td>42.5658</td>
<td>-126.0558</td>
<td>2367</td>
</tr>
<tr>
<td>G36B2</td>
<td>42.5994</td>
<td>-126.9038</td>
<td>2423</td>
</tr>
<tr>
<td>J06B</td>
<td>43.2652</td>
<td>-128.7997</td>
<td>3220</td>
</tr>
<tr>
<td>J09B</td>
<td>40.2011</td>
<td>-124.7270</td>
<td>252</td>
</tr>
<tr>
<td>J11B</td>
<td>43.5414</td>
<td>-126.3673</td>
<td>3030</td>
</tr>
<tr>
<td>M11B</td>
<td>42.9320</td>
<td>-125.0171</td>
<td>1109</td>
</tr>
<tr>
<td>M12B</td>
<td>42.1840</td>
<td>-124.9461</td>
<td>1045</td>
</tr>
<tr>
<td>M14B</td>
<td>40.9850</td>
<td>-124.5897</td>
<td>638</td>
</tr>
</tbody>
</table>

Stations not used due to distance from other stations

<table>
<thead>
<tr>
<th>Stations not used due to distance from other stations</th>
<th>Latitude</th>
<th>Longitude</th>
<th>Depth (m)</th>
</tr>
</thead>
<tbody>
<tr>
<td>J19B</td>
<td>44.1777</td>
<td>-126.2717</td>
<td>2968</td>
</tr>
<tr>
<td>J23B</td>
<td>44.8459</td>
<td>-129.6821</td>
<td>2650</td>
</tr>
<tr>
<td>J27B</td>
<td>44.8476</td>
<td>-126.3069</td>
<td>2824</td>
</tr>
<tr>
<td>J28B</td>
<td>45.0631</td>
<td>-127.1552</td>
<td>2866</td>
</tr>
<tr>
<td>J48B</td>
<td>47.1282</td>
<td>-130.6385</td>
<td>2877</td>
</tr>
<tr>
<td>J63B</td>
<td>48.2054</td>
<td>-130.0019</td>
<td>2857</td>
</tr>
</tbody>
</table>

Stations not used due to data quality

<table>
<thead>
<tr>
<th>Stations not used due to data quality</th>
<th>Latitude</th>
<th>Longitude</th>
<th>Depth (m)</th>
</tr>
</thead>
<tbody>
<tr>
<td>FS01B</td>
<td>40.3268</td>
<td>-124.9492</td>
<td>940</td>
</tr>
<tr>
<td>FS14B</td>
<td>40.4955</td>
<td>-124.5918</td>
<td>107</td>
</tr>
<tr>
<td>G03B</td>
<td>40.0580</td>
<td>-126.1634</td>
<td>4051</td>
</tr>
<tr>
<td>G12B</td>
<td>40.6809</td>
<td>-127.2289</td>
<td>3080</td>
</tr>
<tr>
<td>G19B</td>
<td>41.3074</td>
<td>-125.7736</td>
<td>3071</td>
</tr>
<tr>
<td>G27B</td>
<td>41.9166</td>
<td>-126.0167</td>
<td>3480</td>
</tr>
<tr>
<td>G37B</td>
<td>42.5913</td>
<td>-127.7213</td>
<td>3004</td>
</tr>
<tr>
<td>J10B</td>
<td>43.3494</td>
<td>-125.5435</td>
<td>3093</td>
</tr>
<tr>
<td>J18B</td>
<td>44.0083</td>
<td>-125.4660</td>
<td>3047</td>
</tr>
<tr>
<td>J20B</td>
<td>44.3542</td>
<td>-127.0952</td>
<td>2934</td>
</tr>
</tbody>
</table>
B.2 Published paper and Supporting Information of Chapter 2

Source regions and reflection of infragravity waves offshore of the U.S.s Pacific Northwest
Jennifer Neale1, Nicholas Harmon1, and Meric Srokosz2
1Ocean and Earth Science, National Oceanography Centre Southampton, University of Southampton, Southampton, UK, 2National Oceanography Centre, University of Southampton Waterfront Campus, Southampton, UK

Abstract
Infragravity waves are oceanic surface gravity waves but with wavelengths (tens of km) and periods (10-30 s) much longer than wind waves and swell. Mostly studied in shallow water, knowledge of infragravity waves in deep water has remained limited. Recent interest in deep water infragravity waves has been motivated by the error they may contribute to future high-resolution satellite radar altimetry measurements of sea level. Here deep water infragravity waves offshore of the Pacific Northwest of the U.S. were studied using differential pressure gauges which were deployed as part of the Cascadia Initiative array from September 2012 to May 2013. Cross correlation of the records revealed direction of infragravity wave propagation across the array, from which source regions were inferred. The dominant source was found to be the coastline to the east, associated with large wind waves and swell incident on the eastern side of the basin. The source shifted southward during northern-hemisphere summer, and on several days in the record infragravity waves arrived from the western side of the Pacific. Asymmetry of cross-correlation functions for five of these westerly arrivals was used to calculate the ratio of seaward to shoreward propagating energy, and hence estimate the strength of infragravity wave reflection at periods of 100–200 s. Reflection of these remote arrivals from the west appeared to be strong, with a lower bound estimate of $r = 0.49 \pm 0.06$ and an upper bound estimate of $r = 0.74 \pm 0.06$. These results suggest that reflection at ocean boundaries may be an important consideration for infragravity waves in the deep ocean.

1. Introduction
Low-frequency infragravity waves are associated with wave groups of the higher-frequency sea waves and swell in the coastal zone. Two mechanisms have been proposed for the generation of infragravity waves from the short-wave groups. One is that the interaction of shoreward propagating swell creates “bound” or “forced” infragravity waves (Longuet-Higgins and Stewart, 1962, Herbers et al., 1995a). As the swell waves break, the forced infragravity waves are released as free infragravity waves and are reflected from the beach. Free infragravity waves satisfy the dispersion relation for surface gravity waves and have longer wavelengths than forced waves of the same frequency (Webb et al., 1991). The second mechanism is that infragravity waves are generated by a time-varying breakpoint, with standing waves shoreward of the breakpoint and progressive infragravity waves radiating seaward (Symonds et al., 1982).

In either case, the seaward propagating free infragravity waves can have two fates: those that travel seaward at oblique angles can become refractively trapped along the shoreline as “edge waves” by a sloping beach or shelf (Herbers et al., 1995a, Munk et al., 1964); those that propagate directly seaward can escape into the open ocean as “leaky waves” (Munk et al., 1964). Due to their long wavelength (Xu and Ardhuin, 2013), only a small fraction of the infragravity energy escapes from the coast into the open ocean (<1%) [Webb et al., 1991], with most being trapped within a few hundred meters of the shore. The amount of energy leaked into the open ocean for a given short-wave spectrum and coastline is poorly understood [Aucan and Ardhuin, 2013]. Variation in alongshore topography may be partly responsible (Uchiyama and Godin et al., 2008), although the model of Ardhuin et al. (2014) produces a good prediction of measured infragravity wave levels assuming a locally straight coast.

Infragravity waves that make it into the open ocean propagate with very little attenuation (Godin et al., 2013), and it is possible to observe infragravity waves that have been generated from coasts thousands of kilometers away on the other side of an ocean basin (Herbers et al., 1995a, Normon et al., 2012).
Most studies of infragravity waves have been undertaken in shallow water on continental shelves where they are most energetic [Webb et al., 1991] and instrumentation is more accessible. Here they are also known as “surf beat” or “swash,” and they are important for sediment transport and waternere morphology [Aagaard and Greenwood, 2008; Remene, 2004] and harbor oscillations [Ochiho and Guza, 1996]. The first studies were undertaken by Munk [1949] and Tucker [1950].

Infragravity waves in the deep ocean have received less attention than shallow water infragravity waves partly due to their very small amplitudes in the deep ocean (< 1 cm) [Webb et al., 1991], several cm at most [Aucan and Ardhuin, 2013]. However, there has recently been a resumed interest in infragravity waves in the deep ocean as they have been recognized as important for coupling processes in the ocean, ice, atmosphere, and solid earth [Godin et al., 2013]. Aucan and Ardhuin [2013] have shown that infragravity waves in the deep ocean may add significant error to sea level measurements associated with submesoscale currents, which are due to be collected by future satellite radar altimetry missions. Bronmivi et al. [2010] have recently shown that infragravity waves generated along the Pacific coast propagate transoceanic distances and can be implicated in the failure and subsequent breakup of Antarctic ice shelves. Infragravity waves at frequencies below 0.004 Hz may transfer energy from the ocean to the atmosphere [Crawford et al., 2007; Godin et al., 2015]. The deformation of the seafloor under the pressure of infragravity waves is used in measurements of seafloor compliance to determine the shear velocity structure of the shallow oceanic crust [Battjes, 1977], and the propagation of infragravity waves over a sloping seabed are thought to create low-frequency seismic noise known as Earth’s seismic hum [Rhe and Romanowicz, 2006; Ardhuin et al., 2015].

Pressure sensors (or seismometers) deployed on the seafloor have been the most widely used approach to observe infragravity waves in the deep ocean [Godin et al., 2013]. Using an array of pressure gauges in the southwestern Pacific off the South Island of New Zealand, Godin et al. [2014] observed strong directionality of the infragravity wavefield with the northwest coast of the South Island acting as a net source of infragravity wave energy. Webb et al. [1991] studied deep water infragravity waves in the Pacific during November 1988 and identified infragravity waves originating from the Gulf of Alaska, the northwest Pacific and the southern tip of South America, but little from the southern ocean or tropical western Pacific.

A further and more comprehensive study of infragravity waves in the deep ocean was undertaken by Aucan and Ardhuin [2013]. They analyzed pressure records from 40 locations in the Pacific and Atlantic oceans to determine spatial and temporal variability of infragravity wave energy at depths of 3–6 km. Their inferred infragravity significant wave heights were found to reach larger values than estimated in previous work, reaching over 4 cm in episodic events. Energy levels in the Atlantic and Pacific were found to be similar, and mid to high latitudes in both oceans displayed strong seasonal cycles associated with seasonal variability of wind waves.

Other studies have noted the arrival of infragravity waves which seem to have been generated right across the other side of the ocean basin [Rammon et al., 2012], and a combined observational and modeling study [Ravat et al., 2014] has shown the coherent propagation of large infragravity wave bursts from one side of the basin to the other. The latter study made use of a global numerical model of free infragravity wave generation and propagation that has been under development recently [Ardhuin et al., 2014].

