

Estimating change in areas of indigenous vegetation cover in New Zealand from the New Zealand Land Cover Database (LCDB)

John R. Dymond^{1*}, James D. Shepherd¹, Peter F. Newsome¹ and Stella Belliss²

¹Landcare Research, Private Bag 11052, Palmerston North 4410, New Zealand

²Landcare Research, PO Box 69040, Lincoln 7640, New Zealand

*Author for correspondence (Email: Dymondj@landcareresearch.co.nz)

Published online: 22 September 2016

Abstract: Four iterations of the New Zealand Land Cover Database have been produced from satellite imagery for nominal dates of 1996/97, 2001/02, 2008/09 and 2012/13. These data may be used to estimate changes in area for land cover classes of interest. However, these estimates are subject to uncertainty, which can be significant, particularly when change in area is small. Changes in indigenous vegetation classes are of interest for a number of applications, including monitoring threatened environments. Here we show how the combination of exhaustive sampling of change polygons with random ‘truth’ sampling can be used to estimate the uncertainty of area change. We demonstrate the method on five important indigenous covers: indigenous forest, broadleaved indigenous hardwoods, manuka and/or kanuka, tall tussock grassland, and subalpine shrubland. For these classes, we estimate their area in 2008/09 and the change of area between 2001/2002 and 2008/09. Areas were estimated to within plus or minus 5%. Change in areas were estimated to within plus or minus 10% of change for classes with a large change in area, and to within plus or minus 30% for the classes with a small change in area. We anticipate similar uncertainties for estimated changes in area between the other dates and for other classes. The number of random ‘truth’ samples required for this assessment was very high, in excess of 30 000. Many more samples would be required to further lower the uncertainties.

Keywords: area change; land cover change; LCDB; random sampling; satellite imagery; uncertainty

Introduction

The New Zealand Land Cover Database (LCDB) is a regularly updated national map of land cover. Thus far it has been produced from satellite imagery at four different dates: 1996/97, 2001/02, 2008/09 and 2012/13. The first map was produced by manually drawing polygons on SPOT imagery. Subsequent updates of the database have been made through automatic detection of spectral changes in satellite imagery (both SPOT and LANDSAT), which facilitates an efficient change-site-based update rather than green-fields mapping (i.e. only those areas identified as having changed are remapped rather than everywhere being remapped). The processing of satellite imagery to standardised spectral reflectance, necessary for the change detection, is described by Dymond et al. (2012). Revised polygons may be automatically drawn by the spectral change detection, or redrawn manually if the automatically generated boundaries are not thought sufficiently accurate when overlain on the satellite imagery.

The regular updating of LCDB makes it useful for regional and national reporting of indigenous cover in New Zealand. Walker et al. (2006) used it to assess the proportions of historic ecosystems remaining and this was updated by Cieraad et al. (2015). However, due to an absence of accuracy estimates of change there has been some contention about the conclusions of these studies (Brockerhoff et al. 2008; Walker et al. 2008). The national indicator ‘Indigenous Cover and Protection in Land Environments’ (Statistics New Zealand 2015) also uses the LCDB to estimate the area of land environments remaining in indigenous cover. The LCDB is used because it is regularly updated (~6 yearly) and has a high overall mapping accuracy (>90%; Dunningham et al. 2000; Land Cover Database v3.0).

However, even though the overall mapping accuracy is known, it is not known how well the LCDB can report the *area* of individual classes nor how well it can report the *change in area* of a class.

Methods for reporting map accuracy have been long established (Van Genderen et al. 1978). They are based on a contingency table (i.e. the ‘confusion matrix’) of mapped classes versus ‘true’ classes, which are populated by randomly sampled points of concurrent mapped and ‘truth’ classes. The proportion of samples correctly mapped is used as an indicator of map accuracy, termed the overall map accuracy. However, users often wish to report the areas of individual cover classes and their associated uncertainties. Card (1982) showed how the confusion matrix could also be used to adjust the map area of a cover class for bias and thence infer uncertainty. He showed that an unbiased estimator for class area was merely the sum of proportions in the true class column of the confusion matrix.

Conceptually, the same approach can be applied to a change map constructed from two land cover maps at different dates (Biging et al. 1999; Foody 2002). A confusion matrix can be formed from mapped change classes versus true change classes. However, the number of columns and rows are now many more than the single date confusion matrix, and many more samples are required to achieve reasonable accuracy (Biging et al. 1999). Indeed, many of the change classes are rarely populated with samples, creating significant estimation difficulties. Some authors have suggested methods for improving sampling efficiency of change maps, such as grouping change classes (Olofsson et al. 2014) or exhaustive sampling of some change classes (Dymond et al. 2008).

Here, we estimate how accurately the area and change in the areal extent of individual cover classes may be determined

from the LCDB. We choose five important indigenous classes, based on abundance, as examples: indigenous forest, broadleaved indigenous hardwoods, mānuka and/or kānuka, subalpine shrubland, and tall tussock grassland. The method requires a random sample of points stratified by map class to determine how well the mapping compares with truth (Card 1982). A systematic correction can then be made to area and area change estimates, and the uncertainty of that systematic correction can be determined statistically. However, the number of samples required to achieve a low uncertainty is very high for the area change estimates (Dymond et al. 2008) and significant resources are required. So for efficiency's sake we have incorporated previous 'truth' sampling performed on area change between the dates 2001/02 and 2008/09. Hence results are presented for these two dates only and considered indicative of other dates.

