Appendix: Supplementary materials for Chapter 5

This appendix presents the results of binary classifications of *Azorella* distribution excluded from chapter 5.

Kappa statistics of the OOB errors for the binary classifications were affected by the spatial autocorrelation in the multiple pixel training method. The single pixel and object-based classifications performed similarly for all subsets of input variables. (Fig. S1)

The statistical hybrid and hypothesis-driven subsets of variables had the highest accuracies, performing significantly better than terrain or spectral classifications in isolation, with the spectral classification having the lowest accuracies for all training methods.

*Fig. S1*: Kappa statistics of the OOB accuracy measures for the two-class classifications of the presence and absence of *Azorella*. For the multiple pixel-trained classifications, the kappa statistic is much higher than the other training methods for all sets of inputs. OOB, as implemented in the *randomForest* package, is unsuitable for use with multiple pixel training methods common in remote sensing applications.
Terrain-based classifications

This section presents the confusion matrices (Table S1) and comparative maps (Fig. S2) for the terrain-based classifications for the three image training methods. All three terrain-based classifications produced solid patches of potential Azorella habitat across most of the plateau. The differences among the three maps are marginal, with the multiple pixel-trained map covering a slightly greater extent than the other maps.

Table S1: Confusion matrices for OOB accuracy estimates for single pixel-based (A); object-based (B); and multiple pixel-based (C) RF classifications based on terrain data only. These accuracy estimates are for classifications based on the smallest possible subset of input variables without reducing the accuracy of the classification.

(A) Predicted

<table>
<thead>
<tr>
<th></th>
<th>Absent</th>
<th>Present</th>
<th>Accuracy</th>
<th>Kappa</th>
</tr>
</thead>
<tbody>
<tr>
<td>Observed</td>
<td>123</td>
<td>13</td>
<td>90.4%</td>
<td></td>
</tr>
<tr>
<td></td>
<td>12</td>
<td>53</td>
<td>81.5%</td>
<td></td>
</tr>
<tr>
<td>Accuracy</td>
<td>91.1%</td>
<td>80.3%</td>
<td>87.6%</td>
<td>0.717</td>
</tr>
</tbody>
</table>

(B) Predicted

<table>
<thead>
<tr>
<th></th>
<th>Absent</th>
<th>Present</th>
<th>Accuracy</th>
<th>Kappa</th>
</tr>
</thead>
<tbody>
<tr>
<td>Observed</td>
<td>123</td>
<td>12</td>
<td>91.1%</td>
<td></td>
</tr>
<tr>
<td></td>
<td>15</td>
<td>50</td>
<td>76.9%</td>
<td></td>
</tr>
<tr>
<td>Accuracy</td>
<td>89.1%</td>
<td>80.6%</td>
<td>86.5%</td>
<td>0.689</td>
</tr>
</tbody>
</table>

(C) Predicted

<table>
<thead>
<tr>
<th></th>
<th>Absent</th>
<th>Present</th>
<th>Accuracy</th>
<th>Kappa</th>
</tr>
</thead>
<tbody>
<tr>
<td>Observed</td>
<td>3435</td>
<td>1</td>
<td>100.0%</td>
<td></td>
</tr>
<tr>
<td></td>
<td>0</td>
<td>1660</td>
<td>100.0%</td>
<td></td>
</tr>
<tr>
<td>Accuracy</td>
<td>100.0%</td>
<td>99.9%</td>
<td>100.0%</td>
<td>0.9996</td>
</tr>
</tbody>
</table>
Fig. S2: Predicted Azorella presence on northern Macquarie Island on the basis of terrain variables. The main panel shows the hard classes for Azorella presence based on single pixel, multiple pixel and object-based classifications. The predicted cover layers are partially transparent to demonstrate the overlaps in the predictions. Each of the image training methods resulted in solid blocks of predicted Azorella habitat on the plateau and mountain peaks, which is likely to show potential, rather than actual, Azorella distribution. The inset maps show the probability of Azorella presence based on single pixel (A); object (B); and multiple pixel -based (C) random forest classification of terrain.
Appendix B: Supplementary materials for Ch. 5

Spectral-based classifications

This section presents the confusion matrices (Table S2) and comparative maps (Fig. S3) for the spectral-based classifications for the three image training methods. The three spectral-based classifications produced much more fragmented patches of potential Azorella habitat than the terrain-based classifications. These maps all spuriously included beaches and rocky coastal headlands in the area of predicted *Azorella* presence. The differences among the three maps are subtle, with the multiple pixel-trained map covering a slightly greater extent than the other maps.