The aim of this study was to determine how strongly free infragravity waves reflect when they reach the coastline or shelf of an ocean basin. As far as the authors are aware, no estimate of deep water infragravity reflection has yet been made, although reflection from the shoreline has previously been estimated from pressure gauges in shallow waters (< 13 m in depth) [e.g., Herbers et al., 1995a; Sheremet, 2002] and from laboratory data [Battjes, 2004]. Studies such as these have found that infragravity waves reflect strongly from the shore with reflection coefficients above 0.6. Considering that infragravity waves are capable of propagating right across the oceans, reflection at the ocean boundary, whether at the shoreline or shelf, may be important for infragravity energy in the deep ocean. Here we present estimates of the directionality and reflection coefficient for the infragravity wavefield offshore of the Pacific Northwest of the U.S.

2. Data and Methods

To measure deep water infragravity waves, we used differential Pressure Gauge (DPG) records from the Cascadia Initiative array [Toomey et al., 2014] between September 2012 and May 2013, downloaded from the
The array consisted of 39 DPGs between depths of 107 and 4462 m offshore of the Pacific Northwest of the U.S. (Figure 1). The locations were considered far enough offshore (50–500 km) to be removed from the effects of infragravity edge waves at the coast, which are trapped within a few hundred meters of the shore [Webb et al., 1991].

Monthly spectra of the records were used to identify bad data. In total, 29 stations returned usable data over the whole data period and these stations were used in the study. The station locations are listed in supporting information Table S1. For most of our analysis we exclude the most northerly stations (above 44°N), which fall outside the main cluster of stations.

The daily pressure records were band pass filtered between 0.002 and 0.45 Hz using a second-order Butterworth filter prior to decimation to 1 Hz, then detrended and tapered. In order to characterize local infragravity wave generation we also examined near-shore short-wave parameters using data from the National Data Buoy Center’s data buoys 46015, 46022, and 46027 (http://www.ndbc.noaa.gov), which are also shown in Figure 1. Daily significant wave height, Hₕ, and average wave period, Tₘ₀, were calculated for each buoy from the daily average spectra as follows:

\[ Hₕ = 4\sqrt{m₀} \]  
\[ Tₘ₀ = \sqrt{m₀/m²} \]  
\[ m₀ = \sum_{n=0}^{n} S(f)df \]

where \( f \) is frequency in Hz, \( S(f) \) is the non-directional wave spectrum, \( df \) is the bandwidth of each frequency band, \( m₀ = 0.0325 \) Hz, and \( f₀ = 0.4850 \) Hz. We use \( Hₕ \) as a proxy for local infragravity wave generation because previous studies have found a high correlation between \( Hₕ \) and infragravity wave height in shallow water [Herbers et al., 1995b]. In addition, we calculated \( m₀/Tₘ₀ \sqrt{gD/a} \) for each buoy where \( g = 9.81 \) m s⁻², \( D \) is water depth (m), and \( a \) is a dimensional constant with units of s⁻¹, because this parameter has been found to improve the correlation between infragravity waves and short-wave conditions [Ardhuin et al., 2014]. With \( a = 12 \times 10^{-5} \) s⁻¹, the parameter empirically models the observed free infragravity wave height [Ardhuin et al., 2014; Rowet et al., 2014]. For our purposes, the value of \( a \) does not matter as we are interested in the relative change of infragravity wave generation with short-wave conditions rather than
absolute wave heights, but we used $s = 12 \times 10^{-3} \text{ s}^{-1}$ so that the modeled infragravity wave heights can be compared with other studies. We used both these measures, $H_s$ and $\sqrt{H_s T}$ as proxies for local infragravity wave generation, and we averaged $H_s$ and $\sqrt{H_s T}$ over the three wave buoys.

2.1. Cross Correlation

A cross-correlation function was computed between each DPG station pair on each day to aid in identifying coherent signals between each station pair. A stack over the data period was computed by summing the daily cross-correlation functions to get a sense of the long-term average wavefield (Figure 2). Each cross-correlation function was band pass filtered to a central frequency $f = 0.0015 \text{ Hz}$ (a wider range was found to make the records too spiky, and a narrower range too smooth, to identify the main peaks), and the envelope of the signal was calculated using a Hilbert transform. The central frequencies used were $0.0100$, $0.0080$, $0.0067$, $0.0057$, and $0.0050 \text{ Hz}$ corresponding to periods of 100, 125, 150, 175, and 200 s which covered the main infragravity band. The differential pressure gauges were uncalibrated, so each cross-correlation function was normalized by the maximum of its envelope. Figure 3 shows examples of these band-pass filtered cross-correlation functions.

Asymmetry of the cross-correlation functions gives information on the direction of infragravity wave propagation across the array. For example, if energy travels from Station 1 to Station 2, the time series at Station 2 will lag the time series at Station 1 by $t_1$ seconds, where $t_1$ is the time it takes for energy to travel between the two stations, and a peak in the cross-correlation function at $t_1$ will result. Likewise, if energy is traveling from Station 2 to Station 1, the time series at Station 1 will lag Station 2 by $t_1$, resulting in a peak at $-t_1$ in the cross-correlation function. If energy is traveling perpendicular to the Station 1 to Station 2 alignment, a peak at zero lag would be expected, as both stations receive the signal at the same time. Two methods were used to combine the information contained in all the individual cross-correlation functions from the array: backprojection and beamforming.

As beamforming did not require calculation of unique travel time grids for each frequency, as was necessary for backprojection, it allowed us to examine directionality over many frequencies, and was more appropriate for analyzing temporal changes in wave propagation over the array from sources outside of the array. However, due to uncertainties about the quality of beamforming over varying bathymetry, backprojection was used to verify the beamforming results, and was useful for examining sources along the coast close to the array.

2.2. Backprojection

Backprojection of the infragravity wave energy allows us to examine the spatial distribution of wave generation inside and outside of the array and to determine the direction of energy propagation across the array.
For each frequency of interest, $f$, the enveloped cross-correlation functions were back-projected onto a spatial grid of latitude, $u$, and longitude, $c$, using a method similar to that used by Harmon et al. [2012] and Brzak et al. [2009], and given in equation (4):

$$P(f, l) = \sum_{n=1}^{N} W_n(f) \text{env}(C_n(f, l)),$$

(4)

$P$ is the backprojection as a function of frequency and location index, $(f, l)$, where $l$ is the index of each unique latitude and longitude point on our spatial grid $(1, 2, \ldots, L)$ grid points). $\text{env}(C_n)$ is the enveloped band-pass-filtered cross correlation (with center frequency $f_c = 0.0015$ Hz) for station pair $n$. The envelope was calculated using a Hilbert transform, and the maximum value for the envelope was normalized to 1. $W_n$ is a weighting coefficient for station pair $n$, described below, to reduce the effect of array geometry on the projection. For an unweighted backprojection, $W_n = 1$ for all $n$ and $f$. $T_n$ is the theoretical group lag time for station pair $n$ for a hypothetical source at $l$ with frequency $f$. $T_n$ is calculated using

$$T_n(f, l) = t_i(f, l) - t_j(f, l),$$

(5)

Figure 3. Examples of stacked cross-correlation functions for east-west-aligned station pairs filtered to (top row) 100 s, (middle row) 150 s, and (bottom row) 200 s. Vertical lines are plotted at $t_1$, the theoretical group travel time $T_1$ between the two stations for an infragravity wave at the given period travelling along the direct raypath between the two stations. Positive lags represent waves propagating from Station 1 to Station 2, and in all these cases Station 2 is furthest offshore.
where $t_i (t_j)$ is the group travel time from the source at $l$ with frequency $f$ to station $i (j)$ of the station pair.

Group travel times between the source and station are minimum direct travel times calculated using a ray theoretical approach following Harmon et al. [2012]. This approach is more accurate than a calculation from interstation distance/average group velocity along the great circle path between the two stations, as it takes into account bathymetry and the effects of nongreat circle propagation paths. $t_i (t_j)$ were calculated in the following way: First, group velocity of infragravity waves at each frequency of interest at each grid point, $v_g (f, l)$, was calculated using ETOP01 bathymetry [Amante and Eakins, 2009] and the dispersion relation:

$$\omega^2 = gk \tanh (kh), \quad (6)$$

$$v_g = \frac{\partial \omega}{\partial k}. \quad (7)$$

where $\omega = 2\pi f$ = angular frequency (radians per second), $g$ = acceleration due to gravity (m s$^{-2}$), $k$ = wavenumber (radians per m), $h$ = water depth (m), and $v_g$ = group velocity (m s$^{-1}$).

Second, the group velocity grids $v_g (f, l)$ and station locations were input into an Eikonal travel time solver [Rawlinson and Sambridge, 2004] which output travel times, $t (f, l)$, from each station to each grid point at each frequency of interest.

To obtain our backprojection $P(f, l)$, we sum the envelopes of the individual band-pass-filtered cross correlations, multiplied by $W$, at their respective travel time for hypothetical source at $l$ (equation (4)).

Backprojection was computed over a spatial grid of $35^\circ N \leq \phi < 50^\circ N, 135^\circ W < \gamma < 124^\circ W$ at a spatial resolution of $0.0167^\circ (1$ arc nin) and at frequencies of $0.0100, 0.0080, 0.0067, 0.0057,$ and $0.0050$ Hz (100, 125, 150, 175, and 200 s). Daily cross correlations became less clear at station separation distances below 50 km and above 120 km, so we only used station pairs within this range (shown in supporting information Figure S1). The output of the backprojection technique is a map for each frequency, as shown in Figure 4a.

2.2.1. Backprojection for Isotropic Source Distribution

In order to be certain that the backprojection results were not an artifact of the array geometry, we calculated what the backprojection results would be if the array was subjected to an isotropic wavefield. Backprojection for an isotropic source distribution, $I(f, l)$, was computed from the theoretical cross-correlation function, $R(f, l)$, which has two symmetrical impulses at the positive and negative lags corresponding to the group arrival time of the raypath between the two stations for the given frequency. The theoretical isotropic cross-correlation for each station pair (band pass filtered, enveloped, and normalized to a maximum of 1) can then be backprojected using equation (4) but replacing $C_n$ with $R_n$ and using $W = 1$.

Figure 4. Backprojection of September 2012 to May 2013 stack at 150 s. (a) Backprojection of stack. (b) Backprojection for isotropic source distribution. (c) Weighted stack.

NEALE ET AL. INFRAGRAVITY WAVES OFF PACIFIC NORTHWEST 6479
\[ I(f,l) = \sum_{n=1}^{N} W_n(f) \text{env}(R_n(f,T_n(f,l))). \] (8)

From this, we can see which locations the array would illuminate as sources even if all sources were equal.

Figure 4b shows the isotropic response for a period of 150 s.

