

## Article

# Detection of southern beech heavy flowering using Sentinel-2 imagery

Ben Jolly<sup>1</sup>\*, John R. Dymond<sup>1</sup>, James D. Shepherd<sup>1</sup>, Terry Greene<sup>2</sup>, and Jan Schindler<sup>1</sup>

<sup>1</sup> Manaaki Whenua – Landcare Research;

<sup>2</sup> Department of Conservation, New Zealand;

\* Correspondence: jollyb@landcareresearch.co.nz

Version March 18, 2022 submitted to Remote Sens.

**Abstract:** The southern beech (genus *Fuscospora* and *Lophozonia*) forest in New Zealand periodically has 'mast' years where very large volumes of seed are produced. This excessive seed production results in a population explosion of rodents and mustelids that puts pressure on native birds. To protect the birds, extra pest control, costing in the order of 20 million dollars, is required in mast areas. To plan pest control and keep it cost-effective, it would be helpful to have a map of the mast areas. In this paper, we develop a remote sensing method for making a national map of beech flowering. It uses a temporal sequence of Sentinel-2 satellite imagery to determine areas where a yellow index based on red and green reflectance ( $\text{red-green}/(\text{red} + \text{green})$ ) is higher than normal in spring. The method was used to produce national maps of heavy beech flowering for the years 2017 through to 2021. In 2018, which was a major beech masting year, of the 4.1 million ha of beech forest in New Zealand, 27.6% was observed to flower heavily. The overall classification accuracy of the map was 90.8 %. The method is fully automated and can be used to help identify areas of potentially excessive seed fall, for all New Zealand, several months in advance of when pest control is required.

**Keywords:** southern beech; masting; beech flowering; seed fall; *Fuscospora*; Sentinel-2

## 1. Introduction

New Zealand southern beech (genus *Fuscospora* and *Lophozonia*, formerly *Nothofagus* [1]) forest dominates over 2 million hectares (ha) of New Zealand forest, and features in almost 2 million ha more [2]. It comprises five species: mountain beech (*Fuscospora cliffortioides*), red beech (*Fuscospora fusca*), silver beech (*Lophozonia menziesii*), black beech (*Fuscospora solandri*), and hard beech (*Fuscospora truncata*). The trees reproduce almost yearly, with periodic highly productive seasons known as 'mast' years that produce large volumes of seed [3]. Seed is a significant food source for a number of birds and mammals. During mast years the rodent population increases significantly [4,5], providing an abundant food source for mustelids (esp. stoats – *Mustela erminea*) [6]. All rodent and mustelid species are introduced to New Zealand and also prey on native bird species. When seed begins to run out, increasing pressure is put on native biota as additional food sources are sought. Population control of introduced predators is essential for preserving populations of native and endemic birds, reptiles, and invertebrates in New Zealand, especially during beech mast years [6].

The New Zealand Department of Conservation (DOC) is responsible for managing forests on public land, preserving native species, and coordinating pest control. Prediction of the extent and intensity of significant mast events is a critical component for planning pest control efforts to limit the explosion of predator populations [4]. A number of approaches are currently used for planning management interventions: (i) modelling, (ii) field observations, and (iii) *in situ* sampling of developing seed crops in tree canopies. (i) The 'delta-T' ( $\Delta T$ ) model [7] uses the difference in mean temperature from the previous two summers ( $T_{n-1} - T_{n-2}$ ) to predict likely seed fall for the following autumn

35 at national scale. However, it relies on temperature data, which are currently only available on a  
36 modelled 5 x 5 km grid, so will miss smaller-scale micro-climate effects. While historical temperature  
37 is an important factor in synchronizing mast events [7], other factors such as nutrient availability also  
38 play a role [8]. New research suggests that rising temperatures due to climate change may alter the  
39 mast cycle of beech trees [8,9], increasing the spatial and temporal complexity of mast patterns [8]  
40 and potentially de-synchronizing flowering/seeding, effectively reducing the impact, and predictive  
41 power, of temperature on the timing of the reproductive cycle [10]. (ii) Field staff from DOC are  
42 well-placed to provide observations of beech flowering in certain areas as part of their normal duties.  
43 However, there are large areas of forest that remain unobserved and spatial extent is often difficult to  
44 define, especially at a regional or national scale. (iii) Extensive sampling campaigns are conducted  
45 during years where a heavy mast is expected using helicopters to clip upper branches from trees so  
46 seeds can be counted. This task is expensive, labour-intensive, and dangerous.

47 Remote sensing has proved to be an effective tool for monitoring vegetation phenology,  
48 particularly when a rich time series of imagery is available [11–16]. With sensors like Landsat 8  
49 and Sentinel-2 it is possible to map phenology over millions of hectares to create national maps in  
50 great detail [11,17]. Phenological characteristics are usually derived by first fitting a curve to the time  
51 series of remote sensing data, then using either threshold-based methods, moving averages, inflection  
52 point, or time of maximum increase [18–22]. Seasonality is usually assumed [22]. As mast events are  
53 considered a deviation from the ‘median’ annual cycle, it should be possible to use change detection  
54 techniques [23,24] to identify or even predict mast seasons if identifying features are visible from above  
55 [14,19,21].

56 Vegetation indices, such as the normalized difference vegetation index (NDVI) and the enhanced  
57 vegetation index (EVI), are often used to differentiate flowers or seed pods/cones and to summarize  
58 data as one variable for analysis [14,19,22,25]. Often, multiple indices are combined in an effort to  
59 investigate multiple physical properties. For example, a study by Garcia *et al.* [21] investigating  
60 White Spruce mast events found vegetation indices targeting moisture were more effective than the  
61 traditional color-based ones but still struggled to reliably predict masting. Fernández-Martínez *et al.*  
62 [19] were more successful and used increasing EVI the winter before, along with weather data during  
63 spring, as an indicator of potential mast seasons of Mediterranean oaks. Neither study were able  
64 to find a usable signature of flowering or seed/cone production to map the mast event. Dixon *et al.*  
65 [14] and Chen *et al.* [25] both successfully used multi-scale imagery to map tree-scale flowering over  
66 landscapes, Dixon *et al.* [14] by training a random forest model with drone data from known events  
67 and Chen *et al.* [25] by developing an enhanced bloom index (EBI) from drone data and successfully  
68 translating that approach to CERES, PlanetScope, Sentinel-2, and Landsat data to increase coverage.

