

Article

Detection of southern beech heavy flowering using Sentinel-2 imagery

Ben Jolly¹*, John R. Dymond¹, James D. Shepherd¹, Terry Greene², and Jan Schindler¹

¹ Manaaki Whenua – Landcare Research;

² 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 masting areas. To plan pest control and keep it cost-effective, it would be helpful to have a map of the masting 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*). These 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. Rodents and mustelids are of particular concern [4] because 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 non-native and prey on New Zealand introduced to New Zealand and also prey on native bird species, so increases in their populations are concerning. This compounds when the When seed begins to run out and alternative, increasing pressure is put on native biota as additional food sources are sought as populations become unsustainable. Population control of introduced predators is essential for preserving populations of native and endemic species of birds, reptiles, and invertebrates in New Zealand, especially in during beech mast years [6].

The New Zealand Department of Conservation (DOC) is responsible for managing forests on public land, as well as preserving native species, and coordinating pest control. They need to know when and where Prediction of the extent and intensity of significant mast events occur for targeting 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 targeting planning management interventions: (i) modelling, (ii) field observations, and (iii) sampling of tree branches *in situ* sampling of developing seed crops in tree canopies. (i) The 'delta-T' (ΔT) model [7] uses the difference in mean temperature from the previous two summers ($T_{n-1} - T_{n-2}$) to predict likely seedfall seed fall for the following autumn. This has been used to predict high seed yields at national scale. However, it relies on temperature data, which are currently only available on a modelled 5×5 km grid, so will miss smaller-scale micro-climate effects. While historical temperature is an important factor in synchronizing mast events [7], other factors such as nutrient availability also play a role [8]. New research suggests that rising temperatures due to climate change may change the masting alter the mast cycle of beech trees [8,9], increasing the spatial and temporal complexity of masting mast patterns [8] and potentially desynchronizing de-synchronizing flowering/seeding, effectively reducing the impact, and predictive power, of temperature on the timing of the reproductive cycle [10]. (ii) Field staff from DOC are well-placed to provide observations of beech flowering in certain areas as part of their normal duties. However, there are large areas of forest that remain unobserved and spatial extent is often difficult to define, especially at a regional or national scale. (iii) Finally, extensive Extensive sampling campaigns are conducted during years where a heavy mast is expected using helicopters to clip upper branches from trees so seeds can be counted. This task is expensive, labour-intensive, and dangerous.

Remote sensing has proved to be an effective tool for monitoring vegetation phenology, particularly when a rich time series of imagery is available [11][11–16]. With sensors like Landsat 8 and Sentinel-2 it is possible to map phenology over millions of hectares to create national maps at detailed scale (10m pixels) in great detail [12,17]. Phenological characteristics are usually derived by first fitting a curve to the time series of remote sensing data, then using either threshold-based methods, moving averages, inflection point, or time of maximum increase [18][18–22]. Seasonality is usually assumed [18]. Vegetation indices As mast events are considered a deviation from the 'median' annual cycle, it should be possible to use change detection techniques [23,24] to identify or even predict mast seasons if identifying features are visible from above [14,20,22].

Vegetation indices, such as the normalized difference vegetation index (NDVI) and the enhanced vegetation index (EVI), and the normalized difference yellow index (NDYI), are often used to differentiate flowers and to summarise or seed pods/cones and to summarize data as one variable for analysis [14,18]. [14,18,20,25]. Often, multiple indices are combined in an effort to investigate multiple physical properties. For example, a study by Garcia *et al.* [22] investigating White Spruce mast events found vegetation indices targeting moisture were more effective than the traditional color-based ones but still struggled to reliably predict masting. Fernández-Martínez *et al.* [20] were more successful and used increasing EVI the winter before, along with weather data during spring, as an indicator of potential mast seasons of Mediterranean oaks. Neither study were able to find a usable signature of flowering or seed/cone production to map the mast event. Dixon *et al.* [14] and Chen *et al.* [25] both successfully used multi-scale imagery to map tree-scale flowering over landscapes, Dixon *et al.* [14] by training a random forest model with drone data from known events and Chen *et al.* [25] by developing an enhanced bloom index (EBI) from drone data and successfully translating that approach to CERES, PlanetScope, Sentinel-2, and Landsat data to increase coverage.

In this study, we investigate the use of freely-available satellite remote sensing for identifying large areas of significant southern beech flowering in New Zealand. We use imagery from the European Space Agency (ESA) Sentinel-2 'a' and 'b' satellites to obtain a high rate of repeat passes to maximise maximize the chances of multiple cloud-free observations, and to produce national coverage at detailed scale. Very little ground data on flowering are available. The irregular nature of the flowering events, and ruggedness and remoteness of much of the terrain, makes planning collection of ground data for this study difficult. As significant flowering is an unusual-irregular phenological event, we identify departures of NDYI from a normal annual pattern model the phenology of beech forest per-pixel and identify departures from this image-by-image during spring seasons to identify heavy beech

85 flowering. Using this method, we produce national maps of heavy beech flowering for 5 different
86 years, 2017 through 2021, with a minimum mapping unit of 1 ha. We assess the accuracy of
87 the method by comparison with a human operator at 1000 randomly selected sites. A national map
88 of detectable heavy beech flowering can be produced at the end of every spring to assist DOC in
89 identifying potential masting mast 'hot-spots' and thus aid planning of pest control operations.