### 2.2.2. Backprojection Weighting

The theoretical isotropic cross correlations were also used to calculate a weighting coefficient \( W_n \) for each unique station pair \( (n = 1, 2, \ldots, N \) station pairs) for the backprojection using a least squares regression, seeking to minimize equation (9) below [Widrow et al., 1967; Applebaum and Chapman, 1976]:

\[ \min_{w_j} |G \cdot W - P|^2, \quad x \geq 0. \] (9)

\( P_j \) is 1 for each unique latitude and longitude point on our map \( (l = 1, 2, \ldots, L \) location points). This characterizes an ideal isotropic backprojection for all sources being equal. \( G_{\text{iso}} \) contains the isotropic backprojection of each station pair \( n \) (i.e., equation (8) before the summation over all station pairs). \( W_n \) contains the resulting weighting coefficient for each station pair:

\[ G_{\text{iso}} = \text{env} \left[ R_n(T_n(l)) \right]. \] (10)

\[ \begin{bmatrix} p_1 \\ p_2 \\ \vdots \\ p_L \end{bmatrix} = \begin{bmatrix} 1 \\ 1 \\ \vdots \\ 1 \end{bmatrix} W = \begin{bmatrix} w_1 \\ w_2 \\ \vdots \\ w_L \end{bmatrix}. \] (0)

The backprojection can then be calculated with the solution \( W \) using equation (4) which we call a weighted backprojection. An example is given in Figure 4c.

### 2.3. Beamforming

Beamforming was another method used to estimate the direction of energy propagation across the array. It is similar to backprojection but gives beamformer output as a function of slowness (reciprocal of velocity) and azimuth rather than location, and identifies waves propagating across the array from outside sources.

At each frequency of interest, we generate our beamformer output, \( B \), as function of angular frequency, \( \omega \), group slowness, \( s \), and plane wave back azimuth, \( \phi \), in the following way and given in equation (11):

\[ B(\omega, s, \phi) = \sum_{n=1}^{N} \text{env}(C_n(\omega, T_n(s, \phi))). \] (11)

\( C_n \) is the observed cross correlation for the station pair \( n \) band pass at a center frequency, \( \omega \). We generate \( T_n \), a synthetic plane wave group travel time for station pair \( n \), at each slowness and back azimuth of interest. \( T_n \) is calculated using

\[ T_n(\omega, s, \phi) = \frac{d_n}{s} \cos(\phi - \phi_n). \] (12)

in terms of interstation distance \( d_n \), and azimuth \( \phi \).

We sum the envelopes (env in equation (11), calculated using a Hilbert transform) of the individual band-pass filtered cross correlations with a center frequency, \( \omega_i \), at their respective synthetic travel time. We use a Gaussian band-pass filter \( \exp(-x^2) \times (\omega - \omega_i)/(\omega_o) \), where \( x = 100 \times \sqrt{3}/1000 \) where \( \omega \) is
interstation distance, $\omega$ is angular frequency, and $\omega_c$ is the center angular frequency. We examined from 0.05 to 0.167 Hz, 0 to 360° back azimuth, and 20 to 100 s/km slowness. An example of beamformer output is shown in Figure 5a.

The maximum of the beamformer output identifies the slowness and azimuth of the dominant wavefield for each frequency of interest. However, we note the wavefield may be more complicated than a single plane wave for a given day, with a curved wavefront or multiple components. In the beamformer output, complications present themselves as very broad maxima in azimuth centered on a given slowness or multiple local maxima at the best fitting average array slowness.

The slowness is not so important for our purposes and only verifies that the beamforming is picking up waves that are traveling at the expected velocity for a given frequency and water depth. Our results focus on the azimuth of the dominant wavefield at each frequency of interest.

Our beamforming method works best when there are no large changes in bathymetry (hence slowness) across the array, so we limited the beamforming to stations in deep water west of 126°W where changes in velocity across the array are small (maximum 10% between station pairs).

We also performed the beamforming on theoretical cross correlations for an isotropic distribution of plane waves of a given frequency $R_n(x, t)$ was generated as two symmetrical peaks of value 1 at $t = \pm \left(\frac{s}{2C} + \frac{d_x}{2}\right)$ and 0 at all other $t$. $s$ was obtained as the slowness of maximum power from our backprojection results for the given frequency. The theoretical isotropic cross correlation $R_n$ for each station pair (band pass filtered, enveloped, and normalized to a maximum of 1) was then input into equation (11) in place of the observed cross correlation $C_n$. The ideal result in this case would be an equal distribution of energy across all azimuths at the given slowness. Inevitably, the array geometry results in some azimuths being more sensitive than others, but by comparing our results to the isotropic case, we can make sure our results are not an artifact of the array and the processing. The isotropic response at 146 s is shown in Figure 5b.

2.4. Effects of Normalization on Directionality Estimates
Since the amplitudes of the individual DPG’s are unknown, we normalized each record by the envelope of each trace prior to cross correlation, effectively only retaining the instantaneous phase information. This
procedure will tend to amplify weaker sources and mute the strongest sources. Therefore, our results will tend to pick out the most coherent wavefield, not necessarily the largest so our directionality estimates are biased in this sense. However, the relative amplitudes on a given cross correlation should accurately represent the relative amount of energy propagating in one direction versus another.

2.5. Calculating Reflection Coefficients

2.5.1. Seaward: Shoreward Wave Propagation

For an offshore station pair aligned perpendicular to a coastline, each side of the cross-correlation function represents either shoreward propagating energy or seaward propagating energy. The peaks in the positive and negative enveloped cross correlation at \( \pm t_1 \) gives the ratio of seaward:shoreward energy propagation. In this case, the coastline lies approximately north-south and so east-west-aligned station pairs (azimuth 265°–275° or 85°–95°, shown in Figure 1) were used in the analysis of seaward:shoreward infragravity wave propagation. Again, only station pairs separated by distances >30 and <120 km were used as the quality of the daily cross-correlation functions became much reduced outside this range. This limited the analysis to adjacent stations, and all remaining station pairs were between 60 and 75 km apart.

2.5.2. Assumption of No Local Infragravity Wave Generation

Shoreward propagating infragravity waves were considered to be remote arrivals, while seaward propagating infragravity waves may have been due to reflections at the coast or may have been locally generated "leaky" infragravity waves. In practice, it is difficult to separate the cause of the seaward propagating waves—but previous studies have found that infragravity wave generation is small when either nearshore \( H_s \) [Herbers et al., 1995b] or \( sH_s^p_{max} - 2\sqrt{\frac{g}{D}} \) [Ardhuin et al., 2014] are small. Here we have used both measures, \( H_s \) and \( sH_s^p_{max} - 2\sqrt{\frac{g}{D}} \) and assumed that when these were low, the amount of leaky infragravity waves was minimal. In this case, with local generation assumed to be zero, seaward propagation is reflected shoreward incident energy only (see Figure 6). This reflection may involve both specular reflection and scattering, as well as loss of energy through bottom friction and other processes, so the reflection coefficients we calculate contain the net effect of these processes.

Since the amplitude of the reflected wave \( R \) equals the amplitude of the incident wave \( I \) multiplied by the reflection coefficient \( r \), i.e., \( R = rI \):

\[
R^s = R^p \frac{r^s}{r^p} \frac{1}{\frac{H_s^{max}}{H_s^{max}_s}}
\]

Equations (13) and (14) apply to the case illustrated in Figure 6 for a single wave arrival and reflection. The peaks of observed cross-correlations result from waves propagating from multiple directions [Snedder, 2004], but the result still holds and has been shown formally by Wapenaar and Thorbecke [2013] and Godin et al. [2014].

This assumption of no local generation at low \( H_s \) or \( sH_s^p_{max} - 2\sqrt{\frac{g}{D}} \) may not be perfect, so the reflection coefficients calculated should be considered an upper bound.

2.6. Cases of Reflection

To calculate a reflection coefficient, the data were scanned for days when (1) nearshore \( H_s \) or \( sH_s^p_{max} - 2\sqrt{\frac{g}{D}} \) were low (so that locally generated leaky infragravity waves could assumed to be negligible and (2) a strong arrival of infragravity energy from the west (240°–360°) was observed (so that the arrival and reflection are both observable on an east-west-aligned station pair). The 3 days that best matched these conditions were 17 January 2013, 18 January 2013, and 11 May 2013. These days were analyzed along with another 2 days, 25 October 2012 and 4 April 2013, when arrivals from the west were present but less clear, and \( H_s \) and \( sH_s^p_{max} - 2\sqrt{\frac{g}{D}} \) were (mostly) higher.

For each of the 5 days, the reflection coefficient was calculated from each east-west station pair using equation (14). Values for seaward and shoreward infragravity energy were taken as the peak in the enveloped cross-correlation function at around \( \pm t_1 \). To account for a wave arrival slightly oblique (\( \pm 30° \)) to the station
pair alignment (270°) and velocity errors/scattering, which would result in a peak slightly off \( t_1 \), the maximum value in a 200 s window around \( \pm t_1 \) was used as the peak. Varying this window between 100 and 300 s made very little difference to the results.

3. Results

3.1. Infragravity Wave Energy and Directionality

Figure 2 shows the stacked cross correlations filtered between 60 and 500 s, and Figure 3 shows examples from east-west-aligned station pairs filtered to 100, 150, and 200 s. Peaks at \( \pm t_1 \) of an infragravity wave group between the two stations confirms that over the year there is coherent infragravity wave energy propagating both seaward and shoreward, with more going seaward.

From backprojection and beamforming, the dominant source of infragravity energy was found to be from the coastline to the east/northeast, consistent with local generation, as seen in Figures 4c and 5c (plots for other periods are similar can be seen in supporting information Figures S2 and S3). Backprojection (Figure 4c) highlighted the stretch of coastline between 40°8'N and 44°8'N as the dominant source, while beamforming (Figure 5c) identified the region between 42°8'N and 46°8'N as the main source. The differences are small but may be due to the fact that backprojection used additional stations nearer the coast and was weighted to remove effects to array geometry. There was however a notable change in direction with time of year (Figure 7). The source to the east/northeast dominated from mid-September through to March, but in April and continuing through May the dominant source shifted to the south, coinciding with the switch to austral winter and perhaps indicating a switch to a remote infragravity wave source.

The change in source direction to the south was accompanied by a reduction in power at infragravity wave frequencies (Figure 8a). Power showed a positive correlation with local short-wave height \( H_s \) and \( aH_sT^2 \); \( g \frac{q}{Dq} \) (Figures 8b–8e), indicating the importance of leaky infragravity waves generated at the local coastline to infragravity power offshore in this part of the ocean. Scatter in the relationship could be partly due to remote arrivals. For example, power on 17 January 2013, 18 January 2013, and 11 May 2013 (marked by the black crosses in Figures 8d and 8e) was higher than expected for the local \( H_s \) or \( aH_sT^2 \); \( g \frac{q}{Dq} \), probably due to the infragravity arrival observed from the west.