69 In this study, we investigate the use of freely-available satellite remote sensing for identifying  
70 large areas of significant southern beech flowering in New Zealand. We use imagery from the European  
71 Space Agency (ESA) Sentinel-2 ‘a’ and ‘b’ satellites to obtain a high rate of repeat passes to maximize  
72 the chances of multiple cloud-free observations, and to produce national coverage at detailed scale.  
73 Very little ground data on flowering are available. The irregular nature of the flowering events, and  
74 ruggedness and remoteness of much of the terrain, makes planning collection of ground data for this  
75 study difficult. As significant flowering is an irregular phenological event, we model the phenology  
76 of beech forest per-pixel and identify departures from this image-by-image during spring seasons  
77 to identify heavy beech flowering. Using this method, we produce national maps of heavy beech  
78 flowering for 5 years, 2017 through 2021, with a minimum mapping unit of 1 ha. We assess the accuracy  
79 of the method by comparison with a human operator at 1000 randomly selected sites. A national  
80 map of detectable heavy beech flowering can be produced at the end of every spring to assist DOC in  
81 identifying potential mast ‘hot-spots’ and thus aid planning of pest control operations.

## 82 2. Materials and Methods

### 83 2.1. Area of Interest

84 This study covers the extent of known beech and mixed-beech forest in New Zealand according to  
85 EcoSat Forests [2,26–28], covering both North and South Islands from 36.1°S, 178.0°E to 46.4°S, 166.4°E.  
86 Terrain is largely mountainous with slopes up to 45° as any habitat deemed suitable for agriculture  
87 was cleared in the 1800's and early 1900's. Due to the large variations in both altitude and latitude of  
88 the study area, mean annual temperatures range from 3.8°C to 16.6°C, and mean cumulative annual  
89 rainfall from 465 mm to 9305 mm.

### 90 2.2. Data

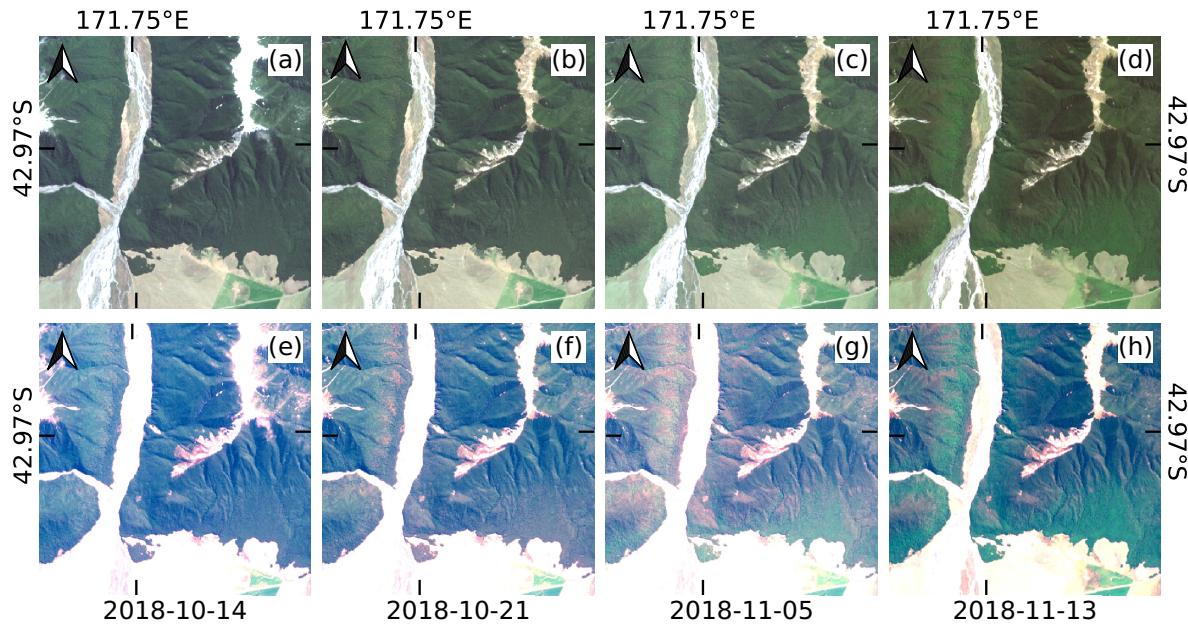
91 This study uses the Copernicus Sentinel-2 Level-1C calibrated top of atmosphere (TOA) reflectance  
92 values as downloaded from the ESA archives, with edges masked to remove pixels that do not contain  
93 data for every band. The Sentinel-2 satellite mission consists of two satellites - Sentinel-2a and -2b - in  
94 sun-synchronous orbits repeating every 10 days. These orbits are 180° out-of-phase with each other,  
95 which produces a 5-day revisit period. The swath width of each pass is 290 km, with five 'passes'  
96 required to cover the mainland of New Zealand, each on a different day. Overlap between passes  
97 means that some areas of the country have a higher revisit rate. Each Sentinel-2 satellite is equipped  
98 with a multi-spectral imaging sensor capturing wavelengths from ultraviolet to short-wave infrared.

99 Analysis uses all available Sentinel-2 data over New Zealand from 2016-09-01 to 2021-12-31, with  
100 images every 10 days before Sentinel-2b came online 2017-07-08 and images every 5 days thereafter.  
101 A 'mega-mast' occurred during spring 2018/autumn 2019 [29,30]. Other years showed some small  
102 flowering events or none at all. Cloudy pixels were excluded from the temporal sequence using the  
103 methodology outlined in Shepherd *et al.* [31].

### 104 2.3. Methods

105 Southern beech flowers produce a reddening of the forest canopy visible from the ground,  
106 especially during a significant mast season. At the 10 m pixel nominal spatial resolution of Sentinel-2,  
107 the reddening of the canopy appears to the human eye as a subtle yellowing in the 'natural color'  
108 (RGB) image as the red flowers increase the red component of a pixel but not to the point of dominating  
109 the green. Higher-wavelength image bands (and associated indices) do not respond to this flowering  
110 with the exception of Band 5 (red edge), however this was also associated with 'flushes' of new foliage  
111 in late spring.

112 The yellowing effect of the canopy is subtle in natural-color renders of Sentinel-2 imagery, as  
113 demonstrated in Figure 1 (a) to (d). Sub-figures (e) to (h) feature an exaggerated 'Red' (B4) band which  
114 highlights the effect, at the expense of non-forested areas. Figure 1 covers one month of imagery in an  
115 area of overlap between orbits 29 and 72 and thus has approximately double the number of overpasses  
116 compared to other areas of the country - in this 30 day period there were 13 overpasses, resulting in 5  
117 usable cloud-free images, 4 of which are displayed. The flowering event is just starting on October  
118 14th (Fig. 1(a,e)), with lower western/southern slopes in full flower 7 days later (b,f), before upper and  
119 eastern slopes are in full flower 15 days later (c,g) as the areas in (b,f) give way to fresh green foliage.  
120 Eight days later on November 13th (d,h) most of the flowering has vanished. Generally there is a two-  
121 to three-week period at most in which a cloud-free image is required to observe flowering, however  
122 there is no way of reliably knowing when this window occurs as it varies with season and location.



**Figure 1.** Sentinel-2 imagery from the Hawdon Valley, South Island, New Zealand showing a flowering event during Spring 2018. (a) - (d) are natural color, (e) - (h) are natural color with an exaggerated stretch to amplify the 'red' band (Band 4). Images are organized by column, e.g. (a) and (e) are natural color and stretched color, respectively, for date 2018-10-14.