90 2. Materials and Methods

91 2.1. *Area of Interest*

92 This study covers the extent of known beech and mixed-beech forest in New Zealand according
93 to EcoSat Forests [2,26–28], covering both North and South Islands from 36.1°S, 178.0°E to 46.4°S,
94 166.4°E. Terrain is largely mountainous with slopes up to 45° as any habitat deemed suitable for
95 agriculture was cleared in the 1800's and early 1900's. Due to the large variations in both altitude
96 and latitude of the study area, mean annual temperatures range from 3.8°C to 16.6°C, and mean
97 cumulative annual rainfall from 465 mm to 9305 mm.

98 2.2. *Data*

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

108 This study uses the Copernicus Sentinel-2 Level-1C calibrated top of atmosphere (TOA)
109 reflectance at the sensor as downloaded from the ESA archives, with edges masked to remove pixels
110 that do not contain data for every band. It covers Analysis uses all available Sentinel-2 data over
111 New Zealand from 2016-09-01 to 2021-12-31, with images every 10 days before Sentinel-2b came
112 online around 2017-07-08 and images every 5 days thereafter. Analysis was restricted to areas of
113 indigenous beech forest assumed as defined by EcoSat Forests [2,26–28]. A 'mega-mast' occurred
114 during spring 2018/autumn 2019 [29,30]. Other years showed some small flowering events or none at
115 all. Cloudy pixels were excluded from the temporal sequence using the methodology outlined in [31]
116 Shepherd et al. [31].

117 No suitable ground data on beech flowering exist to provide a reference dataset for accuracy
118 assessment, however seed counts from a long-running monitoring programme using permanent seed
119 traps [4] were available and supplied by DOC.

120 2.3. *Methods*

121 Southern beech flowers produce a reddening of the forest canopy that is easily visible from the
122 ground, especially during a significant mast season. At the 10 m pixel nominal spatial resolution of
123 Sentinel-2, the reddening of the canopy appears to the human eye as a subtle yellowing in the 'natural
124 color' (RGB) image as the red flowers increase the red component of a pixel but not to the point of
125 dominating the green. Higher-wavelength image bands (and associated indices) were investigated
126 but the only other band that showed a noticeable response was do not respond to this flowering with
127 the exception of Band 5 (red edge), and however this was also associated with 'flushes' of new foliage
128 in late spring.

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

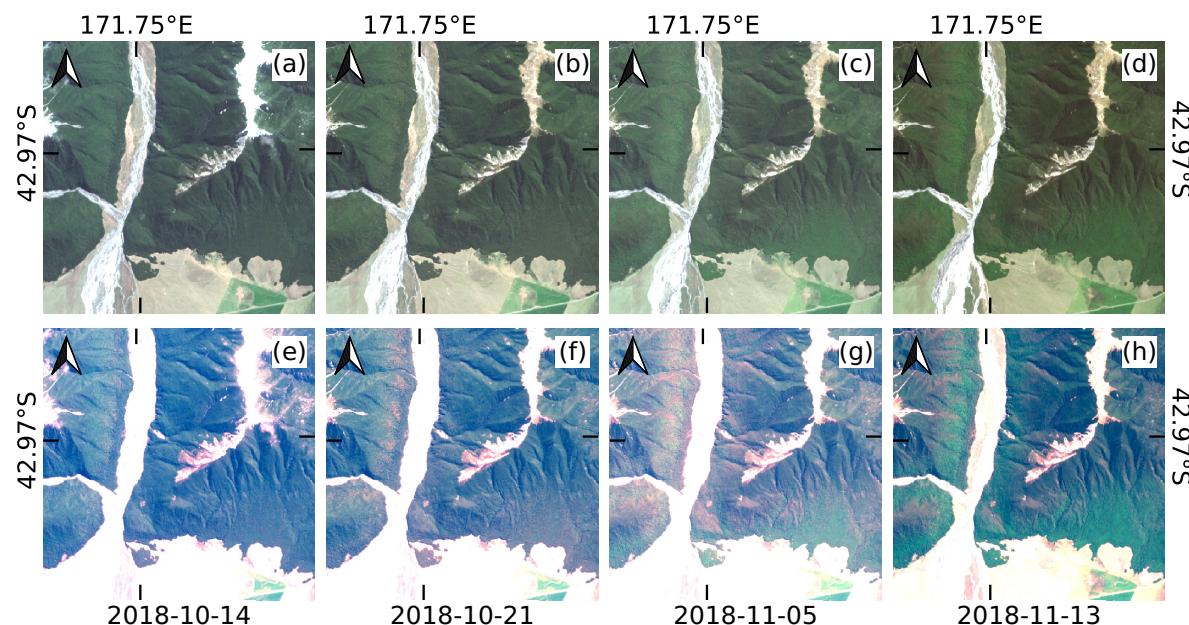


Figure 1. Sentinel-2 imagery from the Hawdon Valley, South Island, New Zealand showing a flowering event during Spring 2018. (a) - (d) are natural color, (e) - (h) are natural color with an exaggerated stretch to amplify the 'red' band (Band 4). Images are organized by column, e.g. (a) and (e) are natural 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 **normalised 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 **normalised 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. ~~NDYI was calculated for cloud-free pixels mapped as beech in EcoSat Forests [2,26–28].~~

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. ~~Part of 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] was adapted to model the yellowing index for each pixel in areas of suspected beech forest (orange line in Fig. 2).~~ The model uses robust regression to calculate unique per-pixel (i, j) coefficients for sine and cosine terms for intra- ($a_1, a_2, a_{1,i,j}, a_{2,i,j}$) and inter-annual ($a_3, a_4, a_{3,i,j}, a_{4,i,j}$) 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:

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

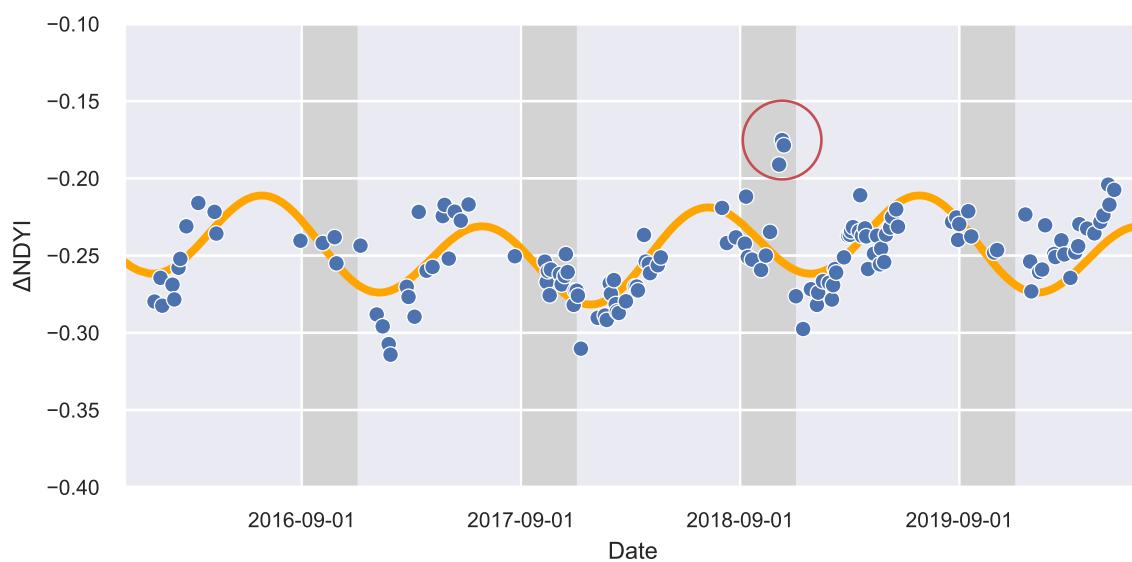


Figure 2. Observed NDYI ~~values~~ time series for an example pixel from Figure 1 in the Hawdon Valley (blue) with modelled values (orange) superimposed. Grey areas indicate ~~Spring austral~~ seasons (September - November) where NDYI is expected to peak during a mast. ~~Red circle shows higher-than-expected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

174 Observed NDYI for each pixel/date were then subtracted from the modelled value ~~to produce~~
 175 ~~ANDYI for that pixel/date to produce~~ Δ NDYI and the maximum value was found for each flowering
 176 season (~~1st September to 10th December~~ ~~September 1st to December 10th~~):

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

177 Finally, ~~ANDYI~~ Δ NDYI values were converted to a map of 'heavy flowering detected' vs 'heavy
 178 flowering not detected' by following a method similar to Shepherd *et al.* [36]. First, a high ~~ANDYI~~
 179 Δ NDYI threshold of 0.08 was chosen by assessing ~~ANDYI against the~~ Δ NDYI ~~against~~ seed trap data
 180 [4] during the 2018 'mega mast'. This threshold was used to create 'seed' areas which were grown
 181 outwards by progressively lowering it to 0.04. The resulting 'flowering' pixels were then buffered by
 182 2 pixels followed by a 5 x 5 majority filter then eroded by 2 pixels. 'Heavy flowering not detected'
 183 patches smaller than 1 ha (~~minimum mapping unit~~) were removed by re-coding as 'heavy flowering
 184 detected' or 'no data' (majority of surrounding pixels), then 'heavy flowering detected' patches smaller
 185 than 1 ha were re-coded 'heavy flowering not detected' ~~to reduce small-scale noise~~.

186 As no ~~ground truth data on flowering exist, the resulting reliable spatial dataset of beech~~
 187 ~~flowering exists beyond occasional field reports from DOC staff, the national-scale map for the 2018~~
 188 mast year was accuracy-assessed by a human operator. At 1000 randomly selected sites, 500 in "heavy
 189 flowering detected" and 500 in "heavy flowering not detected", the operator determined whether heavy
 190 flowering was observed in ~~a temporal sequence of~~ 2018 cloud-free ~~spring~~ imagery in comparison
 191 with a median spring image (excluding 2018). ~~Heavy flowering was easily observed in the temporal~~
 192 ~~progression of spectral reflectance relative to the median image, especially when the spatial extent~~
 193 ~~of flowering moved upward in elevation as the season progressed (using the exaggerated Red band~~
 194 ~~stretch shown in Figure 1).~~ A confusion matrix of proportions was estimated using the method of
 195 Card [37], from which precision and recall were calculated [38].

196 3. Results

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

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

215 In the 2018 map, heavy flowering was detected in 27.6% (~~1144382 ha~~) of the 4.1 million ha of
 216 beech forest in New Zealand. Heavy flowering was not detected in 51.2% (~~2122201 ha~~) of the beech
 217 forest. In the ~~remainder~~ ~~remaining~~ 21.2% of the beech forest (~~878406 ha~~) there was no cloud-free
 218 imagery in spring to make a decision. ~~We assessed the accuracy of the 2018 map for heavy beech~~
 219 ~~flowering.~~ In each of the two classes, "~~heavy flowering detected~~" ~~and~~ "~~and~~" and 'heavy flowering not