3.2. Reflection of Infragravity Waves

The infragravity seaward:shoreward ratio was calculated for each daily cross correlation in the record. Throughout most of the record, the infragravity seaward:shoreward ratio was \( >1 \) (i.e., offshore propagation). A ratio \( >1 \) cannot be explained by incident arrivals and their reflections, but again indicates that for most days of the record the deep water infragravity waves were generated at the local coastline to the east. As with infragravity power (Figures 8d and 8e) the seaward:shoreward ratio increased as significant wave height \( H_s \) and \( aH_sT^2 \); \( g \frac{q}{Dq} \) increased at the coast (Figures 9a and 9b). Again, scatter in the relationship

Figure 6. (a) A simple impulse \( I \) propagates shoreward toward the coastline at normal incidence, passes through Station 2 and Station 1, gets reflected, and propagates back seaward \( R \). \( t_1 \) is the travel time between Station 1 and Station 2, \( t_2 \) is the travel time between Station 1 and the coast. (b) The time series observed at Station 2. (c) The time series observed at Station 1. (d) Cross correlation of records at Station 1 and Station 2.
may be caused by the shoreward propagating infragravity waves (which have no reason to correlate with
$H_s$ or $aH_sT - \sqrt{Dq}$ and seaward propagating waves. With the data binned according to $H_s$ or $aH_sT - \sqrt{Dq}$ the correlation with seaward-shoreward energy became much clearer (Figures 9a and 9b, red circles). At $H_s$ of about 4 m and below, or $aH_sT - \sqrt{Dq}$ of about 0.05 m and below, the relationship between $H_s$ and $aH_sT - \sqrt{Dq}$ and the seaward-shoreward ratio was approximately linear. For $H_s$ a linear regression calculated from the first four binned data points (which all had standard error of $\pm 0.14$) had a slope of $0.69 \pm 0.04$ m$^{-1}$ and intercept of $0.21 \pm 0.10$. For $aH_sT - \sqrt{Dq}$ the first five binned data points all had standard error of $\pm 0.04$ and had a slope of $33.27 \pm 1.48$ m$^{-1}$ and intercept of $0.55 \pm 0.04$. The square root of the intercept is an estimate of the reflection coefficient if the seaward-shoreward ratio is linear, therefore giving an estimate of $r = 0.49 \pm 0.29$ for $H_s$ or $r = 0.74 \pm 0.06$ for $aH_sT - \sqrt{Dq}$. At higher $H_s$ or $aH_sT - \sqrt{Dq}$ the data suggested a potential leveling off of seaward-propagating (leaky) infragravity waves, but the number of data points at these higher values was limited.

Figure 10 shows one of the examples when $H_s$ and $aH_sT - \sqrt{Dq}$ were low (Table 1), and a strong infragravity arrival from the west was observed. The reflection coefficient was calculated from each of the seven east-west-aligned station pairs on this day, at periods of 100, 125, 150, 175, and 200 s. Some examples of cross correlations on this day, from which the reflection coefficient was calculated, are shown in Figure 11. The reflection coefficient calculated varied between the different station pairs. Each cross in Figure 12a shows the reflection coefficient calculated for each station pair at each period. The mean varied between $0.83$ at 100 s and $0.53$ at 200 s, with a mean over all periods and station pairs of 0.66. A visual inspection of the cross-correlation functions found that for some station pairs the two peaks were not evident at all, and so a calculation based on the values at these two peaks was essentially meaningless (see supporting information Figure S5 for some examples of unclear correlations). For this reason, any unclear cross correlations were discarded, and the results were replotted (supporting information Figure S6a). In this case, the mean varied between $0.90$ at 100 s to $0.59$ at 200 s, with a mean over all periods of 0.71 (the mean at each period is plotted in Figure 13). It is understood that this approach risks a bias toward larger reflection coefficients, although an inspection of the differences found that overall both anomalous low and high values were removed, and the remaining data points became less scattered.

The reflection coefficient was calculated in the same way for the other four dates (25 October 2012, 17 January 2013, 4 April 2013, and 11 May 2013) when $H_s < 2$ m (and $aH_sT - \sqrt{Dq} < 0.03$ m).
westerly arrivals were observed (Figures 12b–12e and 13). The number of stations considered to have clear cross correlations on these other 4 days was generally less than on 18 January 2013 (numbers along bottom of figures in supporting information Figure S6). In particular, the arrivals on 25 October 2012 and 4 April 2013 were much less clear and not over all periods (see supporting information Figure S4 for backprojection plots). Therefore, we focus on the “best cases” of 17 January 2013, 18 January 2013, and 11 May 2013.

The mean reflection coefficient at each period from each of the 5 days (Figures 12a–12e) is plotted together in Figure 12f, and the means from supporting information Figures S6a–S6e are plotted together in Figure 13. The means were closely matched, especially for 18 January 2013 and 11 May 2013, in Figure 13. It is difficult to explain why the values obtained for the 17 January 2013 are the lowest obtained despite not having the lowest $H_s$ or $aH_sT^2m^0;22^2g Dq$. (Table 1). The mean reflection coefficient obtained over the three best cases
Figure 9. Same as Figures 8d and 8e but $H_s$ and $aH_{sT}^{2m_0;2}\sqrt{gDq}$ versus daily seaward:shoreward energy. Daily seaward:shoreward energy was calculated as the mean from the seven east-west station pairs over all periods 100–200 s. A linear regression was calculated through the first four (for $H_s$) or five (for $aH_{sT}^{2m_0;2}\sqrt{gDq}$) binned points and the equation of best fit, $R^2$ value, and $p$ value are shown.

Figure 10. (a) $H_s$ and (b) $aH_{sT}^{2m_0;2}\sqrt{gDq}$ at buoys 46015, 46022, and 46027 during January 2013. 18 January is bounded by the dashed lines. Backprojection on 18 January at (c) 100 s, (d) 150 s, and (e) 200 s shows the westerly infragravity wave arrival.
of 17 January 2013, 18 January 2013, and 11 May 2013 is given in Table 2, for both the calculations where all seven station pairs’ cross correlations were used (Figure 12) and for the calculation using only selected clear cross correlations (Figure 13). The means were $0.65 \pm 0.02$ and $0.66 \pm 0.02$, respectively. These estimates lie within the estimates of $r$ from the linear regressions of seaward:shoreward energy with $H_s$ and $aH_sT^2$; $2^{1/2}$, $D_fq$, which are also shown in Table 2.

4. Discussion

The dominant infragravity wave sources found here are in agreement with previous studies. The northwest coast of America has consistently showed up as a source of infragravity waves [Webb et al., 1991; Rhie and Romanowicz, 2006; Aucan and Ardhuin, 2013; Rawat et al., 2014]. In general, the eastern boundaries of basins provide stronger sources of infragravity waves than western boundaries due to larger wave heights and wave periods incident on these coasts [Rawat et al., 2014]. However, a recent study by Crawford et al. [2015] on deep ocean infragravity waves in the Atlantic has shown that other factors such as short-wave incidence angle or spread and coastal morphology may be more important for infragravity wave generation than wave heights and periods. This study has shown some rare examples where infragravity waves arrived from the western side of the Pacific.
The results suggest that infragravity waves that propagated west to east across the Pacific Ocean reflected at the North American Pacific coastline strongly. Perhaps the main limitation of the method used is that it is problematic to distinguish between reflected infragravity waves and locally generated infragravity waves. Here the reflection coefficient was estimated by assuming that there were no locally generated infragravity waves on the days investigated because nearshore significant wave height \( H_s \) or the parameter \( aH_sT^2 \) was low. An alternative method which extrapolated the observed trend of the seaward:shoreward ratio with \( H_s \) and \( aH_sT^2 \) to zero complemented the results obtained on the individual days investigated.

Figure 12. Values of \( r \) (seaward:shoreward) obtained from east-west station pairs on (a) 18 January 2013, (b) 25 October 2012, (c) 4 April 2013, (d) 17 January 2013, and (e) 11 May 2013. The red points mark the mean value of \( r \) from the seven east-west station pairs at each period, with error bars of two standard deviations (1 SD positive and 1 SD negative). The number of station pairs used in the calculation of the mean is given along the bottom of the plot. N/A means no estimate of \( r \) was made because backprojection showed no clear arrival from the west for this period. The mean value of \( r \) over all periods is given by the number in the top right corner of each plot.

(f) The mean values at each period from Figures 12a–12e. The axes for all plots are equal to those given for Figure 12a.

The results suggest that infragravity waves that propagated west to east across the Pacific Ocean reflected at the North American Pacific coastline strongly. Perhaps the main limitation of the method used is that it is problematic to distinguish between reflected infragravity waves and locally generated infragravity waves. Here the reflection coefficient was estimated by assuming that there were no locally generated infragravity waves on the days investigated because nearshore significant wave height \( H_s \) or the parameter \( aH_sT^2 \) was low. An alternative method which extrapolated the observed trend of the seaward:shoreward ratio with \( H_s \) and \( aH_sT^2 \) to zero complemented the results obtained on the individual days investigated.
One approach to tackling the problem of locally generated infragravity waves might be to use calibrated pressure gauges (only uncalibrated gauges were available in this study). Then, the expected seaward propagating infragravity wave energy for a particular sea state could be estimated from an empirical model between sea state and infragravity wave height, for example such as that described in Ardhuin et al. [2014], and then removed from the observed seaward propagating energy. This would have other problems associated with it such as the accuracy of the model for particular coastlines and sea states. However, it might offer some more insight as to the magnitude of the reflection and allow estimates of reflection to be made on days when wave activity at the coast is larger, so local generation cannot be assumed to be zero. Another approach could be to use modeled infragravity energy for the whole ocean basin [Ardhuin et al., 2014] and see if there is a mismatch in modeled and observed energy on days when reflection is expected to be large.

Another way to distinguish reflection from local generation, which should be possible with the method used here, is by identifying a peak in energy that arrives at a greater lag than the main arrival (\(t_1\) + \(2t_2\) peak labeled IR in Figure 6d). This peak has not been clearly identified in this data set. This might be because the schematic in Figure 6 is too simple for the real complicated wavefield. The complexity of the wavefield can be seen by the early arrivals and multiple late arrivals in the cross correlations (Figures 3 and 11), indicating waves propagating obliquely to normal incidence and multiple reflections, respectively. The IR peak would not be clear or symmetrical if reflection occurred from a source unaligned with the station pair, from multiple sources, or occurred gradually over a large geographical area (such as the continental shelf). The beam-forming output does show a late arrival in the direction that would be expected from a reflection, but the isotropic responses in Figure 5 suggest that this is due to aliasing by the array geometry. However, a good example of this reflection peak can be seen in the data of Harmon et al. [2012] for pressure gauges southwest of Sumatra, which were dominated by remote arrivals rather than local generation. The stacked cross correlation between their stations 42 and 45 shows symmetrical secondary peaks at \(2000\) and \(1200\) s lags (supporting information Figure S8a). A quick calculation of the reflection coefficient from the ratio of the IR peak to the \(2^2\) peak gives \(r\) of approximately 0.3, which is much lower than that found here, but is biased only on arrivals at normal incidence to the coast. This result may indicate that reflection coefficients depend strongly on the bathymetry and configuration of coastlines. Indeed, if all coastlines reflected at approximately 0.7, then infragravity energy in the middle of the ocean would be substantially higher than a case with no reflection. However, Webb et al. [1991] found little infragravity waves originating from large parts of the Pacific, which led them to suggest that reflection from coastlines is small. This perhaps suggests that the North American Pacific coastline is an example of a particularly strong reflector for infragravity waves reflecting at normal incidence to the coast (as indeed it is also a strong source [Ardhuin et al., 2014]).

