



Satellite open data to monitor forest damage caused by extreme climate-induced events: a case study of the Vaia storm in Northern Italy

Journal:	<i>Forestry: An International Journal of Forest Research</i>
Manuscript ID	Forestry-2020-121.R2
Manuscript Type:	Original Article
Date Submitted by the Author:	n/a
Complete List of Authors:	VAGLIO LAURIN, Gaia; University of Tuscia, Francini, Saverio; University of Florence; Università degli Studi del Molise Dipartimento di Bioscienze e Territorio, Dipartimento di Bioscienze e Territorio, Luti, Tania; University of Florence Chirici, Gherardo; Università degli Studi di Firenze, GESAAF Pirotti, Francesco; University of Padova, TESAF Papale, Dario; University of Tuscia
Keywords:	Forest damages, Windthrow, Remote Sensing, Sentinel 1, Sentinel 2

SCHOLARONE™
Manuscripts

2 3 **Satellite open data to monitor forest damage caused by extreme** 4 **climate-induced events: a case study of the Vaia storm in** 5 **Northern Italy**

6
7 **Gaia Vaglio Laurin^{1*}, Saverio Francini^{2,5}, Tania Luti³, Gherardo Chirici²,**
8 **Francesco Pirotti⁴, Dario Papale¹**

9 ¹*Department for Innovation in Biological, Agro-Food and Forest Systems, University of*
10 *Tuscia, Viterbo, 01100, Italy*

11 ²*Department of Agricultural, Food and Forestry Systems, Università degli Studi di Firenze,*
12 *Firenze, 50145, Italy*

13 ³*Department of Earth Science, Università degli Studi di Firenze, Firenze, 50121, Italy*

14 ⁴*Interdepartmental Research Center of Geomatics (CIRGEO), University of Padova,*
15 *Legnaro, 35020, Italy*

16 ⁵*Dipartimento di Bioscienze e Territorio, Università degli Studi del Molise, Pesche, Isernia,*
17 *Italy*

18 *Corresponding author: Tel: +39 0761 357394; Fax: +39 0761 357389; Email: gaia.vl@unitus.it

19
20 The frequency of extreme storm events has significantly increased in the
21 past decades, causing significant damage to European forests. To mitigate
22 the impacts of extreme events a rapid assessment of forest damage is
23 crucial, and satellite data are an optimal candidate for this task. The
24 integration of satellite data in the operational phase of monitoring forest
25 damage can be exploit the complementarity of optical and Synthetic
26 Aperture Radar open datasets from the Copernicus programme. This study
27 illustrates the testing of Sentinel 1 and Sentinel 2 data for the detection of
28 areas impacted by the Vaia storm in Northern Italy. The use of multispectral
29 Sentinel 2 provided the best performance, with classification Overall
30 Accuracy values up to 86%; however optical data use are seriously
31 hampered by cloud cover that can persist for months after the event and in
32 most cases cannot be considered an appropriate tool if a fast response is
33 required. The results obtained using Synthetic Aperture Radar Sentinel 1
34 were slightly less accurate (Overall Accuracy up to 68%), but the method
35 was able to provide valuable information rapidly, mainly because the

1
2
3 36 acquisition of this dataset is weather independent. Overall, for a fast
4 37 assessment Sentinel 1 is the better of the two methods where multispectral
5 38 and ground data are able to further refine the initial SAR-based assessment.
6
7
8 39

10 40 **Introduction**

11
12 41 In recent years, extreme climate-induced events have caused significant damage to European
13 42 forests. The occurrence of strong storms has significantly increased in the past decades
14 43 (Usbeck et al. 2010), and the frequency of these events is expected to increase further in
15 44 future years due to changing climate dynamics (Saadet al. 2017; Seidl et al. 2014).
16 45 Windthrow has a major impact on forest dynamics. Forests affected by repeated damage may
17 46 not have enough time to recover and are more vulnerable to other threats. The forest
18 47 regeneration in areas damaged by storms can alter the overall ecosystem succession, with
19 48 consequences on biodiversity (Ellison et al. 2005). The way in which the windthrown timber
20 49 is managed has an impact on biodiversity (Duelli et al. 2019). After a storm, the fungal
21 50 infections over deadwood can increase and expand, promoting stand degradation (McCarthy
22 51 et al. 2012); similarly, the expansion of outbreaks of bark beetles from damaged to healthy
23 52 stands, that commonly occur from one to three years after windthrow, is known to
24 53 additionally impact conifer forests (Havašová et al. 2017). Civil security issues can also be
25 54 very relevant in windthrow areas (Gardiner et al. 2010).

26
27 55 Rapid assessment of forest damage is crucial for decision support regarding actions to be
28 56 taken to prevent further damage and to mitigate the impacts of future extreme events. Field
29 57 operations are commonly the first response from forest authorities for human security, timber
30 58 management, and ecosystem conservation. The planning and execution of forest operations
31 59 takes advantage from the availability of rapid information regarding spatial characteristics of
32 60 the strongly impacted sites and also regarding the extent and severity of damage.
33 61 Accessibility of remote forest areas is also a key factor that influences efforts that are
34 62 required to collect this type of data. Remote sensing (RS) has frequently been employed to
35 63 monitor different forest hazards and it has been previously used also in the case of the
36 64 detection of storm-damaged trees. Various RS data can be used for post-event forest damage
37 65 assessment, each having specific advantages and disadvantages, with their selection driven by
38 66 site characteristics, imagery and resources availability, and the question being answered
39 67 (Schwarz et al. 2003). Several case studies are needed, considering the variety of instruments
40 68 and environmental conditions, to understand how to better support the integration of remote
41 69 sensing tools in forest management practice.

42
43 70 With airborne or UAVs surveys, very detailed information up to single tree level can be
44 71 produced using digital cameras (Duan et al. 2017; Hamdi et al. 2019; Honkavaara et al.
45 72 2013; Mokroš et al. 2017; Pirotti et al. 2016), or laser scanning instruments (Marchi et al.
46 73 2017; Chirici et al. 2018), or commercial multispectral sensors (Jackson et al. 2000). But on-
47 74 demand airborne or UAVs surveys are costly, suited for areas of limited extent, and flights
48 75 can be hampered for weeks after the event by bad weather conditions, such as heavy rain or
49
50
51
52
53
54
55
56
57
58
59
60