The Land Cover Database

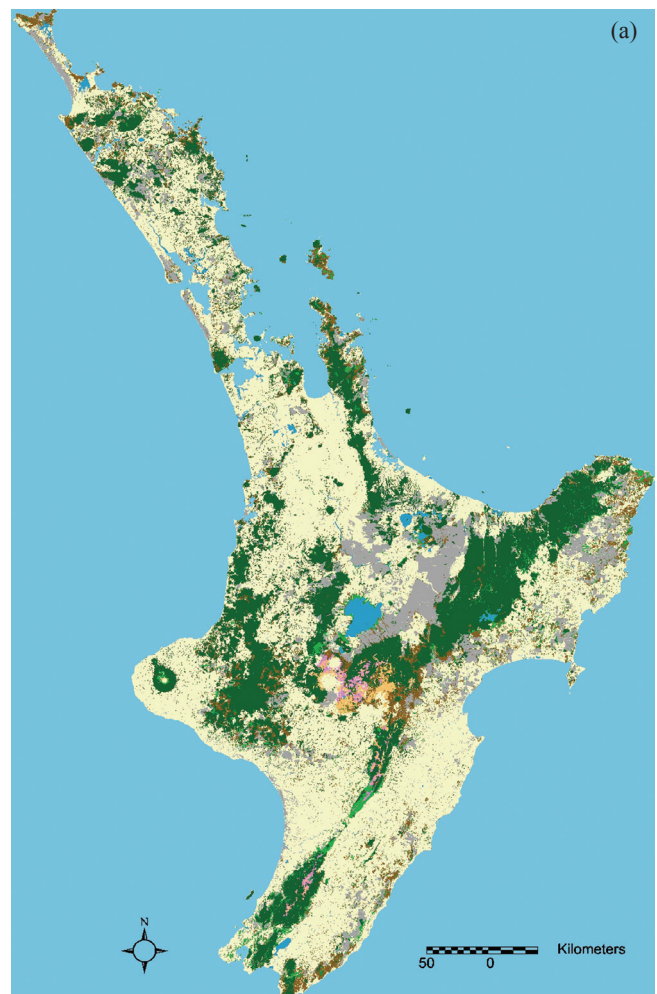
The New Zealand Land Cover Database (LCDB) is a multi-temporal, thematic, spatial classification of land cover that is presently in its fourth edition. Land cover is recorded in the LCDB nominally for the New Zealand summers of 1996/97, 2001/02, 2008/09 and 2012/13, corresponding to versions 1, 2, 3 and 4, respectively. Although the classification has evolved over time, compatibility with earlier classifications has been maintained and each revision of the database is internally consistent in classification for all time steps. For example, version 4 includes revisions of all four dates in the release: 1996/97, 2001/02, 2008/09 and 2012/13. Version 4.0 has now been deprecated in favour of an improved version 4.1. Appendix 1 gives a table of classes and a brief description of each.

All versions of the LCDB have employed visual interpretation and manual digitising, informed and assisted by remote sensing and image processing and classification of multi-spectral satellite imagery. LCDB v1 was created *de novo* from classified SPOT satellite imagery with polygons either captured directly from the classified imagery or manually digitised. The original classification to 16 land cover classes was expanded to 43 for LCDB v2 and polygon boundaries were significantly refined by manual digitising over Landsat 7 ETM+ satellite imagery. Mapping of this second (2001/02) time step was informed by a 'difference layer' created by comparison between 1996/97 and 2001/02 imagery and an extensive field checking phase sought to verify the mapping (Thompson et al. 2003).

LCDB v3 and v4 consolidated and refined the mapping process by first rationalising the classification to 33 mainland classes, then smoothing polygon boundaries to remove latent artefacts of the early raster mapping, and finally aligning the

Figure 1a. Map of main indigenous vegetation classes in the North Island as at 2012/13: indigenous forest (dark green); broadleaved indigenous hardwoods (light green); mānuka and/or kānuka (dark brown); tussock grassland (light brown); subalpine shrubland (pink). Exotic forest is grey and other classes are light yellow.

Figure 1b. Map of main indigenous vegetation classes in the South Island as at 2012/13: indigenous forest (dark green); broadleaved indigenous hardwoods (light green); mānuka and/or kānuka (dark brown); tussock grassland (light brown); subalpine shrubland (pink). Exotic forest is grey and other classes are light yellow.



mapping with the standard topographic coastline. Improved imagery and image classification techniques, combined with resource limitations, prompted a decision to not undertake widespread field checking. Instead, ancillary data and stakeholder reviews were built into the mapping workflow. Change mapping remained a manual on-screen process, but was assisted by improved imagery difference detection, supporting imagery including aerial photography, contributing datasets (such as New Zealand's Kyoto Land Use mapping and regional wetland mapping), and both geographic and thematic error-correction workflows.

Figure 1 shows the five important indigenous classes in LCDB v4.1 for 2012/13. Both the North and South Islands have large tracts of indigenous forest, and the South Island also has large tracts of tall tussock grassland. Subalpine shrubland is also common in the South Island. Both Islands have sporadic distributions of mānuka and/or kānuka and broadleaved indigenous hardwoods. Table 1 shows the area of the five indigenous classes as mapped by LCDB v4.1. Table 2 shows the change in the areas between the four different mapping dates, as well as the change in area between the beginning and end dates. The question is how accurate, or uncertain, are these estimates?