The lag time of this late arrival (IR peak at \(t_1\) + \(2t_2\)) can also offer information on where the reflection occurs—whether at the continental slope or at the shoreline, or a combination of both. A backprojection of

Table 1. Nearshore Short-Wave Conditions for Days of Westerly Infragravity Arrivals*

<table>
<thead>
<tr>
<th>Date</th>
<th>Hs (m)</th>
<th>Average Wave Period (T_m) (s)</th>
<th>(H_sT_m^{1/2}) (m)</th>
<th>Dominant Wave Period (s)</th>
<th>Dominant Wave Direction (°)</th>
</tr>
</thead>
<tbody>
<tr>
<td>4 Apr 2013</td>
<td>1.83</td>
<td>6.11</td>
<td>0.019</td>
<td>8.37</td>
<td>238</td>
</tr>
<tr>
<td>25 Oct 2012</td>
<td>1.63</td>
<td>6.55</td>
<td>0.021</td>
<td>8.84</td>
<td>301</td>
</tr>
<tr>
<td>17 Jan 2013</td>
<td>1.36</td>
<td>6.50</td>
<td>0.018</td>
<td>12.90</td>
<td>286</td>
</tr>
<tr>
<td>18 Jan 2013</td>
<td>1.06</td>
<td>8.82</td>
<td>0.026</td>
<td>12.90</td>
<td>288</td>
</tr>
<tr>
<td>11 May 2013</td>
<td>0.58</td>
<td>6.88</td>
<td>0.009</td>
<td>14.91</td>
<td>213</td>
</tr>
</tbody>
</table>

*Average of NDBC wave buoys 46015, 46016, and 46027. Spectra are shown in supporting information Figure S7.
the Sumatra lag onto a $t = 2t_g$ travel time grid resulting from the relationship between infragravity waves and short-period wave activity. Infragravity wave energy in the deep ocean (i.e., leaky free infragravity waves) increases with short-period wave activity at the coast (indicated by $H_\text{s}$ and $H_\text{m}$), resulting in higher infragravity wave energy during winter when short-period wave activity is highest.

Remote arrivals coming from the west, propagating eastward, were observed but rare. The strength of reflection of these remote arrivals was estimated using the asymmetry of cross-correlation functions calculated between station pairs perpendicular to the coastline. While the method is limited to the assumption of no local infragravity wave generation when short-period wave activity is low, reflection did appear to be strong, with a lower bound estimate of $r = 0.49 \pm 0.29$ (reflection coefficient $\pm$ standard error) and an upper bound estimate of $r = 0.74 \pm 0.06$ for this particular coastline. These results indicate that reflection has the potential to be an important factor to account for infragravity wave energy in the deep ocean.

5. Conclusions

The North American Pacific coastline was found to be the dominant source of infragravity waves observed offshore, although energy and source were found to change seasonally. During northern-hemisphere winter, infragravity wave energy was higher and the waves mostly originated from the nearby coastline to the northeast, while during northern-hemisphere summer, energy decreased and arrivals mostly came from the south. The seasonal pattern can be explained by the relationship between infragravity waves and short-period (2–30 s) wave activity. Infragravity wave energy in the deep ocean (i.e., leaky free infragravity waves) increases with short-period wave activity at the coast (indicated by $H_\text{s}$ and $H_\text{m}$), resulting in higher infragravity wave energy during winter when short-period wave activity is highest.

Remote arrivals coming from the west, propagating eastward, were observed but rare. The strength of reflection of these remote arrivals was estimated using the asymmetry of cross-correlation functions calculated between station pairs perpendicular to the coastline. While the method is limited to the assumption of no local infragravity wave generation when short-period wave activity is low, reflection did appear to be strong, with a lower bound estimate of $r = 0.49 \pm 0.29$ (reflection coefficient $\pm$ standard error) and an upper bound estimate of $r = 0.74 \pm 0.06$ for this particular coastline. These results indicate that reflection has the potential to be an important factor to account for infragravity wave energy in the deep ocean.

References

Supporting Information for Source regions and reflection of infragravity waves offshore of the USA’s Pacific Northwest

Jennifer Neale 1, Nicholas Harmon 1 and Meric Srokosz 2

1 Ocean and Earth Science, University of Southampton, U.K.
2 National Oceanography Centre, Southampton, U.K.

Contents of this file

Figures S1 to S8
Table S1

Introduction

This supporting information contains additional figures referred to in the main text, and a table of station names and locations.
Figure S1. Station map with lines connecting station pairs that have separation distances greater than 30km and less than 120km.
**Figure S2** Weighted backprojection of September 2012-May 2013 stack. First, second and third row for 100s, 150s and 200s. Column 1: Backprojection of stack. Column 2: Response of array to isotropic wavefield. Column 3: Weighted stack.
**Figure S3** Beamforming of September 2012-May 2013 stack. First, second and third row for 100s, 146s and 200s. Column 1: Beamforming of stack. Radial axis is slowness in s/km, angular axis is azimuth of wave arrival. Column 2: Response of array to isotropic wavefield. Column 3: Beamform output at slowness of 12.3704 s/km (for 100s), 8.22 s/km (for 146s), 7.7037 s/km (for 200s) projected onto map of Pacific.
Figure S4 Backprojection for the days with westerly arrivals. (a) 18 January 2013 (b) 25 October 2012 (c) 4 April 2013 (d) 17 January 2013 (e) 11 May 2013. Columns left to right are periods of 100s, 125s, 150s, 175s, 200s.
Figure S5 Examples of unclear cross-correlations.
Figure S6 Same as Figure 12 but only station pairs which showed clear peaks at $\pm t_1$ are plotted and used in the calculations of the mean.
Figure S7 Wave spectra from buoy 46015 on a) 25 October 2012 b) 4 April 2013 c) 17 January 2013 d) 18 January 2013 e) 11 May 2013. The shaded regions show the range of the hourly spectra over the day.
Figure S8 a) Stacked cross-correlation of stations 42-45 from Harmon et al. [2012] band-pass-filtered between 60 and 500s. b) Projection of the cross-correlation onto $t_1 + 2t_2$ traveltime grid (IR peak). For the projection, the positive and negative lags were stacked. c) projection of (b) zoomed in to show relation of projection to bathymetry. The red line is on an azimuth that connects Station 1 and Station 2.
<table>
<thead>
<tr>
<th>Stations Used</th>
<th>Latitude</th>
<th>Longitude</th>
<th>Depth (m)</th>
</tr>
</thead>
<tbody>
<tr>
<td>FS05B</td>
<td>40.3866</td>
<td>-124.8996</td>
<td>2316</td>
</tr>
<tr>
<td>FS06B</td>
<td>40.3812</td>
<td>-124.7855</td>
<td>2198</td>
</tr>
<tr>
<td>FS09B</td>
<td>40.4387</td>
<td>-124.8085</td>
<td>2161</td>
</tr>
<tr>
<td>G02B</td>
<td>40.0486</td>
<td>-125.2869</td>
<td>1920</td>
</tr>
<tr>
<td>G04B</td>
<td>40.0585</td>
<td>-126.9285</td>
<td>4366</td>
</tr>
<tr>
<td>G05B</td>
<td>40.0706</td>
<td>-127.7478</td>
<td>4462</td>
</tr>
<tr>
<td>G10B</td>
<td>40.6779</td>
<td>-125.5533</td>
<td>2936</td>
</tr>
<tr>
<td>G11B</td>
<td>40.6875</td>
<td>-126.3764</td>
<td>3123</td>
</tr>
<tr>
<td>G13B</td>
<td>40.684</td>
<td>-128.0283</td>
<td>3215</td>
</tr>
<tr>
<td>G20B</td>
<td>41.2987</td>
<td>-126.6133</td>
<td>3141</td>
</tr>
<tr>
<td>G21B</td>
<td>41.316</td>
<td>-127.4537</td>
<td>3156</td>
</tr>
<tr>
<td>G22B</td>
<td>41.3092</td>
<td>-128.274</td>
<td>3038</td>
</tr>
<tr>
<td>G28B</td>
<td>41.9428</td>
<td>-126.7339</td>
<td>3327</td>
</tr>
<tr>
<td>G29B</td>
<td>41.9772</td>
<td>-127.4834</td>
<td>3197</td>
</tr>
<tr>
<td>G30B</td>
<td>41.9565</td>
<td>-128.3198</td>
<td>3135</td>
</tr>
<tr>
<td>G35B</td>
<td>42.5658</td>
<td>-126.0558</td>
<td>2367</td>
</tr>
<tr>
<td>G36B2</td>
<td>42.5994</td>
<td>-126.9038</td>
<td>2423</td>
</tr>
<tr>
<td>G37B</td>
<td>42.5913</td>
<td>-127.7213</td>
<td>3004</td>
</tr>
<tr>
<td>J06B</td>
<td>43.2652</td>
<td>-128.7997</td>
<td>3220</td>
</tr>
<tr>
<td>J09B</td>
<td>40.2011</td>
<td>-124.727</td>
<td>252</td>
</tr>
<tr>
<td>J11B</td>
<td>43.5414</td>
<td>-126.3673</td>
<td>3030</td>
</tr>
<tr>
<td>M11B</td>
<td>42.932</td>
<td>-125.0171</td>
<td>3109</td>
</tr>
<tr>
<td>M12B</td>
<td>42.184</td>
<td>-124.9461</td>
<td>1045</td>
</tr>
<tr>
<td>M14B</td>
<td>40.985</td>
<td>-124.5897</td>
<td>638</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Stations not used due to distance from other stations</th>
<th>Latitude</th>
<th>Longitude</th>
<th>Depth (m)</th>
</tr>
</thead>
<tbody>
<tr>
<td>J19B</td>
<td>44.1777</td>
<td>-126.2717</td>
<td>2968</td>
</tr>
<tr>
<td>J23B</td>
<td>44.8459</td>
<td>-129.6821</td>
<td>2650</td>
</tr>
<tr>
<td>J27B</td>
<td>44.8476</td>
<td>-126.3069</td>
<td>2824</td>
</tr>
<tr>
<td>J28B</td>
<td>45.0631</td>
<td>-127.1552</td>
<td>2866</td>
</tr>
<tr>
<td>J48B</td>
<td>47.1292</td>
<td>-130.6385</td>
<td>2877</td>
</tr>
<tr>
<td>J66B</td>
<td>48.2054</td>
<td>-130.0019</td>
<td>2857</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Stations not used due to data quality</th>
<th>Latitude</th>
<th>Longitude</th>
<th>Depth (m)</th>
</tr>
</thead>
<tbody>
<tr>
<td>FS01B</td>
<td>40.3268</td>
<td>-124.9492</td>
<td>940</td>
</tr>
<tr>
<td>FS14B</td>
<td>40.4955</td>
<td>-124.5918</td>
<td>107</td>
</tr>
<tr>
<td>G03B</td>
<td>40.058</td>
<td>-126.1634</td>
<td>4051</td>
</tr>
<tr>
<td>G12B</td>
<td>40.6869</td>
<td>-127.2289</td>
<td>3087</td>
</tr>
<tr>
<td>G13B</td>
<td>41.3074</td>
<td>-125.7736</td>
<td>3071</td>
</tr>
<tr>
<td>G27B</td>
<td>41.946</td>
<td>-126.0167</td>
<td>3480</td>
</tr>
<tr>
<td>G37B</td>
<td>42.5913</td>
<td>-127.7213</td>
<td>3004</td>
</tr>
<tr>
<td>J10B</td>
<td>43.3494</td>
<td>-125.5435</td>
<td>3093</td>
</tr>
<tr>
<td>J18B</td>
<td>44.0083</td>
<td>-125.466</td>
<td>3047</td>
</tr>
<tr>
<td>J20B</td>
<td>44.3542</td>
<td>-127.0952</td>
<td>2954</td>
</tr>
</tbody>
</table>
Table S1. Stations of the Cascadia Array used in the study.