In order to detect the yellowing associated with southern beech flowering, we apply two approaches: the calculation of a normalized difference yellowing index (NDYI) to describe the red/green band relationship; and the modelling of this index over the time-series of the image archive to detect variations from the expected (non-flowering) state. The effect is subtle enough that the technique required is very sensitive to cloud and shadow contamination. Additional to the cloud masking performed above, 'invalid' pixels from the cloud mask (cloud/shadow/snow/water) were buffered by 30 pixels (300 m) and patches of 'valid' pixels smaller than 100 ha were re-classified 'invalid'. Extra spectral filtering was then used to mask remaining pixels too bright or dark to be forest (or useful):  $B4_{red} < 650$ ,  $B3_{green} < 900$ ,  $B2_{blue} < 1000$ ,  $B8_{NIR} > 1000$ ,  $B4_{red} - B5_{rededge} < -1500$ . The result was then further buffered by 3 pixels.

The NDYI is similar to the well-known normalized difference vegetation index (NDVI)[32], using the 'red' (Band 4 665 +/- 15 nm) and 'green' (Band 3 560 +/- 18 nm) Sentinel-2 bands instead of 'near-infrared' and 'red'. It is also very similar to the green-red vegetation index (GRVI)[33,34], with the order of the bands merely reversed:

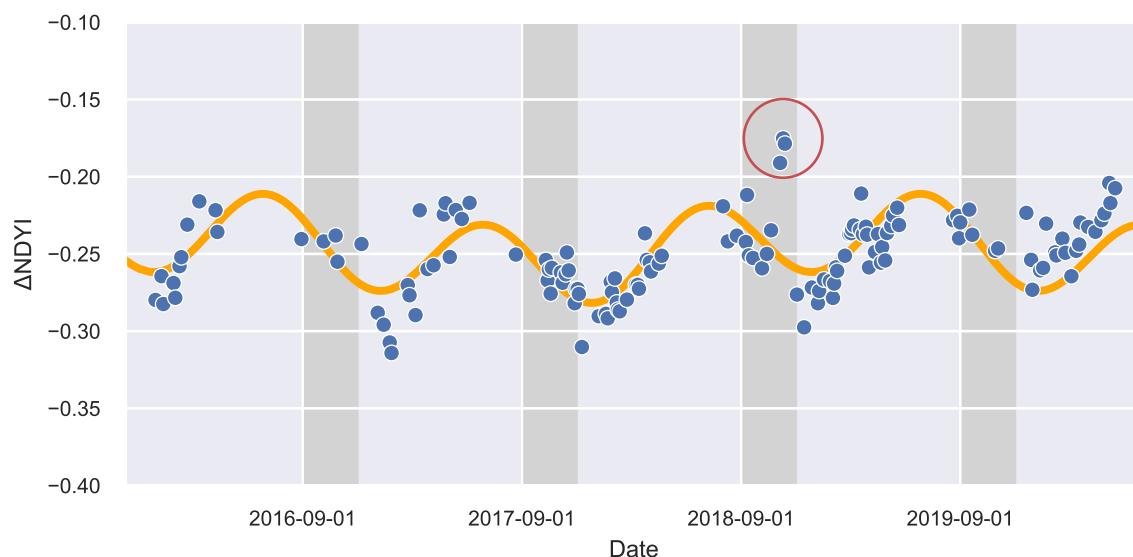
$$NDYI = \frac{(B4_{red} - B3_{green})}{(B4_{red} + B3_{green})} \quad (1)$$

The NDYI is calculated using the Level-1C TOA reflectance product re-projected to the New Zealand Transverse Mercator (NZTM) coordinate reference system (EPSG:2193) and with invalid pixels masked. As the NDYI represents the ratio of red-green, and both are affected similarly by transient atmospheric conditions, it will be negative over most areas of forest most of the time (i.e. when a pixel is 'green'), increasing to near or slightly above 0 during heavy flowering events.

Substantial annual variation is present in TOA reflectance observations over forest due to climatic conditions, vegetation phenology (e.g. new leaves), and sun angle. The NDYI signal visually observed during flowering is subtle enough in the context of a year of data that setting simple thresholds is inadequate to produce a reliable result. Additionally, flowering can occur at different times during the spring season, depending on latitude and altitude [3,8]. The temporal sequence of NDYI for a pixel in the Hawdon Valley (South Island, New Zealand) is shown in Figure 2 as an example. The orange

line is a modelled NDYI time series unique to that pixel, using an approach similar to that used in the TMASK methodology developed for cloud detection in Landsat 8 imagery [35]. The model uses robust regression to calculate unique per-pixel ( $i, j$ ) coefficients for sine and cosine terms for intra- ( $a_{1,i,j}, a_{2,i,j}$ ) and inter-annual ( $a_{3,i,j}, a_{4,i,j}$ ) variability as well as a constant term ( $c_{i,j}$ ), where  $x$  is the number of days since the start of the temporal sequence,  $T_{yr}$  is the number of days per year, and  $T_{all}$  is the number of days in the sequence:

$$\text{NDYI}_{mod}(i, j, x) = c_{i,j} + a_{1,i,j}\sin(2\pi \frac{x}{T_{yr}}) + a_{2,i,j}\cos(2\pi \frac{x}{T_{yr}}) + a_{3,i,j}\sin(2\pi \frac{x}{T_{all}}) + a_{4,i,j}\cos(2\pi \frac{x}{T_{all}}) \quad (2)$$



**Figure 2.** Observed NDYI time series for an example pixel from Figure 1 in the Hawdon Valley (blue) with modelled values (orange) superimposed. Grey areas indicate austral spring seasons (September - November) where NDYI is expected to peak during a mast. Red circle shows higher-than-expected NDYI during spring 2018, the 'mega mast' season.

Observed NDYI for each pixel/date were then subtracted from the modelled value for that pixel/date to produce  $\Delta\text{NDYI}$  and the maximum value was found for each flowering season (September 1st to December 10th):

$$\Delta\text{NDYI} = \text{NDYI} - \text{NDYI}_{mod} \quad (3)$$

Finally,  $\Delta\text{NDYI}$  values were converted to a map of 'heavy flowering detected' vs 'heavy flowering not detected' by following a method similar to Shepherd *et al.* [36]. First, a high  $\Delta\text{NDYI}$  threshold of 0.08 was chosen by assessing  $\Delta\text{NDYI}$  against seed trap data [4] during the 2018 'mega mast'. This threshold was used to create 'seed' areas which were grown outwards by progressively lowering it to 0.04. The resulting 'flowering' pixels were then buffered by 2 pixels followed by a  $5 \times 5$  majority filter then eroded by 2 pixels. 'Heavy flowering not detected' patches smaller than 1 ha (minimum mapping unit) were removed by re-coding as 'heavy flowering detected' or 'no data' (majority of surrounding pixels), then 'heavy flowering detected' patches smaller than 1 ha were re-coded 'heavy flowering not detected' to reduce small-scale noise.