1
2
3 76 post-fire smoke. Natural hazards often affect large areas and cause diffuse impacts,
4 77 consequently the mapping and monitoring of the area can take advantage of the use of data
5 78 from satellite platforms, that can cover broad extents repeatedly in time (Poursanidis and
6 79 Chrysoulakis 2017). For this purpose, on-demand multispectral satellite images at very high
7 80 spatial resolution (< 5 m) were previously successfully used in Russian (Kislov et al. 2020),
8 81 and in German forests (Einzmann et al. 2017; Schwarz et al. 2003). Several forest damage
9 82 assessments were conducted using medium spatial resolution multispectral data, particularly
10 83 the Landsat open archive that offers free data at 30 m pixel size. As an example, with Landsat
11 84 data an old diffusive windthrow caused by a storm was detected in a French forest by Haidu
12 85 et al. (2019); disturbance due to intensive harvesting and strong windthrow was mapped in
13 86 Western Siberia forests by Dyukarev et al. (2011); damage caused by a wind storm were
14 87 assessed in Lithuania forests by Jonikavičius and Mozgeris (2013); windthrow disturbance
15 88 was mapped in the temperate forest zone of European Russia and the southern boreal forest
16 89 zone of the United States by Baumann et al. (2014). Recently, a Pan-European mapping of
17 90 windthrow was generated through a model based on Landsat images, plus ancillary forest
18 91 data from other satellites and national inventory data (Pecchi et al. 2019). Overall, optical-
19 92 based studies have demonstrated the feasibility of detecting windthrow in forests using
20 93 satellite images and that the accuracy of results depends mainly on the spatial and spectral
21 94 resolutions of the datasets. However, the use of optical data for the rapid assessment of forest
22 95 windthrows is not encouraged due to different factors, including the purchase cost and time in
23 96 the case of on-demand images, and, importantly, the presence of clouds that in most cases
24 97 persists for weeks after weather-related events and hamper the use of satellite data.

25
26
27
28
29
30
31
32
33 98 To cope with these limits, the use of Synthetic Aperture Radar (SAR) satellite data are
34 99 recommended for a faster response. SAR data are independent from solar illumination and
35 100 weather. They are therefore available right after the event only depending on the revisit time
36 101 of the carrier, even when adverse weather conditions persist. Different SAR missions are
37 102 available at present, each with specific configurations in terms of the frequency of the active
38 103 signal, the polarization, and the spatial resolution. In forests, the energy backscattered by
39 104 SAR systems at higher frequencies (e.g. X and C-band) mainly comes from the crowns and
40 105 the upper forest strata, while at lower frequencies (e.g. L and P-bands) the contribution from
41 106 branches and trunks increases, together with the signal penetration through the canopy
42 107 (Solimini et al. 2016). The SAR backscatter is also influenced by the water content and the
43 108 geometric features of the target object, thus in the case of forests by the moisture levels (in
44 109 vegetation and soil) and the vegetation structure, including stem, branches and leaf
45 110 characteristics and architecture (Woodhouse 2005). In severely damaged forest areas the
46 111 geometric features suddenly change, as well as the surface roughness, making SAR data
47 112 potentially suitable for forest damage assessment, and specifically for detection of windthrow
48 113 (Eriksson et al. 2012). SAR datasets can bring additional and complementary information
49 114 with respect to optical data (e.g. on canopy roughness, water content, and volume) (Green
50 115 1998). Few studies proved the value of SAR data in the context of detection of forest
51 116 windthrows including: a multisensor based research conducted by Schwarz et al. (2003), who
52 117 compared the results obtained with SAR data against those from optical data; the detection of
53 118 areas affected by wind and insect outbreaks performed with L-band data (Tanase et al. 2018);

1
2
3 119 and the detection of windthrow in Germany and Switzerland based on Sentinel 1 C-band data
4 120 (Rüetschi et al. 2019).

6
7 121 Considering the different characteristics of satellite data, their integration into operational
8 122 forest monitoring after extreme events seeks to exploit the complementary features of optical
9 123 and SAR data. At present, this is feasible using the Copernicus European satellite missions,
10 124 specifically the Sentinel 1 C-band SAR data and the Sentinel 2 optical multispectral data
11 125 (Drusch et al. 2012; Torres et al. 2012). More relevant, these datasets are also available as
12 126 preprocessed products in Google Earth Engine (GEE), an integrated platform designed to
13 127 empower not only traditional remote sensing scientists but also a wider audience with limited
14 128 technical image processing skills (Gorelick et al. 2017). With its dense time series of optical
15 129 and SAR data provided free, already preprocessed, the Copernicus datasets represent an
16 130 optimal tool for the rapid assessment of land processes, including large scale forest damage
17 131 and windthrow, as in this specific case study.

22 132 The Vaia storm hit the North-Eastern part of Italy on the 29th October 2018; with winds
23 133 exceeding 200 km/h and strong rainfall it caused extensive forest damage in 494
24 134 municipalities, destroying or severely damaging forests of about 42,500 ha, with an
25 135 estimated stock of fallen trees of 85 million of cubic metres (Chirici et al. 2019). The
26 136 Copernicus Emergency Mapping system reports only about 4000 ha of damaged areas, about
27 137 10% of the affected area, due to cloud cover presence in the optical images used for mapping
28 138 (data available at: <https://emergency.copernicus.eu/mapping/list-of-components/EMSR334>).
29 139 Following the Vaia storm, the impacted regions assessed the forest damage by means of the
30 140 integration of aerial photographs or very high-resolution optical satellite images with data
31 141 from field surveys.

36 142 The present research tests Sentinel 1 and Sentinel 2 data for the detection of areas impacted
37 143 by the Vaia storm. To classify healthy and damaged areas, different algorithms were
38 144 evaluated, including a Bayesian Generalized Linear Model, a k-Nearest Neighbors approach,
39 145 and Random Forest, using ground data provided by the regional authorities for model
40 146 calibration and validation. Change detection approaches based on pre and post event image
41 147 differencing were frequently used in previous research (Dalponte et al. 2020; Ruetschi et al.
42 148 2019; Tanase et al. 2018). In this work we evaluated the impact of algorithm selection on
43 149 results, to support the selection of proper methods in operational forest monitoring. The
44 150 present research expands on the common monitoring of forest windthrow based on optical
45 151 data, which is ineffective in case of adverse atmospheric conditions, and it introduces testing
46 152 of Sentinel 1 SAR C-band. Sentinel 2 optical data are here tested for the first time, according
47 153 to our knowledge, in the context of the detection of forest damages by storms. Even if the
48 154 sensitivity of SAR signal to forest damages was previously illustrated by various authors
49 155 (Eriksson et al. 2012; Thiele et al. 2012; Ulander et al. 2005), only very few studies exploited
50 156 SAR for windthrow mapping (Ruetschi et al. 2019; Tanase et al. 2018), possibly due to data
51 157 complexities and limited access to user-friendly processing tools. Thus, the present study can
52 158 be of help to understand how SAR can support forestry practice, also considering that these
53 159 data and related tools are increasingly available by different space agencies, and preprocessed
54 160 Sentinel 1 datasets are delivered by GEE.