References
Appendix C

Section C.1 of this Appendix contains additional material for Chapter 3 which were given as Supporting Information in the published paper. Section C.2 contains the formatted paper published in the Journal of Geophysical Research: Oceans in January 2017.
C.1 Additional figures for Chapter 3

Figure C.1: Beamforming output between September 2012 and September 2014. a) Maximum beampower (dB) at each time step. b) Azimuth (clockwise from north in degrees) of the maximum beampower. c) Slowness (s/km) of the maximum beampower.
Figure C.2: Construction of estimated source spectrum for 15 November 2012 00:00-03:00. a) Modelled source spectrum. b) Beamformer spectrum. c) Average shape of modelled source spectrum (Figure 3.8c in Chapter 3). d) Estimated source spectrum.
C.2 Published paper and Supporting Information of Chapter 3

Monitoring remote ocean waves using P-wave microseisms
Jennifer Neale 1, Nicholas Harmon 1, and Meric Srokosz 2

1Ocean and Earth Science, National Oceanography Centre Southampton, University of Southampton, Southampton, UK, 2National Oceanography Centre, University of Southampton Waterfront Campus, Southampton, UK.

Abstract
Oceanic microseisms are generated by the interaction of opposing ocean waves and subsequent coupling with the seabed, so microseisms should contain information on the ocean conditions that generated them. This leads to the possibility of using seismic records as a proxy for the ocean gravity wavefield. Here we investigate the P-wave component of microseisms, which has previously been linked to areas of high wave interaction intensity in mid-ocean regions. We compare modeled P-wave microseismic sources with those observed at an array in California, and also investigate the relationship between observed sources and significant wave height. We found that the time-varying location of microseism sources in the North Pacific, mapped from beamforming and backprojection of seismic data, was accurate to ≤10° in 90% of cases. The modeled sources were found to dominate at ~0.2 Hz which was also reflected in the seismic observations. An empirical relationship between observed beampower and modeled source power allowed source locations during an independent data period to be estimated with a correlation coefficient of 0.63. Likewise, significant wave height was also estimated with a correlation coefficient of 0.63. Our findings suggest that with improvements in resolution and amplitude retrieval from beamforming, correlations up to 0.78 should be possible between observed P-wave microseisms and significant wave height in remote ocean regions.

1. Introduction
Oceanic microseisms are tiny, continuous oscillations of the ground caused by the interaction of ocean waves with the solid earth beneath them. The most energetic microseisms are generated when ocean wave trains of similar frequency but opposite direction interact, producing a pressure fluctuation at seismic frequencies, which propagate as seismic surface waves and body waves. This pressure fluctuation is unattenuated with depth, has twice the frequency of the forcing waves, and a near zero wave number (Longuet-Higgins, 1956; Ardhuin and Herbers, 2013; Traer and Gerstoft, 2014, Ardhuin et al., 2015). This pressure fluctuation couples to the seabed to produce microseisms with typical peak frequencies of about 0.1–0.2 Hz (5–20 sec periods), which propagate as seismic surface waves and body waves; e.g., Toksöz and Lacoss, 1968, Haubrich and McCurry, 1969, Roux et al., 2005, Gerstoft et al., 2006, Koper and de Voz, 2008, Koper et al., 2010, Reading et al., 2014.

The requirement of opposing waves to generate double frequency microseisms means that microseism energy cannot be explicitly related to ocean wave height (Kedar et al., 2008, Ardhuin et al., 2011), as it depends on the directional characteristics of the wave trains as well as the wave energy. However, in cases where the opposing wave trains are related, as a result of directional spreading of one wave system (classified as Class I in Ardhuin et al. (2011, 2012)) or reflection from a coastline (Class II), a strong dependence on the ocean wave height is expected. As such, empirical relationships between microseism energy and significant wave height recorded at nearby coastal wave buoys have been identified (Ardhuin et al., 2012, Bromirski et al., 1999, Ferretti et al., 2013). These relationships were based on the full seismic spectrum which is dominated by seismic surface waves, and work well for seismic stations that are most sensitive to local sea states or coastal reflection sources.

This study investigates the P-wave component of seismic noise, for which origins in the deep ocean have been inferred (Toksöz and Lacoss, 1968, Haubrich and McCurry, 1969, Longuet-Higgins, 1975; Koper et al., 2008; Zhang et al., 2010a, Euler et al., 2014, Koper et al., 2010). Here Class I situations do occur, but the opposing wavefield could also result from the interaction of two independent wave systems: either the crossing of two swells, or the interaction of swell with locally generated wind waves (Class III in Ardhuin et al. (2011, 2012)). Given that two low-energy wave systems directly opposing could produce the same microseismic
The beamformer output is then of stations with given azimuth and white triangle with black border. The array has been used previously to successfully locate P-wave microseisms from October 2012 to September 2014. The Southern California Seismic Network (SCSN) that makes up the bulk of the array is given by

\[
\mathbf{p}_n(f, \mathbf{s}, \mathbf{h}; \mathbf{r}) = \exp\left(2\pi i f \mathbf{r} \cdot \mathbf{s}ight)
\]

where the brackets indicate temporal averaging over a 3 h period, was calculated. The plane-wave response of the array is given by

\[
\mathbf{C}_{138} = \mathbf{v} \mathbf{v}^T
\]

which we adjusted for the number of stations N in the array by dividing by N². Array response functions at frequencies of 0.1 and 0.3 Hz are shown in Figures 2a and 2b.

We first compare observed P-wave microseisms with modeled microseism sources in terms of location, frequency, and amplitude, and derive an empirical relationship between the observed and modeled source. We then reconstruct the ocean wave energy spectrum from the estimated microseism source spectrum by applying assumptions about the degree of wavefield opposition. Finally, the significant wave height calculated from the estimated ocean wave energy spectrum is then compared with modeled significant wave height.

In this study, we investigate the relationship between P-wave microseisms and significant wave height at short (3 h) time periods. Significant wave height is a common and useful parameter to describe sea state, but measurements by wave-buoys far from the coast are very sparse, and coverage by satellites insufficient. Locating P-wave microseisms to these deep ocean regions, and being able to infer wave heights, would allow for the development of real-time monitoring of sea state from seismic records, adding a valuable new record to our current measurements of significant wave height in the deep ocean.

2. Data and Processing

Two years of vertical component seismic data (channel LHZ) from an array in California (Figure 1) were downloaded from the Southern California Earthquake Data Center (SCEDC, 2013) covering the period September 2012 to September 2014. The Southern California Seismic Network (SCSN) that makes up the bulk of the array has been used previously to successfully locate P-wave microseisms (Gerstoft et al., 2006, 2008; Zhang et al., 2010a, 2010b; Obrebski et al., 2013). Instrument response was removed, and each daily record was band-pass filtered between 0.002 and 0.400 Hz. Earthquakes were removed using an automated detection and removal algorithm. Earthquake events were identified using the ISC bulletin (International Seismological Centre, 2014) and removed by setting 1 h of the waveform to zero if the RMS of that window was over 3 times the daily RMS. Daily spectra for each station were then individually examined and any bad quality data were discarded.

A similar beamforming procedure to Gerstoft and Tanimoto [2007] and Gerstoft et al. [2008] was used to examine microseisms as a function of frequency, azimuth, and slowness. Each daily time series was split into 512 s chunks which were Fourier transformed to give a complex valued vector \( \mathbf{v}(f, t) \) containing the response from all stations in the array, where \( f \) is frequency in Hz and \( t \) refers to the start time of the Fourier transform. Only phase was retained by dividing \( \mathbf{v} \) by its magnitude to reduce the influence of local site amplification effects (Gerstoft and Tanimoto, 2007). The cross-spectral density matrix, \( \mathbf{C}_{138} = \mathbf{v} \mathbf{v}^T \), where the brackets indicate temporal averaging over a 3 h period, was calculated. The plane-wave response of the array is given by

\[
\mathbf{C}_{138} = \mathbf{v} \mathbf{v}^T
\]

with given azimuth \( \theta \), \( r \) is the coordinates of the seismometers with respect to their mean, and \( s \) is slowness. The beamformer output is then

\[
\mathbf{p}_n(f, \mathbf{s}, \mathbf{h}; \mathbf{r}) = \exp\left(2\pi i f \mathbf{r} \cdot \mathbf{s}\right)
\]

which we adjusted for the number of stations N in the array by dividing by N² (Euler et al., 2014). Array response functions at frequencies of 0.1 and 0.3 Hz are shown in Figures 2a and 2b.
For each 3 h time period, we therefore have beamformer output as a function of frequency, slowness, and azimuth. An examination of the full 2 years of data revealed a large contrast between the strongest arrivals during winter and summer (supporting information Figure S1). During winter months, the strongest arrivals came from the North Pacific and North Atlantic with slownesses corresponding to P or PP phases, whereas during summer months the strongest arrivals came from southerly azimuths with slownesses of PKP phases.