As no reliable spatial dataset of beech flowering exists beyond occasional field reports from DOC staff, the national-scale map for the 2018 mast year was accuracy-assessed by a human operator. At 1000 randomly selected sites, 500 in "heavy flowering detected" and 500 in "heavy flowering not

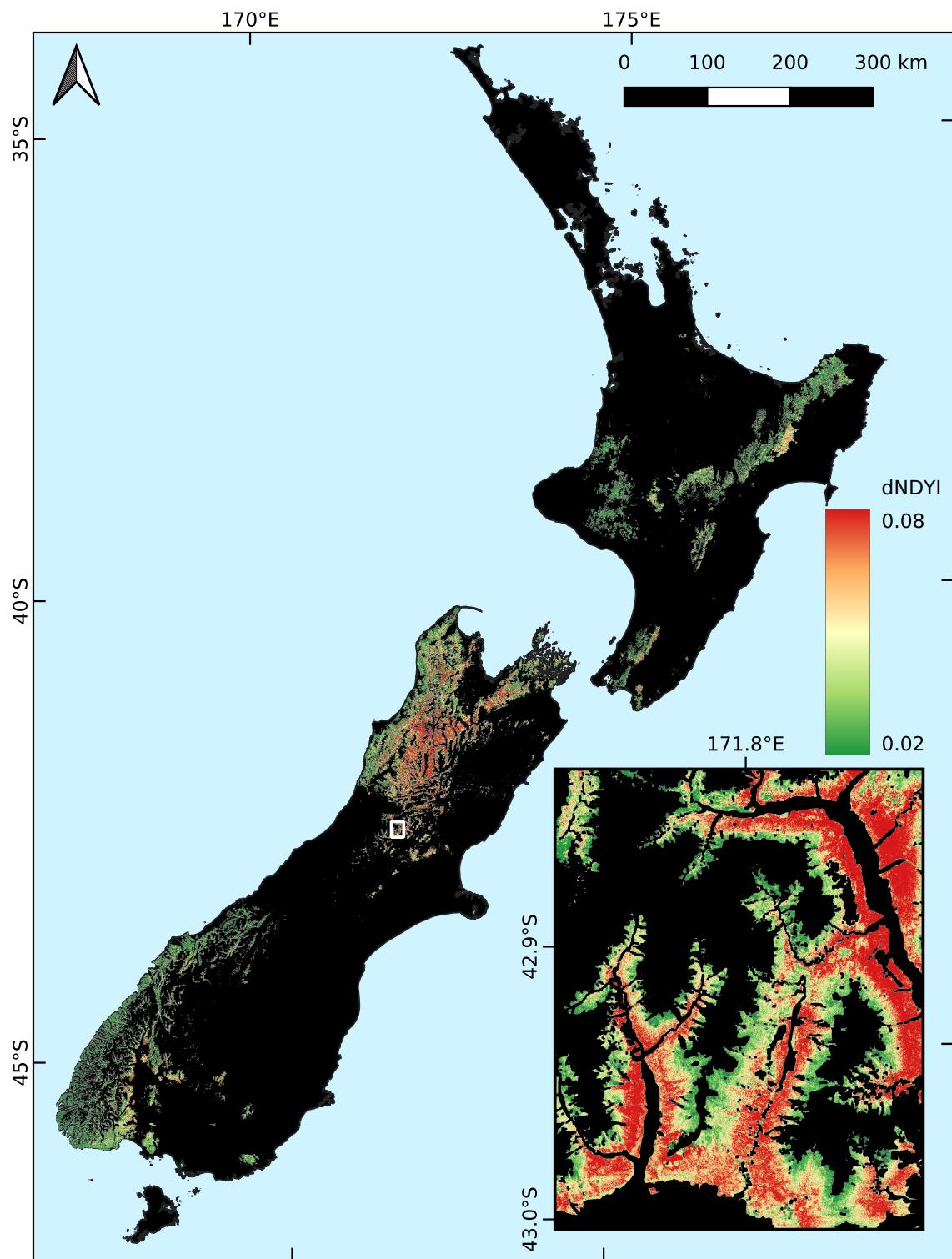
detected", the operator determined whether heavy flowering was observed in a temporal sequence of 2018 cloud-free imagery in comparison with a median spring image (excluding 2018). Heavy flowering was easily observed in the temporal progression of spectral reflectance relative to the median image, especially when the spatial extent of flowering moved upward in elevation as the season progressed (using the exaggerated Red band stretch shown in Figure 1). A confusion matrix of proportions was estimated using the method of Card [37], from which precision and recall were calculated [38].

### 3. Results

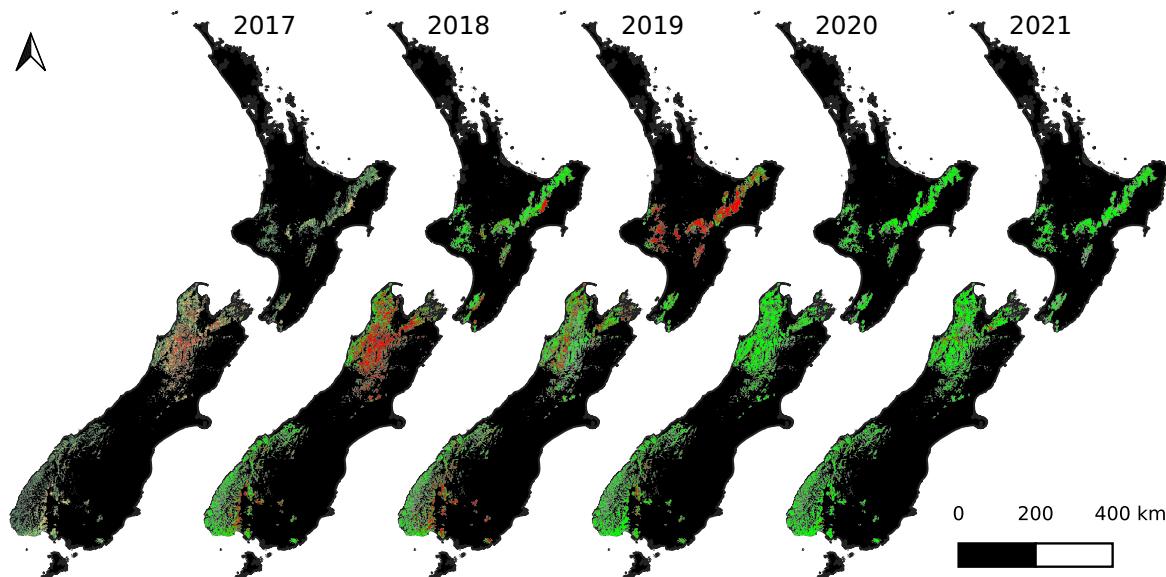
Austral spring (September/October/November) of 2017 was a light flowering season for southern beech in New Zealand. This was followed by a 'mega mast' in 2018 with heavy flowering observed from the ground during spring, and corresponding heavy seed fall the following autumn. Maps of maximum spring  $\Delta$ NDYI were produced for each year of data, with results for the 2018 season shown in Figure 3. Spatial patterns correspond with anecdotal reports from DOC staff based at field offices around New Zealand. There is heavy flowering throughout most of the north-western corner of the South Island, and sporadic heavy flowering in eastern Fiordland. The inset of Figure 3 highlights the level of detail available and shows heavy flowering on the lower slopes of the Hawdon and Poulter Valleys, dissipating as altitude increases up the valley walls (black areas are not beech forest - either alpine or riverbed in this location).