1
2
3 161 The aim of this research is to contribute to forestry practice, developing knowledge useful to
4 162 operational management to exploit satellite open-data, and defining a strategy for the rapid
5 163 detection of forest damage and the further refinement of information. Remote sensing has
6 164 great potential to cost-efficiently map storm-affected regions, but previous research has been
7 165 somewhat limited, as Sentinel 2 imagery was not previously exploited with this aim; and
8 166 Sentinel 1 was only partially examined. Thus, further investigation to assess the potential of
9 167 integrating satellite open-data into forest practical workflows is needed. Here the focus is on
10 168 open-data from the Copernicus programme, exploiting the GEE platform for fast processing,
11 169 demonstrating that this approach can significantly support decision makers with remote
12 170 sensing-based assessment of windthrow damaged areas.

171 **Methods**

172 *Study area and ground data*

173 The research was conducted in Northern Italy, in areas affected by the windthrow and
174 included in two selected Sentinel-2 tiles for which ground truth data were made available by
175 local administrations encompassing three regions: Friuli Venezia Giulia, Trentino Alto
176 Adige, and Veneto (fig. 1). These regions host important forest resources, and different local
177 agencies in charge of their census and management were involved in assessing the Vaia
178 impacts. These data can also be found in open databases (Forzieri et al. 2020).

179 *Insert fig. 1*

180 For the Trentino Alto Adige region, ground data for the Trento Autonomous Province were
181 provided by the local forest service, and for the Autonomous Province of Bozen by the
182 Province authority; in both cases the area of damaged forest were detected on the basis of
183 photointerpretation of aerial orthophotos, and integrated with data from field surveys.
184 Overall, in Trentino Alto Adige there were 1463 discrete areas, covering 5913 ha, and with
185 a mean area of 4 ha. For the Friuli Venezia Giulia region, ground data on forest damage were
186 provided by the local forest service using aerial orthophotos and ground surveys; there were
187 499 damaged areas for this region, covering 3693 ha, with mean area of 7.4 ha. For the
188 Veneto region, the ground data were provided by the Veneto Agency for agriculture
189 payments (AVEPA), who are responsible to provide economic help in case of natural
190 disasters. AVEPA provided a shapefile of the affected areas based on photointerpretation of
191 very high resolution orthophotos (20 cm spatial resolution) and SPOT 6/7 satellite pre and
192 post event images at 1.5 m spatial resolution. The dataset included information on area
193 borders and estimation of percentage of damaged trees in each area. In total, there were 1588
194 damaged areas detected in Veneto, covering 4020 ha, and having a mean surface of 2.5 ha.

195 The data provided for these Italian regions included 3550 polygons that identify any area
196 affected by the windthrow. These polygons were filtered out to create a subset for testing and
197 validation purposes, according to the following inclusion criteria: (i) polygons >2 ha, to
198 include areas compatible with the spatial resolution and the detection capability of the remote
199 sensing data used in this study; (ii) polygons in which the average terrain slope was below
200 20% in at least 85% of the surface, to exclude areas of unreliable SAR signal, according to

201 distortion; (iii) for the Veneto region only, polygons in which the amount of damaged trees
202 resulted > 80%, that were more than 40% of the total Veneto polygons; (iv) polygons
203 included in two Sentinel tiles, to test the methods in the most affected area (3218 polygons).
204 For classification purposes, polygons in forest not impacted by the Vaia storm were also
205 drawn in proximity of the impacted polygons, by on-screen photointerpretation of post-event
206 Google imagery.

207 In particular, the first criterion was guided by the imagery spatial resolution and allowed to
208 retain the larger damaged areas. The second criterion was derived after exploring the SAR
209 distortion masks based on local incidence angle. These maps are a by-product of SAR data
210 processing, that was additionally performed as these layers are not included in the GEE
211 available datasets. The maps indicated frequent distortions above the 20% slope; to facilitate
212 the analysis using GEE data, this single threshold was selected. The third criterion was
213 applied in Veneto, and was introduced due to the different ways the Italian regions assessed
214 damage. In fact, it was noted that Trentino and Friuli Venezia Giulia reported only areas
215 where damage was very significant, while Veneto Region also digitized areas that were
216 partially impacted. The dataset was standardized by keeping only the polygons with damage
217 degree above 80% in Veneto. The application of the mentioned criteria resulted in a
218 standardized dataset including the larger and most affected areas, where SAR data had the
219 higher signal to noise ratio and the forest impacts were similar. The damaged and non-
220 damaged datasets included a total of 209 polygons, 104 from healthy forest stands, and 105
221 from damaged forest areas; the corresponding pixels were extracted from the imagery and
222 averaged at polygon level. In total 90% of the 209 polygons were used to calibrate and
223 validate with the k-fold approach classification algorithms. The remaining 22 polygons were
224 used as independent test set for further evaluation of the overall accuracy. The total area used
225 for calibration, validation and testing the methods was considerable: in Trentino Alto Adige it
226 was equal to 622 ha; in Veneto to 533 ha; and in Friuli 237 ha, representing the different
227 forest types and environmental conditions occurring in the area of interest.

228 *Copernicus Sentinel data*

229 The Sentinel 2 (S2) multispectral images were downloaded from Google Earth Engine as
230 Level-2A orthorectified atmospherically corrected surface reflectance. The S2 Multispectral
231 Instrument (MSI) samples 13 spectral bands: visible and NIR at 10 meters, red edge and
232 SWIR at 20 meters, and atmospheric bands at 60 meters spatial resolution. Only bands at 10 -
233 20 m spatial resolution were used for tests (bands # 2, 3, 4, 5, 6, 7, 8, 8A, 11, 12), resampling
234 at 10 m the 20 m bands with a nearest neighbor approach. The vegetation indices included in
235 Table 1 were also computed.

236 *Insert Table 1*

237 To evaluate the hypothesis that Sentinel 2 data can detect the damaged forest areas with
238 significant accuracy, post-event Sentinel 2 images were used. The possibility to use also a
239 pre-damage image and focus the analysis on the variations in reflectance was also evaluated,
240 but confounding factors such as day-specific atmospheric conditions, including cloud cover,

1
2
3 241 and plant phenology stage at different dates were considered relevant causes of increased
4 242 uncertainty in results and leading to uncertainty. Therefore, we preferred to work only on
5 243 post-damage optical images using a binary classification approach (healthy forest/damaged
6 244 areas).

7
8
9 245 The first available post-event Sentinel 2 imagery is dated June 2019 (7 months after the
10 246 event). The predictor set named S2_Set1 (set of image bands) was used to evaluate the
11 247 contribution of each single band, and the S2_Set2 (vegetation indices) to evaluate the
12 248 contribution of the derived vegetation indices. The latter combination is preferable in case
13 249 different images are used (predictor sets in Table 2).