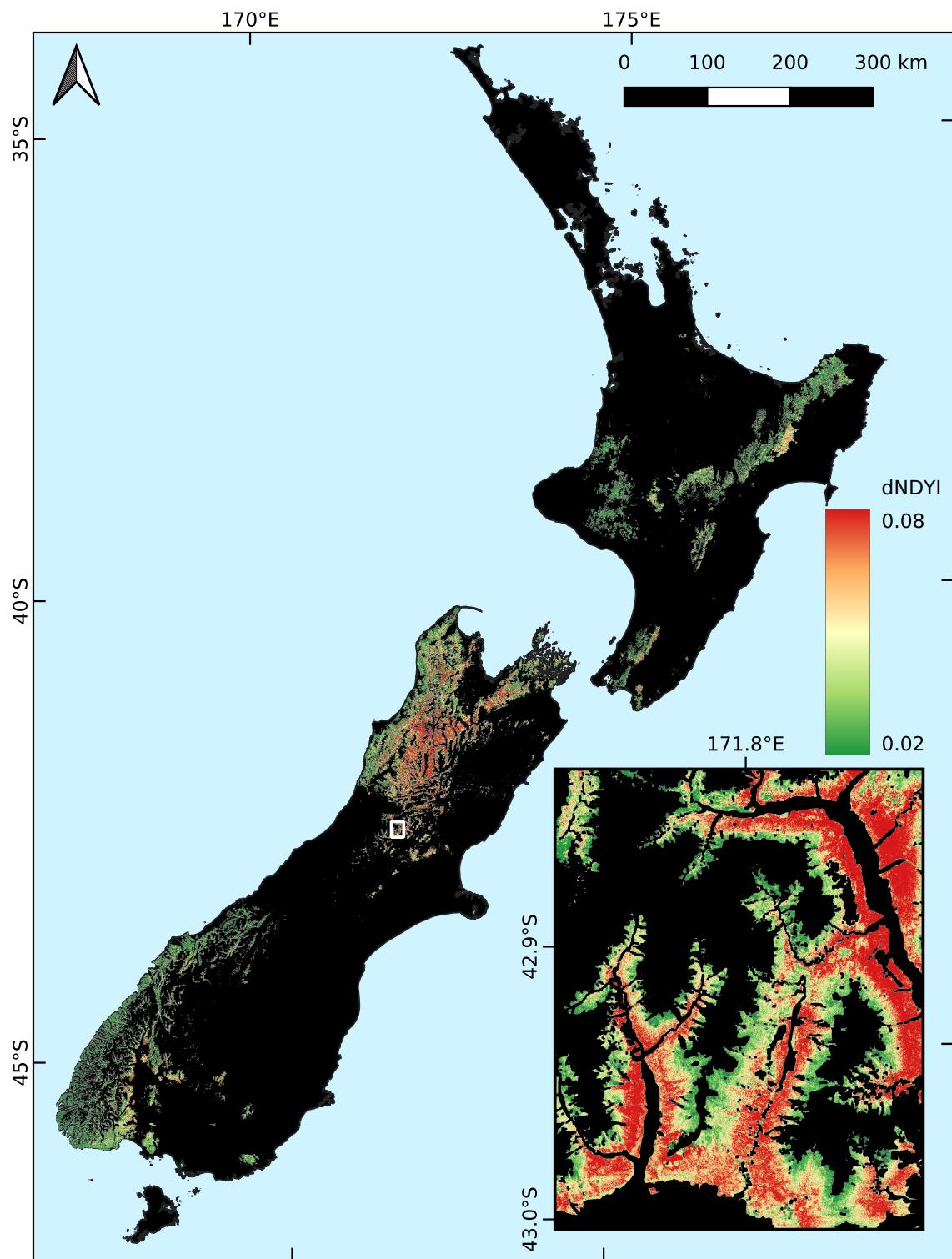


Figure 3. Maximum Δ NDYI (from modelled) for spring 2018 in areas of known southern beech forest in New Zealand. Green denotes areas of low (< 0.02) maximum Δ NDYI, while red is high (> 0.08) and indicates heavy beech flowering. Inset shows Hawdon and Poulter Valleys near Arthur's Pass (1:250,000 at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

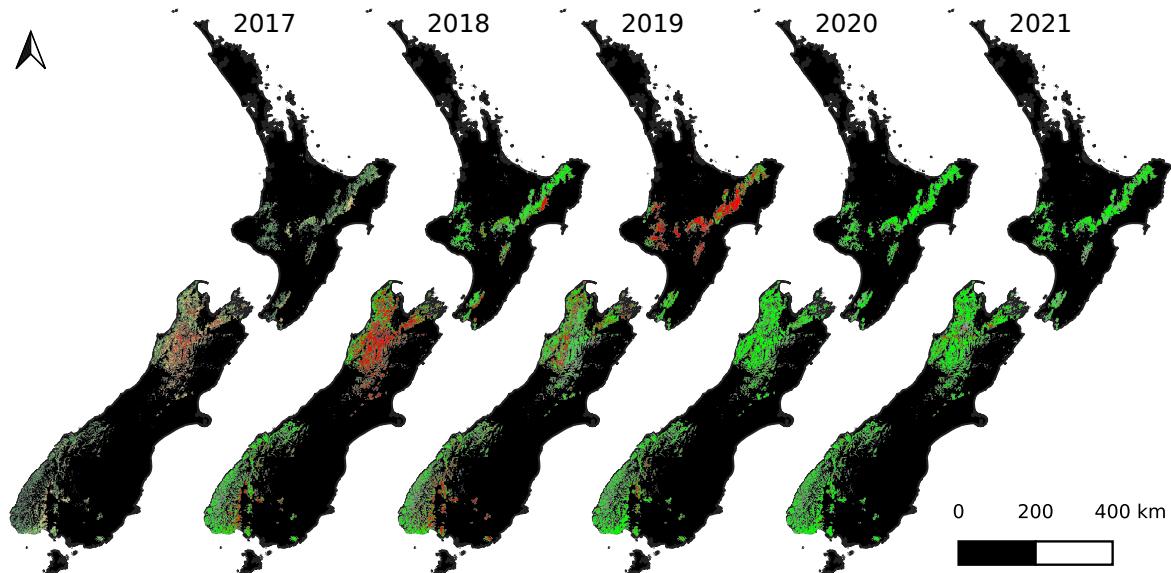


Figure 4. Maps of heavy beech flowering during spring time for four years of Sentinel-2 imagery (2017–2021 inclusive). Classes are "heavy flowering detected" (red), "heavy flowering not detected" (green), and "no cloud-free imagery" (grey).