Methods

Uncertainty of class area

A common method for estimating the uncertainty of the area of a class in a land cover map is to randomly sample a large number of points on the map (Card 1982; Dymond et al. 2008). At each point, the mapped class is compared with the 'true' class and the sample is placed in a contingency table to record the frequency of the possible outcomes. The 'true' classes ideally should be determined on the ground to make 'ground-truth', but this is often too costly when there are many samples so 'true' classes are commonly determined by visual interpretation of imagery (Olofsson et al. 2014), as is the case here. Not only is determining ground-truth costly, but there are often problems with access, such that missed sites

can compromise statistical validity. Visually interpreted data may be used in place of ground-truth provided they are of higher quality than the land cover maps (Olofsson et al. 2014).

Our truth data come from visual interpretation of a time sequence of satellite imagery (1990 Landsat TM, 1996/97 SPOT, 2001/02 Landsat TM, 2008/09 SPOT5) and 2009 natural colour SPOT Maps (Harris MapMart - SPOTMaps), the latter at 2.5 m pixel resolution. The five different dates of imagery gave superior contextual information to the two dates used for the land cover maps (2001/02 and 2008/09) and the higher spatial resolution of the SPOT Maps also gave superior textural information. Interpretation was based on documented experience of how real ground features relate to spectral signatures (Thompson et al. 2003).

Photo-interpreters used customised software to move from sample point to sample point while viewing co-registered windows of imagery. Images were nominally displayed at 1:50 000 scale, but could be zoomed to 1:10 000 when necessary with a single button press. For single sample points, the actual land cover of a 1 ha square surrounding the point was compared with that recorded in the land cover map. For polygon samples, the actual predominant land cover of the polygon was compared with that recorded in the land cover map.

Errors in mapped land cover change primarily come from four sources: (1) failure to identify true change through automatic detection of spectral change; (2) delineation of change polygon boundaries; (3) operator error; and (4) photo-interpretation error.

The visually interpreted truth data will minimise type (2) and type (3) errors through greater care and attention than possible in the mapping process, and will also minimise type (4) errors through access to greater contextual and textural information. Therefore, we assume that our truth data are of higher quality than the mapped data, as required by Olofsson et al. (2014), and the sampling system will assess type (1) errors.

Conversion of frequency to probability by dividing by the number of samples will result in a probability table like Table 3 for an example class, in this case forest. The probabilities in the first column of Table 3, that is, $p1$ and $p2$, may be summed to obtain the probability of a sample actually being forest. It

Table 1. Areas of indigenous vegetation classes (in 1000s of ha) considered here as mapped by the LCDB v4.1

Date	1996/97	2001/02	2008/09	2012/13
Tall tussock grassland	2368	2357	2346	2337
Mānuka and/or kānuka	1192	1192	1185	1173
Broadleaved ind. hardwoods	662	657	657	656
Subalpine shrubland	432	432	433	433
Indigenous forest	6319	6315	6311	6309

Table 2. Area change of indigenous vegetation classes (in 1000s of ha) considered here as mapped by the LCDB v4.1.

Date change	1996/97–2001/02	2001/02–2008/09	2008/09–2012/13	1996/97–2012/13
Tall tussock grassland	−10.7	−11.4	−8.8	−30.9
Mānuka and/or kānuka	−0.3	−7.4	−11.2	−18.9
Broadleaved ind. hardwoods	−5.1	0.0	−0.4	−5.5
Subalpine shrubland	0.1	0.3	0.1	0.5
Indigenous forest	−3.7	−4.6	−2.2	−10.5

Table 3. Probability of a random sample co-occurring as fF , fG , gF or gG . For example, $p1$ is the probability of a sample co-occurring as fF , that is mapped as forest and actually being forest.

	Actually forest F	Actually non-forest G
Mapped as forest f	$p1$	$q1$
Mapped as non-forest g	$p2$	$q2$

follows that an estimate of the ‘true’ area of forest, as opposed to the mapped area, is then given by $(p1+p2)$ times the area of New Zealand, A . The estimated ‘true’ area of forest can either be higher or lower than the mapped area of forest and represents a correction to systematic error in the mapped area. The variance of the ‘true’ area may then be estimated by the variance of $(p1+p2)$.

$$\text{Estimated 'true' area of forest} = A \sum_{i=1}^2 p_i \quad (1)$$

and

$$\text{variance of 'true' area of forest} = A^2 \sum_{i=1}^2 \text{Var}(p_i) \quad (2)$$

where the variance of p_i is that of the well-known binomial distribution:

$$\text{Var}(p_i) = \frac{p_i(1-p_i)}{n-1} \quad (3)$$

where n is the number of random samples.

To make sampling efficient, stratified random sampling is usually adopted. In this methodology, more samples are allocated to larger or more variable cover types (Cochran 1977, p. 109). Thus the probabilities in Table 3 would be the probability of occurrence in the strata weighted by the proportion of the strata in New Zealand. For example, in Table 3, under random sampling stratified by map class:

$$p1 = P(F|f)P(f) \quad (4)$$

where $P(F|f)$ is the probability of occurrence of F in the f stratum and $P(f)$ is the proportion of the f stratum in New Zealand.