Figure 4 shows the maps of heavy beech flowering as detected by the method for years 2017 through 2021. In spring 2018, much heavy beech flowering was detected in the north-west of the South Island, synonymous with a 'mega mast' event in that region. In the North Island and the south-west of the South Island, some pockets of heavy flowering were detected. The following year in spring 2019 the flowering was much reduced in the north-west of the South Island, but in the North Island much heavy flowering was detected, synonymous with another "mega mast". The south-west of the South Island had pockets of heavy flowering, much the same as in 2018. In years 2020 and 2021, minimal beech flowering was detected in most areas.

In the 2018 map, heavy flowering was detected in 27.6% (1144382 ha) of the 4.1 million ha of beech forest in New Zealand. Heavy flowering was not detected in 51.2% (2122201 ha) of the beech forest. In the remaining 21.2% of the beech forest (878406 ha) there was no cloud-free imagery in spring to make a decision. In each of the two classes, 'heavy flowering detected' and 'heavy flowering not detected', we generated 500 random locations at which we compared reference data with map data. Reference data were determined from visual interpretation of all the cloud-free spring imagery for the year. Table 1 shows the confusion matrix of proportions. Overall classification accuracy is 90.8%. Precision (User's Accuracy) scores indicate that 90.4% of the area mapped as 'heavy flowering detected' is actually heavy flowering. Recall (Producer's Accuracy) scores indicate that 84.4% of actual 'heavy flowering detected' is successfully mapped as 'heavy flowering detected'. The overall F1 Score for 'heavy flowering detected' is 0.873, and 'heavy flowering not detected' is 0.928. Thus the  $\Delta$ NDYI method is likely to under-estimate (slightly) rather than over-estimate areas of heavy flowering, however the scores reflect well on the method overall.



**Figure 3.** Maximum  $\Delta$ NDVI (from modelled) for spring 2018 in areas of known southern beech forest in New Zealand. Green denotes areas of low ( $< 0.02$ ) maximum  $\Delta$ NDVI, while red is high ( $> 0.08$ ) and indicates heavy beech flowering. Inset shows Hawdon and Poulter Valleys near Arthur's Pass (1:250,000 at  $42.95^{\circ}$ S,  $171.82^{\circ}$ E, see white box).



**Figure 4.** Maps of heavy beech flowering during spring time for four years of Sentinel-2 imagery (2017–2021 inclusive). Classes are "heavy flowering detected" (red), "heavy flowering not detected" (green), and "no cloud-free imagery" (grey).

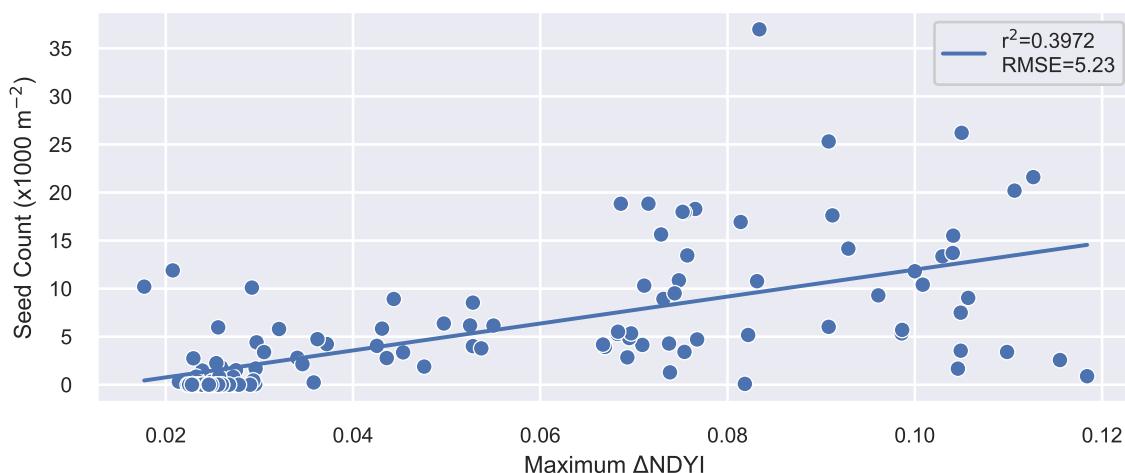
**Table 1.** Confusion matrix showing detected/not detected heavy flowering (proportion of map) for the  $\Delta$ NDYI method (Mapped) assessed against a human operator (Reference). Proportions are estimated from a random sample of 500 locations in 'heavy flowering detected' class (weighted by proportion of 'heavy flowering detected' in map = 0.35) and a random sample of 500 locations in 'heavy flowering not detected' class (weighted by 0.65).

		Reference flowering		<b>Precision</b>
		Detected	Not Det.	
<b>Mapped flowering</b>	Detected	0.316	0.034	0.904
	Not Det.	0.059	0.592	0.910
<b>Recall</b>		0.844	0.946	
<b>F1 Score</b>		0.873	0.928	

#### 207 4. Discussion

208 We developed a method that produces a national map of heavy beech flowering from a temporal  
 209 sequence of Copernicus Sentinel-2 imagery (Fig. 4). The method detects elevated values of a yellow  
 210 index, NDYI, above those normally expected in spring. A  $\Delta$ NDYI value greater than 0.08 indicates  
 211 especially heavy flowering, however these regions can be 'grown' into adjacent pixels where  $\Delta$ NDYI  
 212 is greater than 0.04 to better capture all heavy flowering. The elevated yellow index is caused by  
 213 the production of red flowers obscuring green leaves. The national map of beech flowering may be  
 214 produced at the end of spring, several months before the subsequent mast event actually occurs and  
 215 seed drops to the ground. It is now provided to DOC, the national agency in charge of pest control,  
 216 for planning additional pest control to be implemented several months later. In the 2018 spring, a  
 217 nationwide beech mast event was detected and mapped by this method. A manual accuracy assessment  
 218 determined the heavy flowering map to have an overall accuracy of 90%. The spatial distribution of  
 219 beech flowering as mapped by the method was also consistent with anecdotal observations from DOC  
 220 field staff.