detected^w, we generated 500 random locations at which we compared reference data with map data. Reference data ~~was~~ were determined from visual interpretation of all the cloud-free spring imagery for the year. The 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' class was 90.4% accurate and the^w 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' class was 91.0% accurate (Table 1 shows the confusion matrix) is 0.928. Thus the ANDYI method is likely to under-estimate (slightly) rather than over-estimate areas of heavy flowering, however the scores reflect well on the method overall.

Table 1. Confusion matrix of 2018 showing detected/not detected heavy flowering (proportion of map) for the ANDYI method (Mapped) assessed against a human operator (Reference). Proportions are estimated from a random sample of 500 locations in "Heavy" heavy flowering detected^w class (weighted by proportion of 'heavy flowering detected' in map = 0.35) and a random sample of 500 locations in "Heavy" heavy flowering not detected^w class (weighted by 0.65).

		Reference flowering		Precision	
		Detected	Not Det.		
Mapped flowering	Heavy flowering-detected	0.250 0.276	0.316 0.059	0.026 0.592	
	Not Det.	0.904 0.466	0.034 0.946	0.046 0.910	
		Recall	0.000 0.844	0.910	
		No-cloud-free imagery	0.000 0.873	F1 Score	0.000 0.928

Maximum ANDYI (from modelled) for spring 2018 in areas of known southern beech forest in New Zealand. Green denotes areas of low (< 200) maximum ANDYI, 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 42.95S, 171.82E, see white box).

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).

237 Relationship between maximum ANDYI (spring 2018) and number of seeds collected from
238 seed traps in the permanent trap network (autumn/winter 2019) for the 2018/19 masting season.
239 Locations are filtered to exclude those with fewer than eight valid satellite observations.

240 4. Discussion

241 We developed a method that produces a national map of heavy beech flowering from a temporal
242 sequence of Copernicus Sentinel-2 imagery (Fig. 4). The method detects elevated values of a yellow
243 index, NDYI, above those normally expected in spring—a ANDYI. A ANDYI value greater than
244 0.08 indicates heavy flowering especially heavy flowering, however these regions can be 'grown' into
245 adjacent pixels where ANDYI is greater than 0.04 to better capture all heavy flowering. The elevated
246 yellow index is caused by the production of red flowers obscuring green leaves. The national map of
247 beech flowering may be produced at the end of spring, several months before the subsequent masting
248 mast event actually occurs and seed drops to the ground. This gives It is now provided to DOC, the
249 national agency in charge of pest control, DOC, several months to analyse the spatial distribution
250 and intensity of the flowering in order to plan the extra pest control required for planning additional
251 pest control to be implemented several months later. In the 2018 spring, a nationwide beech masting
252 mast event was detected and mapped by this method. A manual accuracy assessment determined
253 the method to be heavy flowering map to have an overall accuracy of 90% accurate against a human
254 operator with the same imagery. The spatial distribution of beech flowering as mapped by the method
255 was also consistent with anecdotal observations from DOC field staff.

256 The national map of beech flowering can be used to provide extra detail to augment the existing
257 ΔT model [7], as it provides a higher spatial resolution of 10 m as opposed to 5 km. Observations
258 of flowering also help mitigate sources of uncertainty in the previous summer It is also a map of
259 confirmed flowering, one less degree of separation from actual seed fall than the ΔT that could
260 limit seed production model, as physical factors such as carbon availability and soil moisture
261 conditions [39] also affect flowering and seed productivity [39]. However, an issue with our method
262 is the requirement for cloud-free satellite imagery at critical flowering times. This means that in
263 some areas flowering may have been missed, which effectively makes the map a better indicator of
264 'presence' rather than 'absence'. The observations should also mitigate the impact of microclimatic
265 effects not captured in the modelled 5 km temperature grid. In addition to the planning of extra
266 pest control at appropriate scales, the national map of beech flowering can be used for targeting
267 observation/measurement campaigns investigating seedfall.

268 Not all heavy beech flowering in spring will result in heavy seedfall seed fall in the following
269 autumn. Heavy frost or very wet weather can interfere with seed production [3]. Figure 5 shows
270 how well the maximum ANDYI ANDYI compares with seed counts in trays located on the floor of
271 beech forest (seed traps are spread throughout beech forests in New Zealand as part of a long-running
272 monitoring programme conducted by DOC [4]). Data were restricted to locations with at least eight
273 cloud-free valid observations to obtain a fair representation over the majority of the spring season.
274 There is noise in the data, nevertheless high seed counts generally correspond with high maximum
275 ANDYI ANDYI ($r^2 = 0.397$). For reference, the ΔT model had r^2 values between 0.331 and 0.556 for
276 the same species range [7] (that study used the older genus name *Nothofagus*). Reasons for mismatches
277 between ANDYI and seed count include: cloud coverage still obscuring the flowering event (low
278 ANDYI despite high revisits (low ANDYI vs high seed count); exact trap location inaccurate trap
279 location (variable impact on relationship); trap location relative to flowering trees as well as combined
280 with wind direction during seed fall (high ANDYI ANDYI vs low seed count); different beech species
281 (different relationship between ANDYI ANDYI and seed count); climate, adverse weather events, and
282 nutrient availability (lower seed count vs higher ANDYI ANDYI); and inaccuracies in the method
283 (addressed in accuracy assessment). A similar comparison with 2017 data (not shown) showed no
284 relationship, seed counts were all 0 (or close to) and ANDYI values were very low. We recommend the
285 national map of flowering/not flowering be regarded as a map of potentially high seedfall potential