Likewise under random sampling stratified by map class, the variance of a probability in Table 3 would then be the weight squared times the variance of the probability of occurrence in the stratum (Dymond et al. 2008):

$$\text{Var}(p1) = \frac{P(f)^2 P(F|f)(1-P(F|f))}{(n_f-1)} \quad (5)$$

where n_f is the number of samples in the f stratum.

For each of the five indigenous covers to be assessed for uncertainty of area, we collected truth data at 300 random points in each map class and at 9600 random points elsewhere. For example, for indigenous forest we collected truth data at 300 random points in the area mapped as indigenous forest and at 9600 random points elsewhere. Likewise for the four other important indigenous covers.

Uncertainty of change in class area

The method for estimating the uncertainty of change in class

area uses a probability table as described in the previous section. However, there are more classes to consider now, with each class being a pair of classifications (at first and second date): see Table 4. For example, if a sampled point changes from being mapped as forest to being mapped as non-forest, it is denoted fg .

The true area of forest loss is given by the sum of probabilities in column FG times the area of New Zealand, A . And the true area of gain of forest is given by the sum of probabilities in column GF times A . These may be subtracted to get the net change in forest area. The variance may be estimated similarly to equation (2):

$$\text{estimated change in 'true' area of forest} = A(\sum_{i=1}^4 r_i - \sum_{i=1}^4 q_i) \quad (6)$$

and

$$\text{variance of change in 'true' area of forest} = A^2(\sum_{i=1}^4 \text{Var}(r_i) + \sum_{i=1}^4 \text{Var}(q_i)) \quad (7)$$

where variances are again of the binomial distribution.

For each of the five indigenous covers to be assessed for uncertainty of change, we collected truth data *exhaustively* for all change polygons involving change in the indigenous cover. We collected truth data from 20 000 random points elsewhere. For example, for indigenous forest we collected truth data from 20 000 random points in the area mapped as not having a change from or to indigenous forest, and we exhaustively sampled elsewhere.

Results

Area of indigenous cover classes

Table 5 shows the sampling results for indigenous forest. Of the 300 random samples in mapped indigenous forest (i.e. f), three were not actually indigenous forest (i.e. they were actually G) and the rest (297) were actually indigenous forest (i.e. they were actually F). Of the 9600 random samples in non-indigenous forest (i.e. g), 104 were actually indigenous forest and the rest (9496) were not actually indigenous forest. When samples of mapped indigenous forest are actually indigenous forest they are counted as fF , otherwise they are counted as fG . The other 9600 samples of mapped classes not belonging to indigenous forest are counted in either gF or gG , depending on whether they were actually indigenous forest or not.

Table 5 contains the number of samples. These may be converted to area (Table 6) by multiplying the area of the strata (f or g) by the proportion of samples in the strata. For example, the area in fF is 6311 thousand ha times 297/(297+3), which is 6247 thousand ha. Likewise, the area in gF is 20 531 thousand ha times 104/(104+9496), which is 222 thousand ha.

The estimated true area of indigenous forest is simply the addition of the areas in column F , that is 6247 thousand ha plus 222 thousand ha, which makes 6469 thousand ha. The variance of the estimated true area is the variance of ‘6247’ plus the variance of ‘222’. The variance of ‘6247’ is estimated using equation (5) with $n_f = 300$, $P(f) = 6311/(6311+20 531)$, and $P(F|f) = 297/(297+3)$, all from Table 5. Likewise the variance of ‘222’ is estimated using equation (5) with $n_g = 9600$, $P(g) = 20 531/(6311+20 531)$, and $P(F|g) = 104/(104+9496)$, again from Table 5.

Table 4. Probability of a random sample co-occurring as *ff*-*FF*, *ff*-*FG*, etc. For example, r_2 is the probability of a sample co-occurring as *fg*-*GF*, that is, mapped as going from forest to non-forest but actually has gone from non-forest to forest.

Mapped classes	‘Truth’ classes			
	<i>FF</i>	<i>FG</i>	<i>GF</i>	<i>GG</i>
<i>ff</i>	p_1	q_1	r_1	s_1
<i>fg</i>	p_2	q_2	r_2	s_2
<i>gf</i>	p_3	q_3	r_3	s_3
<i>gg</i>	p_4	q_4	r_4	s_4

Table 5. Number of random samples mapped as indigenous forest (*f*) or not (*g*) and that actually are indigenous forest (*F*) or not (*G*) (mapping date: 2008/09). There are 300 random samples in mapped indigenous forest and 9600 samples elsewhere.

	<i>F</i>	<i>G</i>	map area (1000s ha)
<i>f</i>	297	3	6311
<i>g</i>	104	9496	20 531

Table 6. Estimated areas (1000s ha) mapped as indigenous forest (*f*) or not (*g*) and that are actually indigenous forest (*F*) or not (*G*) (mapping date 2008/09).

	<i>F</i>	<i>G</i>	Strata area (1000s ha)
<i>f</i>	6247	64	6311
<i>g</i>	222	20 309	20 531
	6469		

Table 7. Area (1000s ha) of main indigenous cover at 2008/09 showing the effect of adjusting for systematic mapping error as determined by stratified random sampling. Uncertainty is given at the 95% confidence level.