Gerstoft et al. [2008] previously found P-waves from the Pacific to dominate at the SCSN network during winter, so here we limit our analysis to P-waves from the North Pacific from mid-October to mid-March. Additionally we reject any days when the number of stations falls below 170 to minimize errors and biases associated with too few stations. At each frequency of interest, we back-projected azimuth and slowness onto a 2° latitude by 2° longitude geographical grid assuming slownesses of direct P-waves with a source at the Earth’s surface using the ak135 travel time tables of Kennett et al. [1995] for a spherically symmetric Earth model. We projected for distances between 30° and 90° from the array center which is the typical range for teleseismic P-waves [Obrebski et al., 2013]. A synthetic test for a point source located at 35°N 169°E is shown in Figures 2c and 2d. An example output for 23 December 2012 (00:00) is shown in Figure 3a.

Double frequency P-wave microseism sources were modeled over the same 2 year period as the seismic data. The method of Ardhuin et al. [2011] and Ardhuin and Herbers [2013] based on the numerical ocean wave model WAVEWATCH III [Tolman et al., 2014] was followed to calculate the vertical ground displacement associated with P-waves at each source location. From Ardhuin et al. [2011] and Fane et al. [2016], the second-order pressure spectrum at near-zero wavenumber and twice the ocean wave frequency \( f \), due to the interaction of similar frequency waves traveling in opposite directions [Hasselmann, 1963], is given by:

\[
F_p(s_2=2f) = 2\pi \rho_w g \mathcal{E}(f)(f/f) 
\]

which has units of Pa m^-3 s and where \( \rho_w \) is the density of water, \( g \) is gravitational acceleration, \( \mathcal{E} \) is the ocean wave frequency spectrum and \( f(f/f) \) is a non-dimensional function that depends on the wave energy distribution \( \mathcal{E} \) over the directions \( \theta \).
WAVEWATCH III was forced with 6 hourly winds and sea-ice cover from ECMWF’s ERA-interim reanalysis [Dee et al., 2011] and the second-order pressure spectrum was output hourly on a 0.5° longitude-by-0.5° latitude global grid from 78° North to 78° South. The output was smoothed using a 7-by-7 Gaussian low-pass filter and resampled to a 2°-by-2° grid and averaged over 3 h time periods for direct comparison with the beamforming. Coastal wave reflection, which is function of wave amplitude, wave frequency and beach slope, can significantly affect the value of $F_P$ near the coast. Because we are mainly interested in deep-water events where reflection has less influence, we chose the reflection coefficient simply as $R = 0.1$ for continents and large islands and 0.2 for small islands [Ardhuin et al., 2011]. A second model run over a 2 month period with lower reflection coefficients of $R = 0.02$ for continents and large islands and 0.04 for small islands, which more closely correspond to the estimates of Stutzmann et al. [2012], did not produce significantly different values of $F_P$. The map of P-wave sources ($P_a^2 m^2 s$) is calculated by multiplying the second order pressure spectrum at each grid point by the squared source site effect $\frac{C}{C_{138}}$ [Farra et al., 2016]:

$$P(f) = P(f) \cdot \left( \frac{C}{C_{138}} \right)^2$$

(C)
We also output the modeled ocean wave energy spectrum $E(f)$ from which we calculated significant wave height, $H_s$, as observations of $H_s$ (e.g., from wave buoys or satellites) are not available at the temporal resolution or spatial extent required for comparison with the seismic data. Again we smoothed and resampled the WAVEWATCH III output of $E(f)$ to a 2’ by 2’ grid and averaged over 3 h time periods.

3. Results

3.1. Comparison Between Observed and Modeled Source

For each 3 h time period over the first winter (mid-October 2012 to mid-March 2013), the beamformer output and modeled source were integrated over the main double-frequency microseism band between 0.1 and 0.3 Hz. In order to identify multiple peaks in the beamformer maps, caused by more than one source acting in the basin at a given time, we used a standard watershed algorithm that identifies peaks in the image and the regions associated with each of these peaks (see Figure 3c). For each beamformer peak identified, we found the largest modeled source within its watershed region to examine how the beampower varies with source power. By associating the beampower peak to a source located anywhere within its watershed region, rather than using a direct pixel comparison, we allow for errors in the beamformer location which result in it being offset from the true source location. We considered beamform peaks in the North Pacific (above 0°N).

The plot of modeled source power versus beampower in Figure 4 shows that up to a certain limit, an increase in modeled source power had little to no effect on the observed beampower, but beyond this limit beampower increased rapidly with source power, although with a lot of scatter. The correlation coefficient calculated using Spearman’s Rank was 0.71 (strong correlation, >99% significance). After binning the data (shown in red), we defined the noise level as the mean of the first bin plus 2 standard deviations, which corresponds to a beampower of $-28.68$ dB. The binned data were found to be best described by an exponential function (shown in green), which indicated that the beampower only exceeded the noise level when source power $P_f(2C_20)$ reached approximately 102 dB ($1.6 \times 10^{10}$ Pa m$^2$).

Next we calculated the distance between the beamformer peaks and their associated modeled sources, excluding all peaks below the noise threshold of $-28.68$ dB. The locations of the beamformer peaks and the sources are plotted in Figures 5a and 5b with color indicating the number of occurrences at each grid point. It can be seen from Figures 5a and 5b that both the observed and modeled sources occurred over the whole region, although the most energetic 100 events were concentrated around 40°-50°N, 160°-180°E (Figures 5c and 5d), which corresponds well with the locations observed by Obrebski et al. (2013) in the year 2010. The beamformer peak and associated modeled source were found to be located ≤ 10’ (5 grid points) apart in 90% of occurrences, were ≤ 16’ (3 grid points) apart in 67% of occurrences but only matched the same grid point in 4% of occurrences. Obrebski et al. (2013) found similar offsets (between 1.1’ and 9.5’) for the 54 strongest modeled events in their study. Discrepancies may be due to 3-D velocity structure that we are not taking into account in the backprojection (up to 4’) [Euler et al., 2014], contamination from other phases or earthquakes, contamination from array response [Gal et al., 2016], or inaccuracies in the modeled source location (either due to inaccuracies in the wave model or site effect). In some cases, the discrepancy was found to arise from the resolution limits of the array, where two neighboring sources merged into one observed source centered on the average location (a limitation that has been observed previously) [Hillers et al., 2012; Euler et al., 2014].

To compare the frequency content of the modeled and observed sources, we examined
the nonintegrated output of each of the beamformer peaks and associated source peaks. Again we only considered beamformer peaks that were above the noise threshold, and in addition only the cases when the observed and modeled sources were \( \frac{1}{10} \) apart to exclude any spurious observations.

The mean spectra of the observed and modeled sources were calculated (Figures 6a and 6b), and the mean frequency from these spectra was calculated as:

\[
\text{mean frequency} = \sqrt{m_n^2 + 2\frac{1}{m_n}}
\]

where \( m_n \) is either the beamform spectrum \( b \) or the modeled source spectrum \( P \). The mean frequency of the beamformer was 0.21 Hz, whilst the mean frequency of the modeled source was 0.19 Hz, indicating that the seismic observations represented well the frequency content of the sources acting in the ocean basin. To examine the variability among individual spectra, the peak frequency of each spectrum was picked and plotted as a histogram with color corresponding to number of occurrences in Figure 6c. Again it can be seen that both the beamform peak and modeled source peak occurred most often at about 0.2 Hz, with rare instances where the peak occurred down to 0.1 Hz or up to 0.3 Hz.

The results indicate that the seismic beamforming observations reflect the location, frequency content, and amplitude of microseismic sources, with location accurate to \( \frac{1}{10} \) in 90% of occurrences and with a strong correlation coefficient between beampower and source power of 0.71.

Next we used the relationship between beampower and modeled source power found during the first winter to estimate source power over the second winter (mid-October 2013 to mid-March 2014) using the observed beampowers over that period.

To estimate source power from observed beampower, we fitted another curve between beampower and source power this time with beampower as the predictor and source power as the response variable (Figure 7a), and only considering beampowers over the noise threshold. Again we also excluded from the regression any cases when the observed and modeled sources were \( > \frac{1}{10} \) apart. The best-fitting logarithmic function was applied to the seismic observations of the second winter to estimate sources during the second winter if linear fit between beampower and source power for source powers \( > 10^2 \text{ dB} \) was attempted but...
resulted in overestimation of sources at large beampowers). The time series of the modeled source amplitude ($10^{C217}$ source power/10) is plotted along with estimated source amplitude in Figure 7b, where only the largest estimated source at each time step is plotted if there were more than one source acting simultaneously. Pearson's linear correlation coefficient between the estimated source power and modeled source power was 0.63 (>99% significance).

### 3.2. Relating Microseism Source to Significant Wave Height

The results presented have shown that we are able to estimate the integrated microseism sources in the North Pacific from seismic observations with some confidence. If we now consider observations at each frequency, we can reconstruct the full source spectrum, from which the ocean wave energy spectrum can be estimated, and thus significant wave height. This is a rearrangement of equation (3):

$$E(f) = \sqrt{\frac{P(f)}{2 \rho_w \rho_s C_p C_w}}$$  

where the source power spectral density spectrum $P(f)$ is estimated from the observed beampower spectrum, values of $\rho_w$, $\rho_s$, and $C_p$ are taken at the grid points where peaks in beampower are located, and $t(f)$ is unknown.

Significant wave height can then be calculated from the estimated ocean wave energy spectrum.
Here we distinguish our estimate of $H_s$ as $H_s(\text{swell})$ because we are only considering the spectrum between seismic frequencies of 0.1–0.3 Hz (0.05–0.15 Hz ocean wave frequency which equals 6.7–20 s periods) whereas traditionally significant wave height is calculated from a wider band (e.g., 2–30 ocean wave periods).

To estimate the source spectrum $P(f)$ from the observed beampower, beampower and source power were correlated in a similar way as previously, but instead of plotting the integrated powers, the power at each frequency was plotted separately (Figure 8a). From Figure 8a, it was found that values at the different frequencies all lay along the same curve, so we applied one relationship to all frequencies. Again we binned the data points according to source power and fitted an exponential function. The noise level of the beampower (as defined previously) was found to be $-20.16$ dB and the source exceeded the noise level at a value of $117$ dB ($5.0 \times 10^{11}$ Pa$^2$m$^{-2}$s).

Taking only beampower values above this noise threshold, we replotted using beampower as the predictor variable (Figure 8b) and found that source power at each frequency is best estimated from the observed beampower using a logarithmic function given by:

$$P(\text{estimated})_{(\text{dB})} = \text{aloge}(b \cdot (\text{beampower} - x_0)) + y_0$$

with values of $a = 10.3564$, $b = 0.89105$, $x_0 = -20.8629$ and $y_0 = 108.5138$.