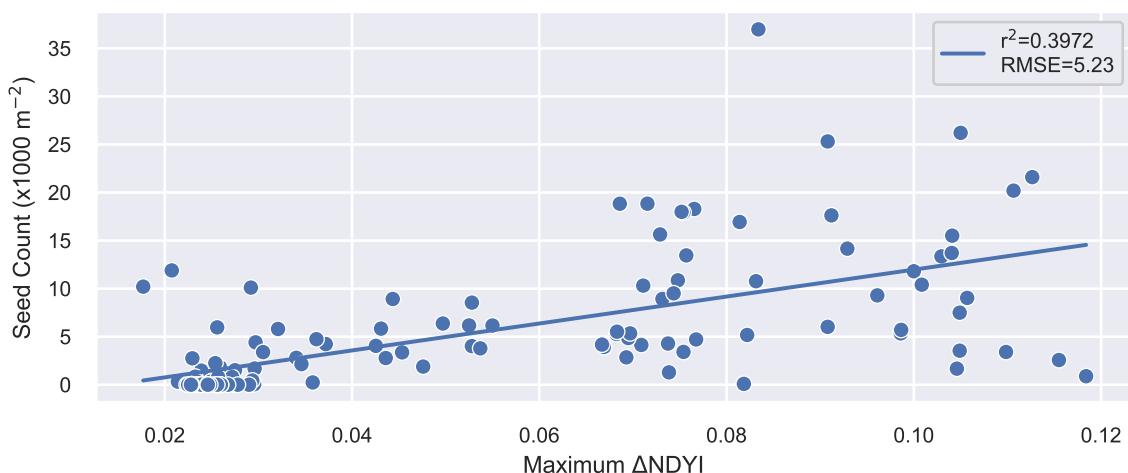


Figure 5. Relationship between maximum Δ NDYI (spring 2018) and number of seeds collected from seed traps in the permanent trap network (autumn/winter 2019) for the 2018/19 masting season. Locations are filtered to exclude those with fewer than eight valid satellite observations.

286 high seed fall for initial planning purposes, to be confirmed later with additional information such as
287 selected field observations.

288 ~~In some areas there is a paucity of satellite observations— even with 5 daily or better repeats,~~
289 ~~many mountainous areas of the country only receive a handful of cloud-free observations at irregular~~
290 ~~intervals for an entire spring. The effect of this is that some areas of mastинг are missed as the~~
291 ~~observations may not occur at times when the flowers are visible. A way to address this shortcoming~~
292 ~~is to~~ One way to address the paucity of valid observations is to add more data sources. As the
293 technique developed in this study relies only on red, green, blue, and NIR (for quality control)
294 wavelengths, it should be possible to include data from commercial satellite constellations with higher
295 revisit rates but lower spectral range or resolution, such as the Planet¹ ‘Dove’ constellation. Adding
296 freely available Landsat-8 data could also increase the probability of obtaining a valid observation
297 at a critical time. Targeted aerial imaging campaigns could also provide valuable information in
298 areas of known data paucity, particularly if they were informed by observations from field staff.
299 This study has shown the resolution requirement is low by aerial imaging standards which would
300 allow higher flight altitudes and larger image footprints, substantially reducing cost. Adding freely
301 available Landsat-8 data could also increase the probability of obtaining a valid observation at a
302 critical time. Multiple studies have shown that fusion of these separate data sources is useful in remote
303 sensing [12,14,16,21,40,41], though the spatial complexity and rugged terrain of the beech forests in
304 New Zealand is likely to reduce the utility of the coarse-resolution MODIS optical imagery.

305 This study successfully mapped the presence of heavy flowering in beech trees at large scale
306 (greater than 4 million ha) using a visible change in canopy color. A similar study by Garcia *et al.* [22]
307 was less successful, but did show that moisture-based indices in the lead-up to a flowering or
308 seed/cone event could provide additional information. Fernández-Martínez *et al.* [20] were also
309 successful in predicting mast events using a combination of the enhanced vegetation index (EVI),
310 and weather data during spring. A number of challenges exist in the context of detecting mast events,
311 and the Δ NDYI approach attempts to minimize these. The NDYI index was chosen to specifically
312 target red and green image bands, avoiding red-edge and near-infrared bands that also respond to
313 vegetation condition and thus increase noise. The effectiveness of the multi-year sine and cosine

¹ <https://www.planet.com/>

314 model for modelling the typical behaviour of NDYI, the utilisation of extreme Δ NDYI values as
315 'seeds' for regions that grew into areas of lower Δ NDYI values, and the ability to tune spectral
316 value constraints have all contributed to the effectiveness of our approach. To further improve the
317 performance of the Δ NDYI method, it would be worth investigating the use of supporting indices like
318 Garcia *et al.* [22] and Fernández-Martínez *et al.* [20], in addition to adding extra data sources. Further
319 work distinguishing different beech species would add greater value to DOC as those with larger
320 seeds (red and hard beech) have a disproportionately-larger impact on rodent irruptions.