Date = 2008/09	LCDB area (1000s ha)	Adjusted area (1000s ha)	Uncertainty (±1000s ha)	Uncertainty (±%)
Tall tussock grassland	2346	2294	70	3.1%
Mānuka and/or kānuka	1185	1325	54	4.1%
Broadleaved ind. hardwoods	657	622	40	6.5%
Subalpine shrubland	433	408	28	6.8%
Indigenous forest	6311	6469	85	1.3%

Table 8. Estimated areas (1000s ha) mapped as loss (*fg*), gain (*gf*), or no change (*ff* or *gg*) of indigenous forest, and are actually loss (*FG*), gain (*GF*), or no change (*FF* or *GG*). The *fg* and *gf* rows were exhaustively sampled, while ‘*ff* or *gg*’ were randomly sampled with 20 000 points. Mapping dates were 2001/02–2008/09.

		<i>FG</i>	<i>GF</i>	<i>FF</i> or <i>GG</i>	Strata area (ha)
all polygons	<i>fg</i>	4.2	0.0	0.6	4.8
all polygons	<i>gf</i>	0.0	0.0	0.2	0.2
	<i>ff</i> or <i>gg</i>	0.0	0.0	26 836.7	26 836.7

The uncertainty of estimated true area of indigenous forest may then be reported approximately at the 95% confidence level as two times the square root of the variance. So the adjusted area of indigenous forest is estimated to be 6469 thousand hectares plus or minus 85 thousand hectares, that is plus or minus 1.3%.

Tables 5 and 6 show the workings for indigenous forest only. Results only for the other important indigenous covers are shown in Table 7. The uncertainties at the 95% confidence level are all less than 7% with that of indigenous forest being the smallest at 1.3%.

Area change of indigenous cover classes

Many samples are required before an adjusted *change* in area can be estimated with certainty. We sampled 20 000 random points on areas mapped as not having changed (i.e. *ff* or *gg*) between 2001/02 and 2008/09. We have minimised the variance in the *fg* and *gf* cells in Table 8 by exhaustively sampling all change polygons (i.e. *fg*, or *gf*). This involved revisiting every polygon that involved a change to or from the mapped class, which is indigenous forest in this example. The effort required for this was reasonable as there are typically in the order of a few thousand polygons of change to check.

Table 8 shows the estimated areas from the sampling of mapped change versus actual change for indigenous forest (2001/02–2008/09). The estimated actual area of forest lost is the sum of 4.2 thousand ha, 0.0 thousand ha, and 0.0 thousand ha, which makes 4.2 thousand ha. The strata *fg*-*FG* and *gf*-*FG* are sampled exhaustively, so their variance is zero. The ‘*ff* or *gg*’ stratum has been randomly sampled with 20 000 points and not one has been recorded as *FG*. This estimated variance of the binomial distribution is zero, but it could easily have been non-zero (i.e. the true probability of map error will always be non-zero), so we conservatively set the variance as if there had been one sample in the *FG* cell. The adjusted area of forest loss is then 4.2 (±0.6) thousand ha. Likewise, the adjusted area of forest gain is given by 0.0 thousand ha (±0.0)

Table 9. Adjusted change in area (1000s ha) between 2001/02 and 2008/09. A negative figure denotes loss.

Dates = 2001/02–2008/09	LCDB mapped change in area (1000s ha)	Adjusted change in area (1000s ha)	Uncertainty (\pm 1000s ha)	Uncertainty (\pm %)
Tall tussock grassland	–11.4	–9.2	0.6	7%
Mānuka and/or kānuka	–7.4	–11.8	2.6	22%
Broadleaved ind. hardwoods	0.0	–0.8	0.3	40%
Subalpine shrubland	0.3	0.3	0.6	200%
Indigenous forest	–4.6	–4.2	0.6	15%

thousand ha. The net area of forest loss is then by 4.2 (\pm 0.6) thousand ha. Table 9 shows the adjusted change in area for the main indigenous covers and their associated uncertainties.

Discussion

We have shown how ‘truth’ sampling of the LCDB was able to produce adjusted estimates of area of land cover classes for 2008/09. These adjustments account for systematic errors in the mapping process. The variance of the adjustments may be estimated and used to characterise uncertainty. For indigenous forest, the adjusted area was estimated accurately, to within \pm 1.3%. The uncertainties for tall tussock grassland and mānuka and/or kānuka were also small, at less than 5%. However, for subalpine shrublands and broadleaved indigenous hardwoods, the uncertainties were moderate, exceeding 5%. This higher uncertainty is primarily due to their smaller areas in total.

How may we use these results to infer uncertainties of the area of indigenous cover at dates other than 2008/09? If we were to undertake ‘truth’ samples of LCDB at those dates, involving in the order of 10 000 samples, we would obtain similar percentage uncertainties because change between the dates is relatively small, at a few percent only. So we expect the uncertainty of classes with an area greater than one million hectares to have an uncertainty of less than \pm 5%. We currently do not have the resources to assess a further 30 000 random samples at each of the other three dates of the LCDB, so we will use plus or minus c. 5% as an indicative uncertainty.