14
15
16
17 250 The Sentinel 1 SAR images were downloaded from Google Earth Engine as Ground Range
18 251 Detected (GRD) scenes, already pre-processed using the Sentinel-1 Toolbox to generate a
19 252 calibrated, ortho-corrected product at 10 m spatial resolution in dual-band cross polarization
20 253 mode (VV – VH). Preprocessing included thermal noise removal, radiometric calibration, and
21 254 terrain correction using a digital terrain model (SRTM 30 m). Five pre-event scenes were
22 255 collected from the period 26 September - 3 October 2018 (pre-event period without frost or
23 256 snow), and 5 post-event scenes were from the period 7 - 15 December 2018. The pre and post
24 257 event scenes were averaged at pixel level, and band ratios (VV/VH, VH/VV), and band
25 258 normalized differences were also computed (VV-VH, VH -VV). In fact, with respect to the
26 259 optical images, the SAR data are less affected by atmospheric condition and vegetation
27 260 phenology and for this reason the use of differences between pre and post event was also
28 261 evaluated to detect forest damaged areas.

29
30
31
32
33 262 Thus, the set of predictors named S1_Set3 (based only on post-event bands), and the S1_Set4
34 263 (based on pre-post event scenes differences) were used in tests (predictor sets in Table 2).

35
36
37 264 *Insert Table 2*

38 39 265 **Classification approaches**

40
41 266 Three different approaches were tested for the classification task: a generalized linear
42 267 Bayesian model and two machine learning models, the k-Nearest Neighbors and Random
43 268 Forest. Using the three models and the four sets of available predictors (Table 2), a total of 12
44 269 models-predictors combinations were developed. The tests were conducted using the
45 270 RFTrainer, KNN, and Bayesglm R packages in R environment (R Core Team 2013).

46
47
48
49 271 Bayesian inference is a method of statistical inference in which Bayes' theorem is used to
50 272 update the probability for a hypothesis as more evidence or information becomes available. It
51 273 facilitates representing and taking full account of the uncertainties related to models and
52 274 parameter values. The Bayesian generalized linear model (BGLM) is based on Bayesian
53 275 functions that finds an approximate posterior mode and variance using extensions of the
54 276 classical generalized linear model computations. The Bayesian function allows the user to
55 277 specify independent prior distributions for the coefficients in the t family, with the default
56 278 being Cauchy distributions with center 0 and scale set to 10 (for the regression intercept), 2.5
57
58
59
60

279 (for binary predictors), or $2.5/(2 \cdot sd)$, where sd is the standard deviation of the predictor in
280 the data (for other numerical predictors) (Berrett and Calder 2016; Gelman et al. 2008)

281 The k-Nearest Neighbors (KNN) technique is a popular method for producing spatially
282 contiguous predictions of forest attributes by combining field and remotely sensed data. KNN
283 are appealing as they can be used for both univariate and multivariate prediction, no
284 assumptions regarding the distributions of response or auxiliary variables are necessary and
285 they can be used with a wide variety of datasets (Chirici et al 2016). The k nearest vectors,
286 used to perform the classification, are found according to Minkowski distance and the
287 classification is performed by means of the maximum of summed kernel densities; both
288 ordinal and continuous variables can be predicted (Wu et al. 2002).

289 Random Forest (RF) is an ensemble of decision trees that learns through a supervised
290 approach and produces multiple models that are aggregated, using a bootstrap aggregating
291 procedure, to produce the result. The models are built using different training subsets,
292 generated by bootstrapping, that are used to build the “forest”. RF is able to reduce the output
293 variance and the overfitting problem with respect to other machine learning approaches,
294 improving model stability and accuracy (Breiman 2001).

295 When a model is trained with data there is the risk of overfitting, i.e. that the parameters are
296 estimated to reproduce closely the training data used, losing the capacity to generalize outside
297 the calibration examples. To avoid overfitting one of the most useful method is k -fold cross
298 validation (k -fold CV) that splits the training set into K number of subsets, called folds: the
299 models are then iteratively fitted K times each time training the data on data from $K-1$ of the
300 folds and evaluating the performances on data from the K th fold. At the end of calibration,
301 the performance on each of the K folds are evaluated in term of Overall Accuracy (OA), i.e.
302 the percentage of cases where the classification as damaged or not damaged was correct. The
303 Overall Accuracy from Cross Validation - OA_{cv} and the relative standard deviation $sd(OA_{cv})$
304 are finally computed averaging the K folds and calculating their standard deviation. This
305 provides more information over how stable the model is by testing it over multiple sets of
306 data.

308 RF and kNN models require the calibration of hyperparameters- The hyperparameters are
309 calculated using the training datasets. BGLM instead does not require hyperparameters
310 calibration. A procedure based on a random search grid was used for the optimization
311 (Bergstra and Bengio, 2012). The procedure defines a grid of hyperparameter ranges, as those
312 defined above. One hundred combinations were randomly sampled from the grid and for each
313 combination a k -Fold CV was performed. For both KNN and RF the optimal
314 hyperparameters combination with the greater OA_{cv} was finally selected. In the RF case, the
315 hyperparameters that were tuned include the maximum depth of each tree (max_depth) in the
316 forest and the number of features ($max_features$) considered by each tree when splitting a
317 node. The number of trees in the forest was set equal to 400, while the minimum number of
318 samples required to split an internal node was set equal to 1.

319 In the kNN case the three hyperparameters that were optimized are: the number of neighbours
320 considered (k), the Minkosky distance, and the kernel to use. The $max_features$ ranged

1
2
3 321 between 2 and n , where n equals the number of predictors used in input; the `max_depth`
4 322 ranges between 1 and 40. k ranged between 1 and 60, the Minkosky distance was equal to the
5 323 Euclidean and Manhattan distances, and the kernel to use were Unweighted, Weighted,
6 324 Inverse, Reciprocal.

7
8 325
9
10 326 The models, once optimized and validated with k -fold approach, were further independently
11 327 evaluated in terms of accuracy using the test set that includes 10% of the polygons never used
12 328 during the optimization procedure. For the evaluation of the classification results, different
13 329 statistics are reported, including overall accuracy, users accuracy, producers accuracy, and
14 330 the percentage of omission and commission errors. User's accuracy represents how reliable
15 331 the classification is in terms of actually finding damage in the real world over an area that
16 332 was classified as "damaged" in the map. Producer's accuracy reports how often a damage that
17 333 is found in the real world is reported in the final classified map (Cohen 1968; Congalton
18 334 1991).

22 335 **Results**

23
24 336 In table 3 we present for each model and for each of the four sets of predictors the averaged
25 337 OA_{cv} obtained with a 9-fold cross validation procedure and the related standard deviation
26 338 obtained averaging the different iterations. For the KNN and RF models the best
27 339 hyperparameters combination, identified using a RandomSearch algorithm are also presented.

30 340 *Insert Table 3*

31
32
33 341 The results in Table 3, also graphically shown in fig. 2, show consistency in different models,
34 342 with negligible differences among the considered approaches and limited variance from
35 343 different iterations.