The estimated source spectrum was constructed from the beampower spectrum using the identified logarithmic relationship in the following way. If the peak of the beampower spectrum was above the noise level of $-20.16$ dB, the source power of this peak, $P(\text{estimated})_{(\text{dB})}$, was estimated, and the source amplitude as $P(\text{estimated})_{(\text{dB})} = 10^{1/20}(P(\text{estimated})_{(\text{dB})}/10)$. It was found that by setting the estimated source amplitude at all other frequencies (below the noise threshold) to zero, the source spectrum and resulting ocean wave spectrum became too narrow. Instead, the mean shape of the modeled source spectrum as a function of $f/f_p$ (where $f_p$ is the peak frequency of the spectrum) was calculated (Figure 8c), and $P(\text{estimated})_{(\text{dB})}$ was multiplied by this. An example of the construction of one spectrum in this way is shown in supporting information Figure S2.
Appendix C

Figure 8. (a) Similar to Figure 4 but with modeled source power versus observed beampower plotted separately at each frequency. Only cases where the integrated value of the beampower was above the noise threshold of $-20$ dB are plotted, and only cases where the observed and modeled source were apart by $\pm 10^3$. The red circles are the mean of each data bin with error bars of $\pm 2$ standard deviations. The dashed line is placed at the top of the error bar of the first bin and is equal to $-20$ dB, which we define as the noise threshold. The green curve is the least-squares fit of the form $y=a(x-x_0)^b+y_0$, where values of $a$, $b$, $x_0$, and $y_0$ are given in the figure. (b) Same as Figure 8a but with the axes switched and only including beampowers $> -20$ dB. The green curve is the least-squares fit of the form $y=a(x-x_0)^b+y_0$. (c) Mean shape of observed source spectrum as a function of $f_p$, which is the sum frequency divided by the peak seismic frequency of the source spectrum. The data are plotted in black and the smoothed data in red (almost identical). (d) Mean spectrum of the wave-interaction intensity.

The wave energy spectrum was estimated from the beamforming observations of the first winter using equation (4), and significant wave height estimated using equation (5). Figure 9 shows the modeled $H_s$(model) versus the estimated $H_s$(well) which had a Pearson’s linear correlation coefficient of 0.48. The scatter points are shaded by point density and binned by modeled $H_s$(well). It can be seen from the figure that at the most common wave heights ($\sim 2-3$ m) the estimation was most accurate, whereas higher modeled wave
heights were underestimated. We used the mismatch between the estimated $H_s(\text{swell})$ bins and the line $y = x$ to define a calibration factor, which was then applied to estimates of $H_s(\text{swell})$ during the second winter (a calibration factor based on a curve which smoothed out the uncertainties at the larger wave heights did not make any significant difference to the results). Figure 10 shows modeled versus estimated $H_s(\text{swell})$ for the second winter which had a Pearson’s linear correlation coefficient of 0.63. With the calibration factor, the means of each data bin lie much closer to the line $y = x$, with underestimation only occurring at $H_s(\text{swell})$ values over about 8 m.

4. Discussion

Errors in the estimate of $E(f)$ using equation (4) and consequently $H_s$ will arise for three main reasons: (1) Inaccurate estimate of source amplitude $P(f^2)$ from beamformer output, (2) Inaccurate location of beamformer peak, therefore inaccurate site effect, (3) Inaccurate estimate of $I(f)$. Another expected source of error is that it is not possible to extract information about the ocean wave spectrum at frequencies for which the wavefield is unidirectional ($I(f) = 0$), because the observed beampower is only sensitive to opposing waves. Consequently, these parts of the ocean wave energy spectrum would be underestimated as would the derived significant wave height.

In order to understand how each of these errors influence the correlation between estimated and modeled $H_s(\text{swell})$, we looked at each of these sources of error in turn. Figure 11a shows the ideal case where the...
source spectrum, site effect, and $I(f)$ are known exactly for the estimation of $E(f)$ using equation (4) (i.e., the modeled values were used). The correlation coefficient in this case between modeled $H_s(\text{swell})$ and estimated $H_s(\text{swell})$ was equal to 1.00. Underestimation would be expected if there was a large portion of unidirectional wave energy in addition to the opposing wavefields, but this was not the case as the modeled $I(f)$ was rarely (<1% of the time) exactly 0. Plot b examines the error due to inaccurate estimate of $I(f)$. In this estimate of $H_s(\text{swell})$, the mean value of $I(f)$ shown in Figure 8d was used in the calculation. There was still a moderately strong correlation of 0.78 but this was much lower than the ideal case (a). There also appears to be a tendency for underestimation which increases with modeled $H_s(\text{swell})$. This would be caused by the assumed $I(f)$ being too large for higher wave heights. An explanation for this may be related to the case for waves generated by local winds, in which $I(f)$ generally decreases with increasing $E(f)$ [Ardhuin et al., 2012]. Plot c examines the error due to inaccurate location (site effect). In this estimate of $H_s(\text{swell})$, the location of the beamformer peak was used to obtain the value of the smoothed site effect. The correlation of 0.90 indicates that the error introduced from inaccurate location is less than the error introduced by inaccurate $I(f)$. Plot d examines the combined error from inaccurate $I(f)$ and location and as expected the correlation drops further, to 0.71.

Plots e-h are the same as plots a-d but with the addition of error caused by inaccurate estimate of the source amplitude. Plot e shows that inaccurate amplitude estimation causes a larger reduction in correlation than errors in $I(f)$ and location combined. With inaccuracies in $I(f)$ and location correlation goes down to 0.48 (Figure 11b).

The large amount of scatter between source amplitude and beampower is unsurprising given the amplitude removal during the beamforming.
process and, as mentioned by Obrebski et al. [2013], because the beampower also depends on the size of the area the source is acting over as well as energy losses along the propagation path (including geometric spreading, attenuation, and transmission through Earth structure such as the Moho, 410 and 660 km discontinuities) [Nishida and Takagi, 2016]. Nevertheless, our results show that a relationship between source amplitude and observed beampower does exist. Furthermore, Figure 8b suggests that with improvements in beampower and location estimation correlations of up to 0.78 are possible even with the uncertainty surrounding \( I(f) \). A direction for these improvements may be found in the recent work of Farra et al. [2016] and Nishida and Takagi [2016]. Farra et al. [2016] used a ray-theoretical approach to estimate P-wave ground displacement for a given source including site, receiver, and propagation effects, and Nishida and Takagi [2016] used a similar formulation to estimate the pressure source by minimising the squared difference between observed and modelled ground displacement.

Finally, throughout the study, the modelled source has been considered the "true" value. Scatter between the beampower and modelled source may be caused by inaccuracies in the wave model itself, for example due to the wind input or parameterization within the model. Although there is currently no other way of estimating wave-interaction intensity over such spatial and temporal scales with which to validate the model output, an idea of error within the WAVEmATCH III model could be found by analysing the spread of results obtained from multiple runs with different wind inputs and parameterization. Scatter may also be caused by inaccuracies in the calculation of the site effect, which may not well represent each 2-by-2° pixel in regions of large bathymetric variability [Hillers et al., 2012], and for which we have not taken into account the effect of sediments or earth structure below the upper crust [Gualtieri et al., 2014]. A thick sediment layer at the source results in reduced amplitudes of land-recorded microseisms [Gualtieri et al., 2015], and may be important for sources close to the coast where sediments are thicker.

It is important to remember however that estimates about significant wave height can only be made where there is wave interaction occurring. Sometimes this does correspond to the largest wave heights in the ocean basin (e.g., Figures 12a and 12b), but this is not necessarily the case (Figures 12c and 12d).

Figure 12. (Top row) An example when maximum source corresponds to location of maximum \( H_s \), and (bottom row) when maximum source does not correspond to location of maximum \( H_s \). (a) Modeled source on 22 December 2012 15:00:00. (b) Modeled \( H_s \) on 22 December 2012 15:00:00. The modeled source has been integrated over \( f = 0.51–3.0 \) Hz whilst \( H_s \) has been calculated from the modelled ocean wave spectra over the full ocean wave frequency range \( f = 0.039–0.5768 \) Hz.

5. Conclusions
Observed seismic P-waves in California were found to relate to modeled microseismic sources in the North Pacific Ocean in terms of location, frequency content, and amplitude. The observed P-waves were located through beamforming and backprojection, and were found to match the location of strong modeled sources by $<10^2$ (5 grid points) in 90% of cases. Both the modeled sources and observed P-waves were dominated by microseisms with a frequency of approximately 0.20 Hz. Beampower was moderately to strongly correlated with the power of the modeled sources, and only exceeded the noise threshold when sources were $>1.0 \times 10^8$ Pa m$^2$, integrated between 0.1 and 0.3 Hz or $>5.0 \times 10^7$ Pa m$^2$. The amplitude correlation between beampower and source power allowed sources during the second winter to be estimated from observed beampower. The resulting estimated sources were found to correlate with the modeled sources with a correlation coefficient of 0.63.

After reconstructing the source spectrum from the beamformer spectrum, and making an assumption about the directional characteristics of the wavefield, the ocean wave energy spectrum was estimated, and from that, significant wave height. During the first year, the modeled and estimated significant wave height correlated by 0.48. An underestimation of wave height at higher modeled values appeared to be introduced in the assumption of the directional wave characteristics. A calibration factor between modeled and estimated significant wave height calculated from the first winter’s results was applied to the second year and the underestimation was largely removed, giving a correlation of 0.63 between modeled and estimated significant wave height: inaccuracy in beampower was found to be the largest source of error, followed by inaccuracy in directional assumption. With improvements in the location and amplitude estimation of sources from beamforming, it should be possible to obtain estimates of significant wave height that correlate with modeled wave heights by up to 0.78.

References

Acknowledgments
All data sets used are freely available to the public and are listed in the references. Source code for WAVESPECTRUM can be requested from http://polar.ncep.noaa.gov/ users/markw/polarw/watch/. J. Neale’s PhD studentship was funded by the Natural Environment Research Council (NERC). M. Schimmel was supported by NERC National Capability funding.


Introduction
This supporting information provides additional figures referred to in the main article.
Figure S1. Beamforming output between September 2012 and September 2014. a) Maximum beampower (dB) at each timestep. b) Azimuth (clockwise from north in degrees) of the maximum beampower. c) Slowness (s/km) of the maximum beampower.
Figure S5a. Construction of estimated source spectrum for 15 Nov 2012 00:00:00. a) Modeled source spectrum. b) Beamformer spectrum. c) Average shape of modeled source spectrum (Figure 8c in main text). d) Estimated source spectrum.