Many more ‘truth’ samples would be required to further reduce uncertainties. For example, to halve the uncertainty of reported change in indigenous forest from 15% to 7.5% would require 80 000 random samples in addition to the exclusively sampled change polygons. This highlights the necessity of using visual interpretation of imagery for ‘truth’ data. Ground inspection of sampled sites, while desirable because it can observe data not possible by visual interpretation, is significantly more labour intensive and therefore less likely to achieve a sufficiently large number of samples. An example of data not observable by interpretation is the partial confusion matrix of Mason et al. (2012) (involving 1243 ground-truth samples), which showed that 29% of samples mapped as shrubland by the LCDB were observed to be indigenous forest. On the ground, canopy height can be observed and a height criterion (of 5 m) used to separate the two classes. Visual interpretation, on the other hand, relies on using the spectral signature, which is primarily controlled by species composition and condition of the canopy. In the future, canopy height models from LiDAR may be available for the LCDB to use a height-based separation between indigenous forest and shrubland.

As a percentage, the uncertainty of change in cover class area is comparatively much higher than the uncertainty of cover class area alone. This is because the areas of change are much smaller than total areas. Total land-cover change in New Zealand over a period of four years is only of the order of a few percent. For tall tussock grassland and indigenous forest, the change area can be reported to less than \pm 15%. For mānuka and/or kānuka and broadleaved indigenous hardwoods, the change in area can only be reported to less than \pm 50%. For subalpine shrublands, which have a small total area, the change area can only be reported to less than plus 200% (and minus 100%). Our reported uncertainties may be acceptable for the tall tussock grassland and indigenous forest classes, but not for the other classes. Hence, much more intensive sampling is required to reduce the uncertainties. Higher classification accuracy would also reduce uncertainties, but this would be difficult to achieve with the current imagery set. In the meantime, an indicative uncertainty of plus or minus \sim 10% may be used for the larger change classes, and plus or minus \sim 30% for the smaller classes (excluding tall tussock grassland).

The results here show there is uncertainty with using the LCDB to estimate areas of land covers in New Zealand. However, we have quantified this uncertainty by comprehensive truth sampling of over 30 000 samples (as opposed to the 67 of Brockerhoff et al. 2008), involving both omission and commission errors as indicated necessary by Walker et al. (2008). This uncertainty should be taken into account in analyses of area of indigenous cover, such as that of Walker et al. (2006) who estimated the proportion of indigenous cover remaining in historical ecosystems. The uncertainty of these proportions should be of the same order as the \pm 5% estimated in this paper for the areas of indigenous cover classes, thus justifying the six threat categories proposed by Walker et al. (2006). However, the higher uncertainty of estimating change in area may have a significant impact on reporting change in threat category for some ecosystems.

Conclusions

A sample of ‘truth’ data may be used to estimate the uncertainty of areal change as given by the LCDB. An exhaustive sample of change polygons improves the efficiency significantly, but the number of samples required to achieve acceptable uncertainty is very high – over 30 000. When the method was applied to the changes in area of five important indigenous covers between 2001/02 and 2008/09, uncertainties of the order \pm 10% were achieved for the larger change classes and \pm 30% for the smaller classes. We anticipate similar uncertainty for area change between other dates and for other classes. When the method was applied to area alone of the five indigenous covers

at 2008/09, uncertainties of the order of $\pm 5\%$ were achieved. We anticipate similar uncertainty for area at the other dates.

References

- Biging GS, Colby DR, Congalton RG 1999. Sampling systems for change detection accuracy assessment. In: Lunetta RS, Elvidge C eds. Remote sensing change detection: environmental monitoring methods and applications. Taylor & Francis, London. Pp. 281–308.
- Brockerhoff EG, Shaw WB, Hock B, Kimberley M, Paul T, Quinn J, Pawson S 2008. Re-examination of recent loss of indigenous cover in New Zealand and the relative contributions of different land uses. *New Zealand Journal of Ecology* 32: 115–126.
- Card DH 1982. Using known map category marginal frequencies to improve estimates of thematic map accuracy. *Photogrammetric Engineering and Remote Sensing* 48: 431–439.
- Cieraad E, Walker S, Price R, Barringer J 2015. An updated assessment of indigenous cover remaining and legal protection in New Zealand land environments. *New Zealand Journal of Ecology* 39: 309–315.
- Cochran WG 1977. Sampling techniques. 3rd edn. New York, John Wiley & Sons. 428 p.
- Dunningham A, Brownlie R, Firth J 2000. Classification accuracy assessment of NZLCBD1: results. Report to the Ministry for the Environment, Wellington. Rotorua, New Zealand Forest Research Institute.
- Dymond JR, Shepherd JD, Arnold GC, Trotter CM 2008. Estimating area of forest change by random sampling of change strata mapped using satellite imagery. *Forest Science* 54: 475–480.
- Dymond JR, Shepherd JD, Newsome PF, Gapare N, Burgess DW, Watt P 2012. Remote sensing of land-use change for Kyoto Protocol reporting. *Environmental Science and Policy* 16: 1–8.
- Foody GM 2002. Status of land cover classification accuracy assessment. *Remote Sensing of Environment* 80: 185–201.
- Harris MapMart – SPOTMaps. <http://cms.mapmart.com/Products/SatelliteImagery/SPOTMaps.aspx> (accessed 5/9/2016).
- Land Cover Database v3.0 Accuracy assessment. <http://www.lcdb.scinfo.org.nz/about-lcdb/accuracy-assessment> (accessed 5/9/2016).
- Land Cover Database v4.1 – Land Cover Database version 4.1, Mainland New Zealand. <https://iris.scinfo.org.nz/layer/423-lcdb-v41-land-cover-database-version-41-mainland-new-zealand/> (accessed 5/9/2016).
- Mason NWH, Carswell FE, Overton JMcC, Briggs CM, Hall GMJ 2012. Estimation of current and potential carbon stocks and Kyoto-compliant carbon gain on conservation land. *Science for Conservation* 317. Wellington, Department of Conservation. 39 p.
- Olofsson P, Foody GM, Herold M, Stehman SV, Woodcock CE, Wulder MA 2014. Good practices for estimating area and assessing accuracy of land change. *Remote Sensing of Environment* 148: 42–57.
- Statistics New Zealand 2015. Environment Aotearoa. Indigenous Cover and Protection in Land Environments. www.stats.govt.nz/browse_for_stats/environment/environmental-reporting-series/environmental-indicators/Home/Land/indigenous-cover.aspx (accessed 5/9/2016).
- Thompson S, Gruner I, Gapare N 2003. New Zealand Land Cover Database version 2: illustrated guide to target classes. Wellington, Ministry for the Environment. <https://iris.scinfo.org.nz/layer/423-lcdb-v41-land-cover-database-version-41-mainland-new-zealand/> (accessed 5/9/2016).
- Van Genderen JL, Lock BF, Vass PA 1978. Remote sensing: statistical testing of thematic map accuracy. *Remote Sensing of Environment* 7: 3–14.
- Walker S, Price R, Rutledge D, Stephens T, Lee WG 2006. Recent loss of indigenous cover in New Zealand. *New Zealand Journal of Ecology* 30: 169–177.
- Walker S, Price R, Rutledge D, Stephens T, Lee WG 2008. Measuring accuracy of land cover data and content of cover classes: a reply to Brockerhoff et al. (2008). *New Zealand Journal of Ecology* 32: 127–129.