36
37 344 The best results are obtained with S2 images, with OA_{cv} always > 0.8 and included in the 0.8-
38 345 0.85 range for either bands or vegetation indices, with the latter reaching a slightly higher
39 346 accuracy. The standard deviation values resulted were always small with a maximum value of
40 347 0.102. This confirms the ability of S2 to detect impacted forest areas.

41
42
43 348 Lower accuracy results –in the 0.6-0.7 OA_{cv} range- are obtained when using SAR data with a
44 349 slightly better scores obtained when using the S1_Set3, that includes only data from post-
45 350 event scene.

46
47
48 351 *Insert fig. 2*

49
50
51 352 The three models were also applied to the independent test set ($n = 22$) to evaluate their final
52 353 performance on new and unseen data and the results are shown in Table 4.

53 354 *Insert Table 4*

54
55
56 355 The results obtained using the independent test set are similar to those obtained with 9-fold
57 356 CV but span, as expected, over a slightly higher range, considering the limited number of
58 357 samples in the test set ($n=22$).

1
2
3 358 For sets 1 and 2, based on S2 data, the accuracy is included in the 0.77-0.86 range, and is
4 359 similar across the three different models. For sets 3 and 4, based on S1 data, the accuracy
5 360 range is 0.5-0.68, with higher results obtained using Random Forest model.
6
7

8 361 Overall, results in Table 4 confirm the higher accuracies obtained with the use of Sentinel 2
9 362 data with respect to those obtained with SAR data.
10

11 363 **Discussion**

12
13 364 The availability of satellite open-data that is also partly preprocessed adds significant value to
14 365 the procedure of the assessment of forest damages, from windthrow or other sources of
15 366 damage that change the landscape. The final product is a classified damage map that supports
16 367 rapid responses in terms of forest management. The results indicate that data from the
17 368 Copernicus Sentinel 1 and 2 missions are suited for the detection of damaged forest areas.
18 369 SAR is especially useful for a fast evaluation, providing useful information for
19 370 immediate/short-term response actions for risk mitigation. Sentinel 2 can be used to refine the
20 371 SAR initial information unless post event data are immediately available. The use of cloud-
21 372 based platforms like Google Earth Engine helps to reduce the time that operators need for
22 373 image download and standard pre-processing. A pre-defined workflow over ready-to-use
23 374 imagery can avoid requiring highly skilled operators for processing imagery. The workflow
24 375 can be partly automatic, providing maps useful to multiple end-users, even those less familiar
25 376 with image processing techniques.
26
27
28
29
30

31 377 Focusing on applications, the present research suggests that the sequential use of GEE
32 378 Sentinel 1 and 2 data for better windthrow information provision is an optimal combination.
33 379 Specifically, the testing of the different predictors from S1 and S2 data provided useful
34 380 insights on the advantages and limits of these datasets.
35
36
37

38 381 The best detection of the forest areas impacted by the Vaia storm is always obtained using
39 382 Sentinel 2 images. Using a 9-fold cross validation approach and either S2 bands or vegetation
40 383 indices as input, the obtained overall accuracies were $> 80\%$, with limited differences among
41 384 modeling approaches, low variance from iterations, and results included in the 80-85% range.
42 385 The use of vegetation indices with KNN and RF approaches provided the higher OA_{CV}
43 386 values, equal to 85 and 84 %, respectively.
44
45

46 387 Very similar results are obtained when the parameterized models were validated against the
47 388 independent test set, represented by 22 samples not used to calibrate the models. The
48 389 obtained results, although the number of test samples is relatively low, are in a very similar
49 390 OA range (77- 86%) compared with those reported for 9-fold cross validation. The highest
50 391 OA score (86%) is obtained either using S2 bands with KNN model or using VIs with
51 392 Random Forest. User accuracies were over 90% whereas producer's accuracies were in the
52 393 71-83% range, with higher scores obtained with RF and VIs. This indicates that commission
53 394 errors were lower than omission errors, in other words some damaged areas were not
54 395 correctly detected by the classifier, thus leaving out some areas from the final map, but most
55 396 of the areas classified as damaged were really damaged. It might be due to canopy of felled
56 397 trees still significantly showing in the image, or also reflectance from water vapor, that
57
58
59
60

398 commonly rises in the morning in mountainous areas, mixing with the reflectance values
399 from the tree trunks.

400 It should be considered that the post-event S2 imagery employed in the study were from a
401 single date and about 7 months after the Vaia event. The time gap between the storm and the
402 S2 imagery allows the greening of the ground in damaged areas, from herbal and shrubs
403 vegetation regrowth that can produce a change in reflectance values and a consequent
404 negative impact in the classification accuracy. The results are however in line with what has
405 been already found in the past with Landsat data, thus at 30 m spatial resolution (30 m): an
406 OA equal to 86% was reached in the detection of windthrows in Voges mountains in France
407 (Haidu et al. 2019); in European Russia and United States the OA was about 75%, with more
408 accurate results reported for larger areas (Baumann et al. 2014); and with an automatic
409 algorithm based on Landsat time series, historical disturbance from windthrow and logging
410 was detected in United States forest with an accuracy about 80%. According to our
411 knowledge, there are no studies based on the use of S2 data for forest windthrow detection,
412 except two abstracts where the accuracy of the obtained results is not reported (Cenci et al.
413 2019; Valt et al. 2019).

414 Vegetation indices are especially useful when multiple images are used (as in the case of
415 change detection analysis), or when the study area is large and covered by different image
416 tiles, or even when a mosaic from different dates is composed to mitigate cloud cover issues.
417 In fact, VIs are designed to maximize sensitivity to the vegetation characteristics while
418 minimizing confounding factors such as soil background reflectance, directional, or
419 atmospheric effects, that change among different acquisitions (Fang and Liang 2008).
420 According to this and based on the reported classification results, the use of VIs from S2 data,
421 fitting the model with the RF method, appears the best solution to detect damaged forest area
422 in this case study. An improvement in accuracy is expected if optical imagery becomes
423 available in dates close to the time of the damage event.

424 The use of SAR inputs produced lower accuracies compared to S2 inputs with an OA values
425 in the range 0.66-0.71 according to 9-fold cross validation and a low standard deviation score
426 (< 0.14). Differences among the three models are minor, as well as those when using post-
427 event data only or pre-post event backscattering difference. The highest OA_{cv} score is
428 obtained with pre-event data and KNN approach ($OA_{cv} = 0.71$).

429 According to results from the independent test set, the OAs slightly decrease and the
430 variability in values increases, being included in the 0.5-0.68 range. The best scores are
431 obtained using RF: 64% and 68% in OA with post-event data and pre-post event difference,
432 respectively. With RF, the user's accuracy is low (54%) when using S1 pre-post event scenes
433 differences, but the producer's accuracy is in line with the one obtained using the S2 data
434 (75%). When using post-event data, the user's and producer's accuracies are on the same
435 order (63%). The results suggest the use of RF as classification model, and the combination
436 of pre- and post-event SAR scenes to better meet the user's needs.