221 The national map of beech flowering can be used to provide extra detail to augment the existing  
 222  $\Delta T$  model [7], as it provides a higher spatial resolution of 10 m as opposed to 5 km. It is also a map  
 223 of confirmed flowering, one less degree of separation from actual seed fall than the  $\Delta T$  model, as



**Figure 5.** Relationship between maximum  $\Delta\text{NDYI}$  (spring 2018) and number of seeds collected from seed traps in the permanent trap network (autumn/winter 2019) for the 2018/19 mast season. Locations are filtered to exclude those with fewer than eight valid satellite observations.

physical factors such as carbon availability and soil moisture conditions also affect flowering and seed productivity [39]. However, an issue with our method is the requirement for cloud-free satellite imagery at critical flowering times. This means that in some areas flowering may have been missed, which effectively makes the map a better indicator of ‘presence’ rather than ‘absence’.

Not all heavy beech flowering in spring will result in heavy seed fall in the following autumn. Heavy frost or very wet weather can interfere with seed production [3]. Figure 5 shows how well the maximum  $\Delta\text{NDYI}$  compares with seed counts in trays located on the floor of beech forest (seed traps are spread throughout beech forests in New Zealand as part of a long-running monitoring programme conducted by DOC [4]). Data were restricted to locations with at least eight valid observations to obtain representation over the majority of the spring season. There is noise in the data, nevertheless high seed counts generally correspond with high maximum  $\Delta\text{NDYI}$  ( $r^2 = 0.397$ ). For reference, the  $\Delta T$  model had  $r^2$  values between 0.331 and 0.556 for the same species range [7] (that study used the older genus name *Nothofagus*). Reasons for mismatches between  $\Delta\text{NDYI}$  and seed count include: cloud coverage still obscuring the flowering event despite high revisits (low  $\Delta\text{NDYI}$  vs high seed count); inaccurate trap location (variable impact on relationship); trap location relative to flowering trees combined with wind direction during seed fall (high  $\Delta\text{NDYI}$  vs low seed count); different beech species (different relationship between  $\Delta\text{NDYI}$  and seed count); climate, adverse weather events, and nutrient availability (lower seed count vs higher  $\Delta\text{NDYI}$ ); and inaccuracies in the method (addressed in accuracy assessment). We recommend the national map of flowering/not flowering be regarded as a map of potential high seed fall for initial planning purposes, to be confirmed later with additional information such as selected field observations.

One way to address the paucity of valid observations is to add more data sources. As the technique developed in this study relies only on red, green, blue, and NIR (for quality control) wavelengths, it should be possible to include data from commercial satellite constellations with higher revisit rates but lower spectral range or resolution, such as the Planet<sup>1</sup> ‘Dove’ constellation. Adding freely available Landsat-8 data could also increase the probability of obtaining a valid observation at a critical time. Targeted aerial imaging campaigns could also provide valuable information in areas of known data paucity, particularly if they were informed by observations from field staff. This study has

<sup>1</sup> <https://www.planet.com/>

252 shown the resolution requirement is low by aerial imaging standards which would allow higher flight  
253 altitudes and larger image footprints, substantially reducing cost. Multiple studies have shown that  
254 fusion of these separate data sources is useful in remote sensing [11,14,16,20,40,41], though the spatial  
255 complexity and rugged terrain of the beech forests in New Zealand is likely to reduce the utility of the  
256 coarse-resolution MODIS optical imagery.

257 This study successfully mapped the presence of heavy flowering in beech trees at large scale  
258 (greater than 4 million ha) using a visible change in canopy color. A similar study by Garcia *et al.*  
259 [21] was less successful, but did show that moisture-based indices in the lead-up to a flowering or  
260 seed/cone event could provide additional information. Fernández-Martínez *et al.* [19] were also  
261 successful in predicting mast events using a combination of the enhanced vegetation index (EVI), and  
262 weather data during spring. A number of challenges exist in the context of detecting mast events,  
263 and the  $\Delta$ NDYI approach attempts to minimize these. The NDYI index was chosen to specifically  
264 target red and green image bands, avoiding red-edge and near-infrared bands that also respond to  
265 vegetation condition and thus increase noise. The effectiveness of the multi-year sine and cosine model  
266 for modelling the typical behaviour of NDYI, the utilisation of extreme  $\Delta$ NDYI values as 'seeds' for  
267 regions that grew into areas of lower  $\Delta$ NDYI values, and the ability to tune spectral value constraints  
268 have all contributed to the effectiveness of our approach. To further improve the performance of the  
269  $\Delta$ NDYI method, it would be worth investigating the use of supporting indices like Garcia *et al.* [21] and  
270 Fernández-Martínez *et al.* [19], in addition to adding extra data sources. Further work distinguishing  
271 different beech species would add greater value to DOC as those with larger seeds (red and hard  
272 beech) have a disproportionately-larger impact on rodent irruptions.

273 Temporal analysis of Sentinel-2 satellite imagery has proved successful at detecting heavy  
274 flowering in New Zealand beech forests. To achieve this, cloud clearing had to be accurate (because  
275 the yellow index is sensitive to missed cloud) and automated (because many images are required).  
276 Automation of the cloud clearing [31] and other processing means that beech flowering maps can be  
277 produced in a timely and cost-effective way. In future, we plan to produce a national map of heavy  
278 beech flowering at the end of each spring. This would give several months for analysis to plan the  
279 extra pest control required in autumn, improving the targeting of pest control in masting areas, and  
280 leading to better outcomes for native fauna.

## 281 5. Conclusions

282 This study used Sentinel-2 top-of-atmosphere (TOA) imagery to detect and map atypical yellowing  
283 associated with heavy flowering of southern beech (*Fuscospora* and *Lophozonia*) in New Zealand over 4.1  
284 million ha at an unprecedented 10 m spatial resolution. This was achieved by modelling a normalized  
285 difference yellowing index (NDYI) over 5 years of observations and investigating deviations from  
286 expected values during spring months (September–November). A 'threshold'  $\Delta$ NDYI value of 0.08  
287 may be used to identify areas of heavy flowering, with connected areas of  $\Delta$ NDYI > 0.04 also likely  
288 flowering. The method has been automated and can be run for all of New Zealand in less than a  
289 day on a cluster of approximately 1000 CPU cores. Using Sentinel-2 imagery, the method typically  
290 provides information on heavy flowering for 80% of the beech forests in New Zealand with a high  
291 overall classification accuracy of 90.8%, producing useful information for planning national-scale pest  
292 control efforts.