Editorial board member: Hannah Buckley

Received 10 September 2015; accepted 9 June 2016

Appendix 1. LCDB V4.1 classes and descriptions

Class code	Class name	Class description
1	Built-up area (settlement)	Commercial, industrial or residential buildings, including associated infrastructure and amenities not resolvable as other classes. Low density 'lifestyle' residential areas are included where hard surfaces, landscaping and gardens dominate other land covers.
2	Urban parkland/open space	Open, mainly grassed or sparsely treed, amenity, utility and recreation areas. The class includes parks and playing fields, public gardens, cemeteries, golf courses, berms and other herbaceous areas usually within or associated with built-up areas.
5	Transport infrastructure	Artificial surfaces associated with transport such as arterial roads, rail-yards and airport runways. Skid sites and landings associated with forest logging are sometimes also included.
6	Surface mine or dump	Bare surfaces arising from open-cast and other surface mining activities, quarries, gravel-pits and areas of solid waste disposal such as refuse dumps, clean-fill dumps and active reclamation sites.
10	Sand or gravel	Bare surfaces dominated by unconsolidated materials of texture generally finer than coarse gravel (60 mm). Typically mapped along sandy seashores and the margins of lagoons and estuaries, lakes and rivers; and some areas subject to surficial erosion, soil toxicity and extreme exposure.
12	Landslide	Bare surfaces arising from mass-movement erosion generally in mountain-lands and steep hill-country.
14	Permanent snow and ice	Areas where ice and snow persists through late summer. Typically occurring above 1800 m but also at lower elevations as glaciers.
16	Gravel or rock	Bare surfaces dominated by unconsolidated or consolidated materials of texture generally coarser than coarse gravel (60 mm). Typically mapped along rocky seashores and rivers, sub-alpine and alpine areas, scree slopes and erosion pavements.
15	Alpine grass/herbfield	Typically sparse communities above the actual or theoretical treeline dominated by herbaceous cushion, mat, turf, and rosette plants and lichens. Grasses are a minor component or infrequent, whereas stones, boulders, and bare rock are usually conspicuous.
20	Lake or pond	Essentially permanent, open, fresh-water without emerging vegetation including artificial features such as oxidation ponds; amenity, farm and fire ponds and reservoirs; and natural lakes, ponds and tarns.
21	River	Flowing open fresh-water generally more than 30 m wide and without emerging vegetation. It includes artificial features such as canals and channels as well as natural rivers and streams.
22	Estuarine open water	Standing or flowing saline water without emerging vegetation including estuaries, lagoons and occasional lakes occurring in saline situations such as inter-dune hollows and coastal depressions.
30	Short-rotation cropland	Land regularly cultivated for the production of cereal, root and seed crops, hops, vegetables, strawberries and field nurseries, often including intervening grassland, fallow land, and other covers not delineated separately.
33	Orchards, vineyards or other perennial crops	Land managed for the production of grapes, pip, citrus and stone fruit, nuts, olives, berries, kiwifruit, hops, and other perennial crops. Cultivation for crop renewal is infrequent and irregular but is sometimes practised for weed control.
40	High producing exotic grassland	Exotic sward grassland of good pastoral quality and vigour reflecting relatively high soil fertility and intensive grazing management. Clover species, ryegrass and cocksfoot dominate, with lucerne and plantain locally important but also including lower-producing grasses exhibiting vigour in areas of good soil moisture and fertility.
41	Low producing grassland	Exotic sward grassland and indigenous short tussock grassland of poor pastoral quality reflecting low soil fertility and extensive grazing management or non-agricultural use. Browntop, sweet vernal, danthonia, fescue, and Yorkshire fog dominate, with indigenous short tussocks (hard tussock, blue tussock, and silver tussock) common in the eastern South Island and locally elsewhere.
43	Tall tussock grassland	Indigenous snow tussocks in mainly alpine mountain-lands and red tussock in the central North Island and locally in poorly drained valley floors, terraces, and basins of both islands.