1
2
3 437 Previous studies conducted with C-band ERS 1/2 and RADARSAT 1 at 30 m spatial
4 438 resolution were not successful for the detection of forest windthrows (Schwarz et al. 2003;
5 439 Ulander et al. 2005). However, with L-band data - that better penetrates into the forest - OAs
6 440 included in the 69-84% range were obtained in the Bavarian Forest National Park, with
7 441 accuracy values depending on the acquisition date and environmental conditions (Tanase et
8 442 al. 2018). Using C-band S1 and a change detection approach, the producer's accuracy
9 443 reached 88% in a German validation site, but the user's accuracy was quite low (21%) and
10 444 limitations consisted in a minimum area of 0.5 ha and the requirement of 10 post-event
11 445 images (Rüetschi et al. 2019). Positive results were also obtained using X-band data with
12 446 very high spatial resolution (Thiele et al. 2012).

13
14
15
16
17 447 The results here presented outline the relevance of SAR spatial resolution for forest
18 448 windthrow detection, and confirm the ability of S1 data to produce fast preliminary
19 449 information on impacted areas, with the obtained OA and user's accuracy included in a range
20 450 of values similar to those reported by other studies.

21
22
23 451 It is important to note the limitations of the present study in terms of suitability in certain
24 452 cases. First, the areas tested were all above 2 ha, to cope with the 10 m spatial resolution of
25 453 both S1 and S2 datasets; however the average size of damaged areas in the three considered
26 454 regions resulted higher than 2 ha. Then, the ground truth was filtered out to exclude slopes
27 455 using a low threshold ($> 20^\circ$) where the first SAR distortion effects were observed. Producing
28 456 the distortion masks to filter out data, instead of a fixed threshold that excluded most slopes
29 457 for easiness of analysis in GEE, could result in a detailed mapping of unreliable SAR pixels
30 458 and a larger availability of reliable data over slopes. Similarly, the use of temporal series can
31 459 improve the amount of area with reliable SAR data, as at each pass the acquisition angle may
32 460 vary. When SAR data are used, it is also important to detect and mask pixels with wet or dry
33 461 snow over the canopy, as it changes the backscatter values (Koskinen et al. 1997). This
34 462 implies an added complexity of the method in areas seasonally covered by snow.

35
36
37
38
39
40 463 For future operational use, the application of a pre-disturbance forest/non-forest map can help
41 464 to perform semi-automatic classification. Further tests are also needed to understand the
42 465 response of satellite data over less impacted forests, where a mixture of healthy and damaged
43 466 trees is present, and how to minimize the impact of SAR distortion areas.

44
45
46 467 The combined use of S1 and S2 was not investigated here, as for data integration the optical
47 468 and radar imagery should be from same period. The S2 images used in this investigation were
48 469 dated months after the storm, when the herbal and shrub vegetation renovation influence both
49 470 the optical reflectance and the backscattering in C-band SAR. These are confounding factors,
50 471 but a data integration approach is feasible if optical data are available soon after the event.
51 472 Higher accuracy in classification is known to occur from combining SAR with optical with
52 473 respect to use single sensor type (Clerici et al. 2017; Vaglio Laurin et al. 2012), so this might
53 474 be another strategy to improve the information accuracy.

54 55 56 57 475 **Conclusion**

1
2
3 476 This study showed the suitability of Copernicus S1 and S2 data for the detection of areas
4 477 affected by windthrow. Sentinel 2 provided the best performance for detection of windthrow
5 478 areas, but its use was seriously hampered by cloud cover. For events occurring in winter,
6 479 Sentinel 2 data might only be available after several months. In those cases, the use of
7 480 Sentinel 1 data, being independent with respect to atmospheric condition and with a fast
8 481 return time, becomes the best option for a first and rapid evaluation of the forest damage, to
9 482 support field operations and the formation of management response plans.

10
11
12
13 483 Thus, for operational monitoring, the results suggest a sequential approach, based initially on
14 484 S1 for fast response. This initial SAR assessment can be refined in later dates, integrating S2
15 485 imagery when available and data from ground or aerial surveys, for a more accurate mapping
16 486 also over steep slopes.

17 18 19 487 **Data availability statement**

20
21 488 Remote sensing data are freely available by the Copernicus facilities. Ground data were
22 489 provided by local authorities and can be requested directly to them. Part of the data can be
23 490 found in the open database published by Forzieri et al. (2020).

24 25 26 491 **Funding**

27
28 492 This work was supported by internal resources of the participating institutes. In particular F.P.'s work
29 493 was supported by University of Padova's VAIA-FRONT (VAIA, FROM lessons learnt to future
30 494 options) project.

31 32 33 495 **Acknowledgements**

34
35 496 For the Trentino Autonomous Province, we acknowledge Dr. A. Wolynski from the local forest
36 497 service who coordinated the acquisition of ground data on the basis of aerial photography and field
37 498 surveys.

38
39 499 For the Alto Adige region, we acknowledge Dr. F. Maistrelli from the Autonomous Province of
40 500 Bozen who provided the ground data delimited on the basis of aerial photography and field surveys.

41
42 501 For the Friuli Venezia Giulia region, we acknowledge Dr. R. Comino from the local forest service
43 502 who coordinated the acquisition of ground data on the basis of aerial photography and ground
44 503 surveys.

45
46 504 For the Veneto Region we acknowledge the Agenzia Veneta per i Pagamenti in Agricoltura (AVEPA)
47 505 that provided the data.

48 49 50 506 **Conflict of interest statement**

51
52 507 'None declared.'

53 54 55 508 **References**

56
57 509 Baumann, M.; Ozdogan, M.; Wolter, P.T.; Krylov, A.; Vladimirova, N.; Radeloff, V.C.
58 510 (2014) Landsat remote sensing of forest windfall disturbance. *Remote Sens. Environ.*, 143,
59 511 171–179.