*Table S2: Confusion matrices for the pixel-based (A) and object-based (B) classifications of the field sites on the basis of spectral data only.*

<table>
<thead>
<tr>
<th></th>
<th>Predicted</th>
<th>Percentage</th>
<th>Kappa</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Observed</strong></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Absence</td>
<td>123</td>
<td>13</td>
<td>90.4%</td>
</tr>
<tr>
<td>Presence</td>
<td>16</td>
<td>49</td>
<td>75.4%</td>
</tr>
<tr>
<td>Percentage</td>
<td>88.5%</td>
<td>79.0%</td>
<td>85.6%</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th></th>
<th>Predicted</th>
<th>Percentage</th>
<th>Kappa</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Observed</strong></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Absence</td>
<td>122</td>
<td>13</td>
<td>90.4%</td>
</tr>
<tr>
<td>Presence</td>
<td>16</td>
<td>49</td>
<td>75.4%</td>
</tr>
<tr>
<td>Percentage</td>
<td>88.4%</td>
<td>79.0%</td>
<td>85.5%</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th></th>
<th>Predicted</th>
<th>Accuracy</th>
<th>Kappa</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Observed</strong></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Absent</td>
<td>3381</td>
<td>55</td>
<td>98.4%</td>
</tr>
<tr>
<td>Present</td>
<td>60</td>
<td>1600</td>
<td>96.4%</td>
</tr>
<tr>
<td>Accuracy</td>
<td>98.3%</td>
<td>96.7%</td>
<td>97.7%</td>
</tr>
</tbody>
</table>
Fig. S3: Predicted Azorella presence on northern Macquarie Island on the basis of spectral variables. The main panel shows the hard classes for Azorella presence based on single pixel, multiple pixel and object-based classifications. The predicted layers are partly transparent to demonstrate the overlaps among the predictions. The differences among the three image training methods are subtle. The spectral classifications resulted in much more fragmented predicted distributions than the terrain-based classifications, and predicted Azorella in coastal areas where the species is known not grow. The inset maps show the probability of Azorella presence based on single pixel (A); multiple pixel (B); and object-based (C) random forest classification of terrain.
Hybrid Classifications

This section presents the confusion matrices (Table S3) and comparative maps (Fig. S4) for the hybrid classifications for the three image training methods, with the input variables chosen on statistical grounds. The three hybrid classifications all performed as compromises between the hybrid and spectral classifications. The areas of predicted presence were all confined to the higher parts of the plateau, and were more fragmented than those in the terrain-based classification. The differences among the classifications based on the three training methods were subtle, with the single pixel-trained classification having the most restricted distribution, and the multiple pixel and object-based classifications extending around the fringes of the single pixel-trained presence class.

Table S3: Confusion matrices for the single pixel-based (A); object-based (B); and multiple pixel-based (C) RF classifications, using both terrain derivatives and spectral data.

<table>
<thead>
<tr>
<th></th>
<th>Predicted</th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Absent</td>
<td>Present</td>
<td>Accuracy</td>
</tr>
<tr>
<td><strong>Observed</strong></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Absent</td>
<td>131</td>
<td>5</td>
<td>96.3%</td>
</tr>
<tr>
<td>Present</td>
<td>7</td>
<td>58</td>
<td>89.2%</td>
</tr>
<tr>
<td><strong>Accuracy</strong></td>
<td>94.9%</td>
<td>92.1%</td>
<td>94.0%</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th></th>
<th>Predicted</th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Absent</td>
<td>Present</td>
<td>Accuracy</td>
</tr>
<tr>
<td><strong>Observed</strong></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Absent</td>
<td>129</td>
<td>6</td>
<td>95.6%</td>
</tr>
<tr>
<td>Present</td>
<td>7</td>
<td>58</td>
<td>89.2%</td>
</tr>
<tr>
<td><strong>Accuracy</strong></td>
<td>94.9%</td>
<td>90.6%</td>
<td>93.5%</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th></th>
<th>Predicted</th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Absent</td>
<td>Present</td>
<td>Accuracy</td>
</tr>
<tr>
<td><strong>Observed</strong></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Absent</td>
<td>3435</td>
<td>1</td>
<td>100.0</td>
</tr>
<tr>
<td>Present</td>
<td>1</td>
<td>1659</td>
<td>99.9</td>
</tr>
<tr>
<td><strong>Accuracy</strong></td>
<td>100.0</td>
<td>99.9</td>
<td>100.0</td>
</tr>
</tbody>
</table>
Fig S4: Predicted Azorella presence on northern Macquarie Island, from single pixel, object and multiple pixel based classifications of the hybrid variables. The predicted distribution layers are partly transparent, to demonstrate the overlaps among the predicted distributions. The hybrid RF classifications were more fragmented than the terrain-based classifications and lacked the spurious predictions of Azorella presence in coastal areas that the spectral-based classifications produced. The main map shows the hard classes for Azorella presence while the inset maps show the probability of Azorella presence based on single-pixel (A); multiple pixel (B); and object (C) trained classifications.
Hypothesis-driven Classifications

This section presents the confusion matrices (Table S4), partial dependence plots (Fig. S5) and comparative maps (Fig. S6) for the classifications based on a hypothesis-driven subset of the available input variables for the three image training methods. These three maps appeared very similar to those based on statistically-derived hybrid sets of variables, though the accuracies were slightly higher.