Continued...

Appendix 1. *continued*

Class code	Class name	Class description
44	Depleted grassland	Areas, mainly former short tussock grassland in the drier eastern South Island high country, degraded by over-grazing, fire, rabbits, and weed invasion, among which <i>Hieracium</i> species are conspicuous. Short tussocks usually occur, as do exotic grasses, but bare ground is more prominent.
45	Herbaceous freshwater vegetation	Herbaceous wetland communities occurring in freshwater habitats where the water table is above or just below the substrate surface for most of the year. The class includes rush, sedge, restiad, and sphagnum communities and other wetland species, but not flax or willows, which are mapped as flaxland and deciduous hardwoods respectively.
46	Herbaceous saline vegetation	Herbaceous wetland communities occurring in saline habitats subject to tidal inundation or saltwater intrusion. Commonly includes club rush, wire rush and glasswort, but not mangrove which is mapped separately.
47	Flaxland	Areas dominated by New Zealand flax usually swamp flax (harakeke) in damp sites but occasionally mountain flax (wharariki) on cliffs and mountain slopes.
50	Fernland	Bracken fern, umbrella fern or ring fern, commonly on sites with low fertility and a history of burning. Mānuka, gorse, and/or other shrubs are often a component of these communities and will succeed fernland if left undisturbed.
51	Gorse and/or broom	Scrub communities dominated by gorse or Scotch broom generally occurring on sites of low fertility, often with a history of fire, and insufficient grazing pressure to control spread. Left undisturbed, this class can be transitional to broadleaved indigenous hardwoods.
52	Mānuka and/or kānuka	Scrub dominated by mānuka and/or kānuka, typically as a successional community in a reversion toward forest. Mānuka has a wider ecological tolerance and distribution than kānuka, with the latter somewhat concentrated in the north with particular prominence on the volcanic soils of the central volcanic plateau.
54	Broadleaved indigenous hardwoods	Lowland scrub communities dominated by indigenous mixed broadleaved shrubs such as wineberry, mahoe, five-finger, <i>Pittosporum</i> spp., fuchsia, tutu, titoki, and tree ferns. This class is usually indicative of advanced succession toward indigenous forest.
55	Subalpine shrubland	Highland scrub dominated by indigenous low-growing shrubs including species of <i>Hebe</i> , <i>Dracophyllum</i> , <i>Olearia</i> , and <i>Cassinia</i> . Predominantly occurring above the actual or theoretical treeline, this class is also recorded where temperature inversions have created cooler micro-climates at lower elevations, e.g. the 'frost flats' of the central North Island.
56	Mixed exotic shrubland	Communities of introduced shrubs and climbers such as boxthorn, hawthorn, elderberry, blackberry, sweet brier, buddleja and old man's beard.
58	Matagouri or grey scrub	Scrub and shrubland comprising small-leaved, often divaricating shrubs such as matagouri, <i>Coprosma</i> spp., <i>Muehlenbeckia</i> spp., <i>Casinnia</i> spp. and <i>Parsonia</i> spp. which, from a distance, often have a grey appearance.
80	Peat shrubland (Chatham Is)	Low-growing shrubland communities usually dominated by <i>Dracophyllum</i> spp., in association with <i>Cyathodes</i> spp. and ground ferns. Mapped only on Chatham Island.
81	Dune shrubland (Chatham Is)	Low-growing shrubland communities dominated by <i>Leucopogon</i> spp. <i>Pimelia arenaria</i> , and <i>Coprosma</i> spp. in association with sedges and scattered herbs and grasses. Mapped only on Chatham Island.
70	Mangrove	Shrubs or small trees of the New Zealand mangrove (<i>Avicennia marina</i> subspecies <i>australascia</i>) growing in harbours, estuaries, tidal creeks and rivers north of Kawhia on the west coast and Ohiwa on the east coast.
64	Forest – harvested	Predominantly bare ground arising from the harvesting of exotic forest or clearing of indigenous forest. Replanting (or conversion to a new land use) of exotic forest is not evident and nor is the future use of land cleared of indigenous forest.
68	Deciduous hardwoods	Exotic deciduous woodlands, predominantly of willows or poplars but also of oak, elm, ash or other species. Commonly alongside inland water (or as part of wetlands), or as erosion-control, shelter and amenity plantings.
69	Indigenous forest	Tall forest dominated by indigenous conifer, broadleaved and beech species.
71	Exotic forest	Planted or naturalised forest predominantly of radiata pine but including other pine species, douglas fir, cypress, larch, acacia, and eucalyptus. Production forestry is the main land use in this class with minor areas devoted to mass-movement erosion-control and other areas of naturalised (wildling) establishment.