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

329 5. Conclusions

330 This study used Sentinel-2 top-of-atmosphere (TOA) imagery to detect and map atypical yellowing
331 associated with heavy flowering of southern beech (*Fuscospora* and *Lophozonia*) in New
332 Zealand over 4.1 million ha at an unprecedented 10 m spatial resolution. This was achieved by
333 modelling a ~~normalised~~ *normalized* difference yellowing index (NDYI) over 5 years of observations
334 and investigating deviations from expected values during spring months (September–November).
335 A 'threshold' Δ NDYI value of 0.08 may be used to identify areas of heavy flowering, with
336 connected areas of Δ NDYI > 0.04 also likely flowering. The method has been automated and
337 can be run for all of New Zealand in less than a day on a cluster of approximately 1000 CPU cores.
338 Using Sentinel-2 imagery, the method typically provides information on heavy flowering for 80%
339 of the beech forests in New Zealand with a high ~~accuracy of over 90~~ overall classification accuracy of
340 ~~90.8%~~, producing ~~helpful information for useful information for planning~~ national-scale pest control
341 efforts.

342 **Author Contributions:** Conceptualization, Ben Jolly, John Dymond, Terry Greene, and Jan Schindler; Data
343 curation, Ben Jolly, James Shepherd and Jan Schindler; Formal analysis, Ben Jolly; Funding acquisition, John
344 Dymond; Investigation, Ben Jolly, John Dymond, and Terry Greene; Methodology, Ben Jolly, John Dymond, James
345 Shepherd, and Terry Greene; Project administration, John Dymond; Resources, John Dymond; Software, Ben Jolly,
346 James Shepherd and Jan Schindler; Supervision, John Dymond; Visualization, Ben Jolly; Writing – original draft,
347 Ben Jolly; Writing – review & editing, John Dymond, James Shepherd, and Terry Greene.

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

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

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

355 Abbreviations

356 The following abbreviations are used in this manuscript:

357

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

359

- 360 1. 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.
- 363 2. 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.
- 365 3. Wardle, J.A. *The New Zealand beeches. Ecology, utilisation and management*; New Zealand Forest Service:
Wellington, New Zealand, 1984; p. 477.
- 367 4. 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.
- 369 5. 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.; Internationl Rice Research Institute: Los Baños
(Philippines), 2010; pp. 239–251.
- 372 6. 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.
- 375 7. 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.
- 379 8. 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.
- 381 9. 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.
- 384 10. 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.
- 387 11. Zeng, L.; Wardlow, B.D.; Xiang, D.; Hu, S.; Li, D. Remote Sensing of Environment A review of vegetation
phenological metrics extraction using time-series , multispectral satellite data. *Remote Sensing of Environment*
2020, *237*, 111511. doi:10.1016/j.rse.2019.111511.
- 390 12. 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.
- 393 13. Misra, G.; Cawkwell, F.; Wingler, A. Status of phenological research using sentinel-2 data: A review. *Remote
Sensing* **2020**, *12*. doi:10.3390/RS12172760.
- 395 14. Dixon, D.J.; Callow, J.N.; Duncan, J.M.; Setterfield, S.A.; Pauli, N. Satellite prediction of forest flowering
phenology. *Remote Sensing of Environment* **2021**, *255*, 112197. doi:10.1016/j.rse.2020.112197.
- 397 15. Moon, M.; Seyednasrollah, B.; Richardson, A.D.; Friedl, M.A. Using time series of MODIS land surface
phenology to model temperature and photoperiod controls on spring greenup in North American
deciduous forests. *Remote Sensing of Environment* **2021**, *260*. doi:10.1016/j.rse.2021.112466.