- 1
2
3 512 Bergstra, J., & Bengio, Y. (2012). Random search for hyper-parameter optimization. *Journal*
4 513 *of machine learning research*, 13, 281-305.
- 5
6
7 514 Berrett, C., & Calder, C. A. (2016). Bayesian spatial binary classification. *Spatial Statistics*,
8 515 16, 72-102.
- 9
10 516 Breiman, L. (2001). Random forests. *Machine learning*, 45(1), 5-32.
- 11
12 517 Cenci, L., De Giorgi, A., Squicciarino, G., Pulvirenti, L., Moser, G., & Boni, G. (2019).
13 518 Exploiting Sentinel 2 data for mapping windstorm damages in forested areas. Case Study: the
14 519 event of October 2018 occurred in Northeast Italy. *Geophysical Research Abstracts*, vol. 21.
- 15
16
17 520 Chirici, G., Mura, M., McInerney, D., Py, N., Tomppo, E. O., Waser, L. T., ... & McRoberts,
18 521 R. E. (2016). A meta-analysis and review of the literature on the k-Nearest Neighbors
19 522 technique for forestry applications that use remotely sensed data. *Remote Sensing of*
20 523 *Environment*, 176, 282-294.
- 21
22
23 524 Chirici, G., Bottalico, F., Giannetti, F., Del Perugia, B., Travaglini, D., Nocentini, S., ... &
24 525 Fattorini, L. (2018). Assessing forest windthrow damage using single-date, post-event
25 526 airborne laser scanning data. *Forestry: An International Journal of Forest Research*, 91(1), 27-
26 527 37.
- 27
28
29 528 Chirici, G., Giannetti, F., Travaglini, D., Nocentini, S., Francini, S., D'Amico, G., ... &
30 529 Tonner, J. (2019). Stima dei danni della tempesta "Vaia" alle foreste in Italia. *Forest@-*
31 530 *Journal of Silviculture and Forest Ecology*, 16(1), 3.
- 32
33
34 531 Clerici, N., Valbuena Calderón, C. A., & Posada, J. M. (2017). Fusion of Sentinel-1A and
35 532 Sentinel-2A data for land cover mapping: a case study in the lower Magdalena region,
36 533 Colombia. *Journal of Maps*, 13(2), 718-726.
- 37
38
39 534 Cohen, J. (1968) Weighted kappa: Nominal scale agreement with provision for scaled
40 535 disagreement or partial credit. *Psychological Bulletin* 70 (4):213-220.
- 41
42
43 536 Congalton, R.G. (1991) A review of assessing the accuracy of classification of remotely
44 537 sensed data. *Remote Sensing of Environment* 37:35-46
- 45
46 538 Dalponte, M., Marzini, S., Solano-Correa, Y. T., Tonon, G., Vescovo, L., & Gianelle, D.
47 539 (2020). Mapping forest windthrows using high spatial resolution multispectral satellite
48 540 images. *International Journal of Applied Earth Observation and Geoinformation*, 93, 102206.
- 49
50
51 541 Drusch, M., Del Bello, U., Carlier, S., Colin, O., Fernandez, V., Gascon, F., ... & Meygret, A.
52 542 (2012) Sentinel-2: ESA's optical high-resolution mission for GMES operational services.
53 543 *Remote sensing of Environment*, 120, 25-36.
- 54
55
56 544 Duan, F.; Wan, Y.; Deng, L. (2017) A novel approach for coarse-to-fine windthrown tree
57 545 extraction based on unmanned aerial vehicle images. *Remote Sens.*, 9, 306.
- 58
59
60

- 1
2
3 546 Duelli, P., Wermelinger, B., Moretti, M., & Obrist, M. K. (2019). Fire and windthrow in
4 547 forests: Winners and losers in Neuropterida and Mecoptera. *Alpine Entomology*, 3, 39.
- 6
7 548 Dyukarev, E.A.; Pologova, N.N.; Golovatskaya, E.A.; Dyukarev, A.G. (2011) Forest cover
8 549 disturbances in the South Taiga of West Siberia. *Environ. Res. Lett.*, 68.
- 10
11 550 Einzmann, K., Immitzer, M., Böck, S., Bauer, O., Schmitt, A., & Atzberger, C. (2017)
12 551 Windthrow detection in European forests with very high-resolution optical data. *Forests*, 8(1),
13 552 21.
- 15
16 553 Ellison, A. M. et al. (2005) Loss of foundation species: Consequences for the structure and
17 554 dynamics of forested ecosystems. *Front. Ecol. Environ.* 3, 479–486.
- 18
19 555 Eriksson, L.E.B.; Fransson, J.E.S.; Soja, M.J.; Santoro, M. (2012) Backscatter signatures of
20 556 wind-thrown forest in satellite SAR images. In *Proceedings of the International Geoscience
21 557 and Remote Sensing Symposium (IGARSS)*, Munich, Germany, 22–27 July; pp. 6435–6438.
- 23
24 558 Fang H. and S. Liang. (2008) Leaf Area Index Models. *Encyclopedia of Ecology*, pp. 2139-
25 559 2148
- 27
28 560 Forzieri, G., Pecchi, M., Girardello, M., Mauri, A., Klaus, M., Nikolov, C., ... Beck, P. S. A.
29 561 (2020). A spatially explicit database of wind disturbances in European forests over the period
30 562 2000 - 2018. *Earth System Science Data*, 12(1), 257–276.
- 31
32 563 Gardiner, B., Blennow, K., Carnus, J. M., Fleischer, P., Ingemarsson, F., Landmann, G., ... &
33 564 Peyron, J. L. (2010). Destructive storms in European forests: past and forthcoming impacts.
34 565 European Forest Institute, Efiatlantic.
- 36
37 566 Gelman, A., Jakulin, A., Pittau, M. G., & Su, Y. S. (2008) A weakly informative default prior
38 567 distribution for logistic and other regression models. *The annals of applied statistics*, 2(4),
39 568 1360-1383.
- 41
42 569 Gorelick, N., Hancher, M., Dixon, M., Ilyushchenko, S., Thau, D., & Moore, R. (2017)
43 570 Google Earth Engine: Planetary-scale geospatial analysis for everyone. *Remote sensing of
44 571 Environment*, 202, 18-27.
- 46
47 572 Green, R.M. (1998) The sensitivity of SAR backscatter to forest windthrow gaps. *Int. J.
48 573 Remote Sens.*, 19, 2419–2425.
- 49
50 574 Haidu, I., Furtuna, P. R., & Lebaut, S. (2019) Detection of old scattered windthrow using low
51 575 cost resources. The case of Storm Xynthia in the Vosges Mountains, 28 February 2010. *Open
52 576 Geosciences*, 11(1), 492-504.
- 54
55 577 Hamdi, Z. M., Brandmeier, M., & Straub, C. (2019) Forest Damage Assessment Using Deep
56 578 Learning on High Resolution Remote Sensing Data. *Remote Sensing*, 11(17), 1976.
- 57
58
59
60