293 **Author Contributions:** Conceptualization, Ben Jolly, John Dymond, Terry Greene, and Jan Schindler; Data  
294 curation, Ben Jolly, James Shepherd and Jan Schindler; Formal analysis, Ben Jolly; Funding acquisition, John  
295 Dymond; Investigation, Ben Jolly, John Dymond, and Terry Greene; Methodology, Ben Jolly, John Dymond, James  
296 Shepherd, and Terry Greene; Project administration, John Dymond; Resources, John Dymond; Software, Ben Jolly,  
297 James Shepherd and Jan Schindler; Supervision, John Dymond; Visualization, Ben Jolly; Writing – original draft,  
298 Ben Jolly; Writing – review & editing, John Dymond, James Shepherd, and Terry Greene.

299 **Funding:** This research was funded by the New Zealand Ministry of Business, Innovation and Employment  
300 (MBIE) Endeavour Fund under the Advanced Remote Sensing of Aotearoa research programme (C09X1709).

**Acknowledgments:** The authors would like to thank the New Zealand Department of Conservation (DOC) for providing advice and validation data.

**Conflicts of Interest:** The authors declare no conflict of interest. The funders had no role in the design of the study; in the collection, analyses, or interpretation of data; in the writing of the manuscript, or in the decision to publish the results.

### Abbreviations

The following abbreviations are used in this manuscript:

DOC	Department of Conservation
ESA	European Space Agency
EVI	Enhanced vegetation index
GRVI	Green-red vegetation index
NDVI	Normalized difference vegetation index
NDYI	Normalized difference yellowing index
NZTM	New Zealand Transverse Mercator
TOA	Top of atmosphere

310

- Heenan, P.B.; Smissen, R.D. Revised circumscription of *Nothofagus* and recognition of the segregate genera *Fuscospora*, *Lophozonia*, and *Trisyngyne* (*Nothofagaceae*). *Phytotaxa* **2013**, *146*, 1–31. doi:10.11646/phytotaxa.146.1.1.
- Shepherd, J.D.; Ausseil, A.G.; Dymond, J.R. *EcoSat Forests: a 1: 750,000 scale map of indigenous forest classes in New Zealand*; Manaaki Whenua Press, 2005.
- Wardle, J.A. *The New Zealand beeches. Ecology, utilisation and management*; New Zealand Forest Service: Wellington, New Zealand, 1984; p. 477.
- Elliott, G.; Kemp, J. Large-scale pest control in New Zealand beech forests. *Ecological Management and Restoration* **2016**, *17*, 200–209. doi:10.1111/emr.12227.
- Ruscoe, W.A.; Pech, R.P. Rodent Outbreaks: Ecology and Impacts. In *Rodent outbreaks: ecology and impacts*; Singleton, G.; Belmain, S.; Brown, P.; Hardy, B., Eds.; International Rice Research Institute: Los Baños (Philippines), 2010; pp. 239–251.
- King, C.M.; Powell, R.A. Managing an invasive predator pre-adapted to a pulsed resource: A model of stoat (*Mustela erminea*) irruptions in New Zealand beech forests. *Biological Invasions* **2011**, *13*, 3039–3055. doi:10.1007/s10530-011-9993-y.
- Kelly, D.; Geldenhuis, A.; James, A.; Penelope Holland, E.; Plank, M.J.; Brockie, R.E.; Cowan, P.E.; Harper, G.A.; Lee, W.G.; Maitland, M.J.; Mark, A.F.; Mills, J.A.; Wilson, P.R.; Byrom, A.E. Of mast and mean: Differential-temperature cue makes mast seeding insensitive to climate change. *Ecology Letters* **2013**, *16*, 90–98. doi:10.1111/ele.12020.
- Allen, R.B.; Hurst, J.M.; Portier, J.; Richardson, S.J. Elevation-dependent responses of tree mast seeding to climate change over 45 years. *Ecology and Evolution* **2014**, *4*, 3525–3537. doi:10.1002/ece3.1210.
- Bogdziewicz, M.; Kelly, D.; Thomas, P.A.; Lageard, J.G.; Hacket-Pain, A. Climate warming disrupts mast seeding and its fitness benefits in European beech. *Nature Plants* **2020**, *6*, 88–94. doi:10.1038/s41477-020-0592-8.
- Bogdziewicz, M.; Hacket-Pain, A.; Kelly, D.; Thomas, P.A.; Lageard, J.; Tanentzap, A.J. Climate warming causes mast seeding to break down by reducing sensitivity to weather cues. *Global Change Biology* **2021**, p. gcb.15560. doi:10.1111/gcb.15560.
- Bolton, D.K.; Gray, J.M.; Melaas, E.K.; Moon, M.; Eklundh, L.; Friedl, M.A. Continental-scale land surface phenology from harmonized Landsat 8 and Sentinel-2 imagery. *Remote Sensing of Environment* **2020**, *240*. doi:10.1016/j.rse.2020.111685.
- Misra, G.; Cawkwell, F.; Wingler, A. Status of phenological research using sentinel-2 data: A review. *Remote Sensing* **2020**, *12*. doi:10.3390/RS12172760.