- 400 16. Moon, M.; Richardson, A.D.; Friedl, M.A. Multiscale assessment of land surface phenology from
401 harmonized Landsat 8 and Sentinel-2, PlanetScope, and PhenoCam imagery. *Remote Sensing of Environment*
402 **2021**, *266*. doi:10.1016/j.rse.2021.112716.
- 403 17. Browning, D.M.; Russell, E.S.; Ponce-Campos, G.E.; Kaplan, N.; Richardson, A.D.; Seyednasrollah, B.;
404 Spiegal, S.; Saliendra, N.; Alfieri, J.G.; Baker, J.; Bernacchi, C.; Bestelmeyer, B.T.; Bosch, D.; Boughton, E.H.;
405 Boughton, R.K.; Clark, P.; Flerchinger, G.; Gomez-Casanovas, N.; Goslee, S.; Haddad, N.M.; Hoover, D.;
406 Jaradat, A.; Mauritz, M.; McCarty, G.W.; Miller, G.R.; Sadler, J.; Saha, A.; Scott, R.L.; Suyker, A.; Tweedie, C.;
407 Wood, J.D.; Zhang, X.; Taylor, S.D. Monitoring agroecosystem productivity and phenology at a national
408 scale: A metric assessment framework. *Ecological Indicators* **2021**, *131*. doi:10.1016/j.ecolind.2021.108147.
- 409 18. Noumonvi, K.D.; Oblišar, G.; Žust, A.; Vilhar, U. Empirical approach for modelling tree phenology in
410 mixed forests using remote sensing. *Remote Sensing* **2021**, *13*. doi:10.3390/rs13153015.
- 411 19. Atkinson, P.M.; Jeganathan, C.; Dash, J.; Atzberger, C. Inter-comparison of four models for smoothing
412 satellite sensor time-series data to estimate vegetation phenology. *Remote Sensing of Environment* **2012**,
413 *123*, 400–417. doi:10.1016/j.rse.2012.04.001.
- 414 20. Fernández-Martínez, M.; Garbulsky, M.; Peñuelas, J.; Peguero, G.; Espelta, J.M. Temporal trends in the
415 enhanced vegetation index and spring weather predict seed production in Mediterranean oaks. *Plant
416 Ecology* **2015**, *216*, 1061–1072. doi:10.1007/s1
- 417 21. Cheng, Y.; Vrieling, A.; Fava, F.; Meroni, M.; Marshall, M.; Gachoki, S. Phenology of short vegetation
418 cycles in a Kenyan rangeland from PlanetScope and Sentinel-2. *Remote Sensing of Environment* **2020**, *248*.
419 doi:10.1016/j.rse.2020.112004.
- 420 22. Garcia, M.; Zuckerberg, B.; Lamontagne, J.M.; Townsend, P.A. Landsat-based detection of mast
421 events in white spruce (*Picea glauca*) forests. *Remote Sensing of Environment* **2021**, *254*, 112278.
422 doi:10.1016/j.rse.2020.112278.
- 423 23. Asokan, A.; Anitha, J. Change detection techniques for remote sensing applications: a survey. *Earth Science
424 Informatics* **2019**. doi:10.1007/s12145-019-00380-5.
- 425 24. Panuju, D.R.; Paull, D.J.; Griffin, A.L. Change detection techniques based on multispectral images for
426 investigating land cover dynamics. *Remote Sensing* **2020**, *12*. doi:10.3390/rs12111781.
- 427 25. Chen, B.; Jin, Y.; Brown, P. An enhanced bloom index for quantifying floral phenology using multi-scale
428 remote sensing observations. *ISPRS Journal of Photogrammetry and Remote Sensing* **2019**, *156*, 108–120.
429 doi:10.1016/j.isprsjprs.2019.08.006.
- 430 26. Dymond, J.R.; Shepherd, J.D. The spatial distribution of indigenous forest and its composition in the
431 Wellington region, New Zealand, from ETM+ satellite imagery. *Remote Sensing of Environment* **2004**,
432 *90*, 116–125. doi:10.1016/j.rse.2003.11.013.
- 433 27. Landcare Research Ltd. EcoSat Forests (North Island), 2014. doi:10.26060/STZ2-Z482.
- 434 28. Landcare Research Ltd. EcoSat Forest (South Island), 2014. doi:10.26060/2EPD-WE90.
- 435 29. Department of Conservation. Flowering and fruit production, 2019.
- 436 30. Department of Conservation. Department of Conservation Te Papa Atawhai Annual Report 2019. Technical
437 report, Department of Conservation, 2019.
- 438 31. Shepherd, J.D.; Schindler, J.; Dymond, J.R. Automated Mosaicking of Sentinel-2 Satellite Imagery. *Remote
439 Sensing* **2020**, *12*, 3680. doi:10.3390/rs12223680.
- 440 32. Rouse, W.; Haas, H.; Deering, W. Monitoring vegetation systems in the Great Plains with ERTS. NASA.
441 Goddard Space Flight Center 3d ERTS-1 Symposium, 1974.
- 442 33. Tucker, C.J. Red and Photographic Infrared I , Linear Combinations for Monitoring Vegetation. *Remote
443 Sensing of Environment* **1979**, *8*, 127–150. doi:10.1016/0034-4257(79)90013-0.
- 444 34. Motohka, T.; Nasahara, K.N.; Oguma, H.; Tsuchida, S. Applicability of Green-Red Vegetation Index for
445 remote sensing of vegetation phenology. *Remote Sensing* **2010**, *2*, 2369–2387. doi:10.3390/rs2102369.
- 446 35. Zhu, Z.; Woodcock, C.E. Automated cloud, cloud shadow, and snow detection in multitemporal Landsat
447 data: An algorithm designed specifically for monitoring land cover change. *Remote Sensing of Environment*
448 **2014**, *152*, 217–234. doi:10.1016/j.rse.2014.06.012.
- 449 36. Shepherd, J.D.; Dymond, J.R.; Cuff, J.R. Monitoring scrub weed change in the Canterbury region using
450 satellite imagery. *New Zealand Plant Protection* **2007**, *60*, 137–140. doi:10.30843/nzpp.2007.60.4671.
- 451 37. Card, D.H. Using Known Map Category Marginal Frequencies to Improve Estimates of Thematic Map
452 Accuracy. *Photogrammetric Engineering and Remote Sensing* **1982**, *48*, 431–439.

- 453 38. Maxwell, A.E.; Warner, T.A.; Guillén, L.A. Accuracy assessment in convolutional neural network-based
454 deep learning remote sensing studies—part 1: Literature review. *Remote Sensing* **2021**, *13*.
455 doi:10.3390/rs13132450.
- 456 39. Uscooe, W.E.A.R.; Latt, K.E.H.P.; Richardson, S.J.; Allen, R.B.; Whitehead, D.; Carswell, F.E.; Ruscoe,
457 W.A.; Platt, K.H. Climate and net carbon availability determine temporal patterns of seed production by
458 Nothofagus. *Ecology* **2005**, *86*, 972–981.
- 459 40. Thapa, S.; Millan, V.E.G.; Eklundh, L. Assessing forest phenology: A multi-scale comparison of near-surface
460 (UAV, spectral reflectance sensor, phenocam) and satellite (MODIS, sentinel-2) remote sensing. *Remote
461 Sensing* **2021**, *13*. doi:10.3390/rs13081597.
- 462 41. Peng, D.; Wang, Y.; Xian, G.; Huete, A.R.; Huang, W.; Shen, M.; Wang, F.; Yu, L.; Liu, L.; Xie, Q.; Liu, L.;
463 Zhang, X. Investigation of land surface phenology detections in shrublands using multiple scale satellite
464 data. *Remote Sensing of Environment* **2021**, *252*. doi:10.1016/j.rse.2020.112133.

465 © 2022 by the authors. Submitted to *Remote Sens.* for possible open access publication
466 under the terms and conditions of the Creative Commons Attribution (CC BY) license
467 (<http://creativecommons.org/licenses/by/4.0/>).