Classifications based on a hypothesis-driven subset of variables showed slightly improved accuracies compared to the hybrid models, when measured by the independent validation and slightly lower when measured by the OOB accuracy estimate. The single and multiple pixel classifications both had accuracies greater than 90% on both measures, while the object-based classification had lower accuracy, with 80.7% by independent validation, and 87.7% by OOB for the present class.

Table S4: Confusion matrices for the single pixel-based (A); object-based (B); and multiple pixel-based (C) RF classifications, using a hypothesis-drive subset of the terrain derivatives and spectral data.

<table>
<thead>
<tr>
<th></th>
<th>Predicted</th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Absent</td>
<td>Present</td>
<td>Accuracy</td>
</tr>
<tr>
<td><strong>Observed</strong></td>
<td><strong>Absent</strong></td>
<td><strong>Present</strong></td>
<td><strong>Accuracy</strong></td>
</tr>
<tr>
<td>Absent</td>
<td>129</td>
<td>7</td>
<td>94.9%</td>
</tr>
<tr>
<td>Present</td>
<td>9</td>
<td>56</td>
<td>86.2%</td>
</tr>
<tr>
<td>Accuracy</td>
<td>93.5%</td>
<td>88.9%</td>
<td>92.0%</td>
</tr>
</tbody>
</table>

**Table S4 continued:**

<table>
<thead>
<tr>
<th></th>
<th>Predicted</th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Absent</td>
<td>Present</td>
<td>Accuracy</td>
</tr>
<tr>
<td><strong>Observed</strong></td>
<td><strong>Absent</strong></td>
<td><strong>Present</strong></td>
<td><strong>Accuracy</strong></td>
</tr>
<tr>
<td>Absent</td>
<td>129</td>
<td>6</td>
<td>95.6%</td>
</tr>
<tr>
<td>Present</td>
<td>8</td>
<td>57</td>
<td>87.7%</td>
</tr>
<tr>
<td>Accuracy</td>
<td>94.2%</td>
<td>90.5%</td>
<td>93.0%</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th></th>
<th>Predicted</th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Absent</td>
<td>Present</td>
<td>Accuracy</td>
</tr>
<tr>
<td><strong>Observed</strong></td>
<td><strong>Absent</strong></td>
<td><strong>Present</strong></td>
<td><strong>Accuracy</strong></td>
</tr>
<tr>
<td>Absent</td>
<td>3432</td>
<td>4</td>
<td>99.9</td>
</tr>
<tr>
<td>Present</td>
<td>4</td>
<td>1656</td>
<td>99.8</td>
</tr>
<tr>
<td>Accuracy</td>
<td>99.9</td>
<td>99.8</td>
<td>99.8</td>
</tr>
</tbody>
</table>
The partial dependence plots showed that many of the input variables had non-linear relationships with the probability of *Azorella* presence. These plots showed that *Azorella* presence was linked to high values for elevation, solar radiation, slope, solar radiation, and blue and yellow reflectance; with lower values for red edge and NIR 2 reflectance, GLCM Mean; and with extreme values for green reflectance, GLCM homogeneity and entropy, curvature, valleyness and wetness index.

Fig. S5: Partial dependence plots for the present class for the multiple pixel-based classification of hypothesis-driven subset of input variables. *Azorella* presence was associated with high values for blue and yellow reflectance, elevation, ridgeness, slope, and solar radiation; with low levels of red edge and NIR2 reflectance, and the GLCM mean texture measure; and with mixed values for green reflectance, GLCM homogeneity and entropy, curvature, valleyness, and wetness index.
Fig. S6: Predicted Azorella presence on northern Macquarie Island, from single pixel, object and multiple pixel based classifications of the hypothesis-driven subset of variables. The main map shows the hard classes for Azorella presence, with the three predictive layers made partially transparent to demonstrate the overlaps among the predictions. Each of these predictions was similar to the classifications based on the statistically-chosen subsets of hybrid variables. The inset maps show the probability of Azorella presence based on single-pixel (A); multiple pixel (B); and object (C) trained classifications.
References


Beyer HL (2009) Geospatial Modelling Environment (0.6.2.0).


References


Chiarucci A (2007) To sample or not to sample? That is the question ... for the vegetation scientist. Folia Geobotanica 42:209-216


Department, of Primary Industries P and W (2009) Tasmanian Threatened Species Listing Statement: Azorella macquariensis (macquarie cushions). Tasmania, Australia


References


Gray AN, Azuma DL (2005) Repeatability and implementation of a forest vegetation indicator. Ecological Indicators 5:57-71


191
References

Harris GP, Heathwaite AL (2012) Why is achieving good ecological outcomes in rivers so difficult? Freshwater Biology 57:91-107


Läjer K (2007) Statistical tests as inappropriate tools for data analysis performed on non-random samples of plant communities. Folia Geobotanica 42:115-122

References


References


References


Smith VR (2002) Climate change in the sub-Antarctic: an illustration from Marion Island. Climatic Change 52:345-357


Wallace L (2006) GIS based snow accumulation model for the Windmill Islands. School of Geography and Environmental Studies, University of Tasmania, Studio Project.


References