- 343 13. Zeng, L.; Wardlow, B.D.; Xiang, D.; Hu, S.; Li, D. Remote Sensing of Environment A review of vegetation  
344 phenological metrics extraction using time-series , multispectral satellite data. *Remote Sensing of Environment*  
345 2020, 237, 111511. doi:10.1016/j.rse.2019.111511.
- 346 14. Dixon, D.J.; Callow, J.N.; Duncan, J.M.; Setterfield, S.A.; Pauli, N. Satellite prediction of forest flowering  
347 phenology. *Remote Sensing of Environment* 2021, 255, 112197. doi:10.1016/j.rse.2020.112197.
- 348 15. Moon, M.; Seyednasrollah, B.; Richardson, A.D.; Friedl, M.A. Using time series of MODIS land surface  
349 phenology to model temperature and photoperiod controls on spring greenup in North American  
350 deciduous forests. *Remote Sensing of Environment* 2021, 260. doi:10.1016/j.rse.2021.112466.
- 351 16. Moon, M.; Richardson, A.D.; Friedl, M.A. Multiscale assessment of land surface phenology from  
352 harmonized Landsat 8 and Sentinel-2, PlanetScope, and PhenoCam imagery. *Remote Sensing of Environment*  
353 2021, 266. doi:10.1016/j.rse.2021.112716.
- 354 17. Browning, D.M.; Russell, E.S.; Ponce-Campos, G.E.; Kaplan, N.; Richardson, A.D.; Seyednasrollah, B.;  
355 Spiegal, S.; Saliendra, N.; Alfieri, J.G.; Baker, J.; Bernacchi, C.; Bestelmeyer, B.T.; Bosch, D.; Boughton, E.H.;  
356 Boughton, R.K.; Clark, P.; Flerchinger, G.; Gomez-Casanovas, N.; Goslee, S.; Haddad, N.M.; Hoover, D.;  
357 Jaradat, A.; Mauritz, M.; McCarty, G.W.; Miller, G.R.; Sadler, J.; Saha, A.; Scott, R.L.; Suyker, A.; Tweedie, C.;  
358 Wood, J.D.; Zhang, X.; Taylor, S.D. Monitoring agroecosystem productivity and phenology at a national  
359 scale: A metric assessment framework. *Ecological Indicators* 2021, 131. doi:10.1016/j.ecolind.2021.108147.
- 360 18. Atkinson, P.M.; Jeganathan, C.; Dash, J.; Atzberger, C. Inter-comparison of four models for smoothing  
361 satellite sensor time-series data to estimate vegetation phenology. *Remote Sensing of Environment* 2012,  
362 123, 400–417. doi:10.1016/j.rse.2012.04.001.
- 363 19. Fernández-Martínez, M.; Garbulsky, M.; Peñuelas, J.; Peguero, G.; Espelta, J.M. Temporal trends in the  
364 enhanced vegetation index and spring weather predict seed production in Mediterranean oaks. *Plant  
365 Ecology* 2015, 216, 1061–1072. doi:10.1007/sl.
- 366 20. Cheng, Y.; Vrieling, A.; Fava, F.; Meroni, M.; Marshall, M.; Gachoki, S. Phenology of short vegetation  
367 cycles in a Kenyan rangeland from PlanetScope and Sentinel-2. *Remote Sensing of Environment* 2020, 248.  
368 doi:10.1016/j.rse.2020.112004.
- 369 21. Garcia, M.; Zuckerberg, B.; Lamontagne, J.M.; Townsend, P.A. Landsat-based detection of mast  
370 events in white spruce (*Picea glauca*) forests. *Remote Sensing of Environment* 2021, 254, 112278.  
371 doi:10.1016/j.rse.2020.112278.
- 372 22. Noumonvi, K.D.; Oblišar, G.; Žust, A.; Vilhar, U. Empirical approach for modelling tree phenology in  
373 mixed forests using remote sensing. *Remote Sensing* 2021, 13. doi:10.3390/rs13153015.
- 374 23. Asokan, A.; Anitha, J. Change detection techniques for remote sensing applications: a survey. *Earth Science  
375 Informatics* 2019. doi:10.1007/s12145-019-00380-5.
- 376 24. Panuju, D.R.; Paull, D.J.; Griffin, A.L. Change detection techniques based on multispectral images for  
377 investigating land cover dynamics. *Remote Sensing* 2020, 12. doi:10.3390/rs12111781.
- 378 25. Chen, B.; Jin, Y.; Brown, P. An enhanced bloom index for quantifying floral phenology using multi-scale  
379 remote sensing observations. *ISPRS Journal of Photogrammetry and Remote Sensing* 2019, 156, 108–120.  
380 doi:10.1016/j.isprsjprs.2019.08.006.
- 381 26. Dymond, J.R.; Shepherd, J.D. The spatial distribution of indigenous forest and its composition in the  
382 Wellington region, New Zealand, from ETM+ satellite imagery. *Remote Sensing of Environment* 2004,  
383 90, 116–125. doi:10.1016/j.rse.2003.11.013.
- 384 27. Landcare Research Ltd. EcoSat Forests (North Island), 2014. doi:10.26060/STZ2-Z482.
- 385 28. Landcare Research Ltd. EcoSat Forest (South Island), 2014. doi:10.26060/2EPD-WE90.
- 386 29. Department of Conservation. Flowering and fruit production, 2019.
- 387 30. Department of Conservation. Department of Conservation Te Papa Atawhai Annual Report 2019. Technical  
388 report, Department of Conservation, 2019.
- 389 31. Shepherd, J.D.; Schindler, J.; Dymond, J.R. Automated Mosaicking of Sentinel-2 Satellite Imagery. *Remote  
390 Sensing* 2020, 12, 3680. doi:10.3390/rs12223680.
- 391 32. Rouse, W.; Haas, H.; Deering, W. Monitoring vegetation systems in the Great Plains with ERTS. NASA.  
392 Goddard Space Flight Center 3d ERTS-1 Symposium, 1974.
- 393 33. Tucker, C.J. Red and Photographic Infrared 1 , linear Combinations for Monitoring Vegetation. *Remote  
394 Sensing of Environment* 1979, 8, 127–150. doi:10.1016/0034-4257(79)90013-0.

- 395 34. Motohka, T.; Nasahara, K.N.; Oguma, H.; Tsuchida, S. Applicability of Green-Red Vegetation Index for  
396 remote sensing of vegetation phenology. *Remote Sensing* **2010**, *2*, 2369–2387. doi:10.3390/rs2102369.
- 397 35. Zhu, Z.; Woodcock, C.E. Automated cloud, cloud shadow, and snow detection in multitemporal Landsat  
398 data: An algorithm designed specifically for monitoring land cover change. *Remote Sensing of Environment*  
399 **2014**, *152*, 217–234. doi:10.1016/j.rse.2014.06.012.
- 400 36. Shepherd, J.D.; Dymond, J.R.; Cuff, J.R. Monitoring scrub weed change in the Canterbury region using  
401 satellite imagery. *New Zealand Plant Protection* **2007**, *60*, 137–140. doi:10.30843/nzpp.2007.60.4671.
- 402 37. Card, D.H. Using Known Map Category Marginal Frequencies to Improve Estimates of Thematic Map  
403 Accuracy. *Photogrammetric Engineering and Remote Sensing* **1982**, *48*, 431–439.
- 404 38. Maxwell, A.E.; Warner, T.A.; Guillén, L.A. Accuracy assessment in convolutional neural network-based  
405 deep learning remote sensing studies—part 1: Literature review. *Remote Sensing* **2021**, *13*.  
406 doi:10.3390/rs13132450.
- 407 39. Uscoe, W.E.A.R.; Latt, K.E.H.P.; Richardson, S.J.; Allen, R.B.; Whitehead, D.; Carswell, F.E.; Ruscoe,  
408 W.A.; Platt, K.H. Climate and net carbon availability determine temporal patterns of seed production by  
409 *Nothofagus*. *Ecology* **2005**, *86*, 972–981.
- 410 40. Thapa, S.; Millan, V.E.G.; Eklundh, L. Assessing forest phenology: A multi-scale comparison of near-surface  
411 (UAV, spectral reflectance sensor, phenocam) and satellite (MODIS, sentinel-2) remote sensing. *Remote  
412 Sensing* **2021**, *13*. doi:10.3390/rs13081597.
- 413 41. Peng, D.; Wang, Y.; Xian, G.; Huete, A.R.; Huang, W.; Shen, M.; Wang, F.; Yu, L.; Liu, L.; Xie, Q.; Liu, L.;  
414 Zhang, X. Investigation of land surface phenology detections in shrublands using multiple scale satellite  
415 data. *Remote Sensing of Environment* **2021**, *252*. doi:10.1016/j.rse.2020.112133.

416 © 2022 by the authors. Submitted to *Remote Sens.* for possible open access publication  
417 under the terms and conditions of the Creative Commons Attribution (CC BY) license  
418 (<http://creativecommons.org/licenses/by/4.0/>).