- 1
2
3 579 Havašová, M., Ferencík, J., & Jakuš, R. (2017) Interactions between windthrow, bark beetles
4 and forest management in the Tatra national parks. *Forest Ecology and Management*, 391,
5 580 349-361.
6 581
- 7
8 582 Honkavaara, E.; Litkey, P.; Nurminen, K. (2013) Automatic Storm Damage Detection in
9 583 Forests Using High-Altitude. Photogrammetric Imagery. *Remote Sens.*, 5, 1405–1424.
- 10
11 584 Huang, C., Goward, S. N., Masek, J. G., Thomas, N., Zhu, Z., & Vogelmann, J. E. (2010) An
12 585 automated approach for reconstructing recent forest disturbance history using dense Landsat
13 586 time series stacks. *Remote Sensing of Environment*, 114(1), 183-198.
- 14
15
16 587 Jackson, R. G., Foody, G. M., & Quine, C. P. (2000) Characterizing windthrown gaps from
17 588 fine spatial resolution remotely sensed data. *Forest Ecology and Management*, 135(1-3), 253-
18 589 260.
- 19
20
21 590 Jonikavičius, D., & Mozgeris, G. (2013) Rapid assessment of wind storm-caused forest
22 591 damage using satellite images and stand-wise forest inventory data. *iForest-Biogeosciences*
23 592 and *Forestry*, 6(3), 150.
- 24
25
26 593 Kislov, D. E., & Korznikov, K. A. (2020) Automatic Windthrow Detection Using Very-
27 594 High-Resolution Satellite Imagery and Deep Learning. *Remote Sensing*, 12(7), 1145.
- 28
29 595 Koskinen, J.T.; Pulliainen, J.T.; Hallikainen, M.T. (1997) The Use of ERS-1 SAR Data in
30 596 Snow Melt Monitoring. *IEEE Trans. Geosci. Remote Sens.* 35, 601–610.
- 31
32
33 597 Marchi, N., Pirotti, F., & Lingua, E. (2018). Airborne and Terrestrial Laser Scanning Data for
34 598 the Assessment of Standing and Lying Deadwood: Current Situation and New Perspectives.
35 599 *Remote Sensing*, 10(9), 1356.
- 36
37
38 600 McCarthy, J. K., Hood, I. A., Kimberley, M. O., Didham, R. K., Bakys, R., Fleet, K. R., ... &
39 601 Brockerhoff, E. G. (2012) Effects of season and region on sapstain and wood degrade
40 602 following simulated storm damage in *Pinus radiata* plantations. *Forest ecology and*
41 603 *management*, 277, 81-89.
- 42
43
44 604 Mokroš, M.; Výbošťok, J.; Merganič, J.; Hollaus, M.; Barton, I.; Koren, M.; Tomaščík, J.;
45 605 Černá, J. (2017) Early stage forest windthrow estimation based on unmanned aircraft
46 606 system imagery. *Forests* 8, 306.
- 47
48
49 607 Pecchi, M., Forzieri, G., Ceccherini, G., Spinoni, J., Feyen, L., Cescatti, A., & Chirici, G.
50 608 (2019) Pan-European mapping of windthrows. *Geophysical Research Abstracts*, vol. 21.
- 51
52 609 Pirotti, F., Travaglini, D., Giannetti, F., Kutchartt, E., Bottalico, F., & Chirici, G. (2016).
53 610 Kernel feature cross-correlation for unsupervised quantification of damage from windthrow
54 611 in forests. *ISPRS - International Archives of the Photogrammetry, Remote Sensing and*
55 612 *Spatial Information Sciences*, XLI-B7, 17–22.
- 56
57
58
59
60

- 1
2
3 613 Poursanidis, D., & Chrysoulakis, N. (2017) Remote Sensing, natural hazards and the
4 614 contribution of ESA Sentinels missions. *Remote Sensing Applications: Society and*
5 615 *Environment*, 6, 25-38.
- 6
7
8 616 R Core Team (2013). R: A language and environment for statistical computing. R Foundation
9 617 for Statistical Computing, Vienna, Austria. URL <http://www.R-project.org/>.
- 10
11 618 Rüetschi, M., Small, D., & Waser, L. T. (2019) Rapid detection of windthrows using
12 619 sentinel-1 c-band sar data. *Remote Sensing*, 11(2), 115.
- 13
14
15 620 Saad, C., Boulanger, Y., Beaudet, M., Gachon, P., Ruel, J. C., & Gauthier, S. (2017) Potential
16 621 impact of climate change on the risk of windthrow in eastern Canada's forests. *Climatic*
17 622 *Change*, 143(3-4), 487-501.
- 18
19
20 623 Schwarz, M., Steinmeier, C., Holecz, F., Stebler, O., & Wagner, H. (2003) Detection of
21 624 windthrow in mountainous regions with different remote sensing data and classification
22 625 methods. *Scandinavian journal of forest research*, 18(6), 525-536.
- 23
24
25 626 Seidl, R., Schelhaas, M. J., Rammer, W., & Verkerk, P. J. (2014) Increasing forest
26 627 disturbances in Europe and their impact on carbon storage. *Nature climate change*, 4(9), 806-
27 628 810.
- 28
29 629 Solimini, D. (2016) *Understanding Earth Observation*; Springer International Publishing:
30 630 Basel, Switzerland, pp. 1–703, ISBN 978-3-319-25633-7.
- 31
32
33 631 Tanase, M. A., Aponte, C., Mermoz, S., Bouvet, A., Le Toan, T., & Heurich, M. (2018)
34 632 Detection of windthrows and insect outbreaks by L-band SAR: A case study in the Bavarian
35 633 Forest National Park. *Remote Sensing of Environment*, 209, 700-711.
- 36
37
38 634 Thiele, A., Boldt, M., & Hinz, S. (2012) Automated detection of storm damage in forest areas
39 635 by analyzing TerraSAR-X data. In: 2012 IEEE International Geoscience and Remote Sensing
40 636 Symposium (pp. 1672-1675). IEEE.
- 41
42
43 637 Torres, R.; Snoeij, P.; Geudtner, D.; Bibby, D.; Davidson, M.; Attema, E.; Potin, P.;
44 638 Rommen, B.Ö.; Floury, N.; Brown, M.; et al. (2012) GMES Sentinel-1 mission. *Remote*
45 639 *Sens. Environ.*, 120, 9–24.
- 46
47
48 640 Ulander, L.M.H.; Smith, G.; Eriksson, L.; Folkesson, K.; Fransson, J.E.S.; Gustavsson, A.;
49 641 Hallberg, B.; Joyce, S.; Magnusson, M.; Olsson, H.; et al. (2005) Mapping of wind-thrown
50 642 forests in southern Sweden using space- and airborne SAR. In *Proceedings of the*
51 643 *International Geoscience and Remote Sensing Symposium (IGARSS)*, Seoul, South Korea,
52 644 25–29 July 2005; pp. 3619–3622.
- 53
54
55 645 Usbeck, T., Wohlgemuth, T., Dobbertin, M., Pfister, C., Bürgi, A., & Rebetez, M. (2010).
56 646 Increasing storm damage to forests in Switzerland from 1858 to 2007. *Agricultural and Forest*
57 647 *Meteorology*, 150(1), 47-55.
- 58
59
60

1
2
3 648 Vaglio Laurin, G., Liesenberg, V., Chen, Q., Guerriero, L., Del Frate, F., Bartolini, A., ... &
4 649 Valentini, R. (2013) Optical and SAR sensor synergies for forest and land cover mapping in a
5 650 tropical site in West Africa. *International Journal of Applied Earth Observation and*
6 651 *Geoinformation*, 21, 7-16.

7
8
9 652 Valt, M., Salvatori, R., & Salzano, R. (2019) Mapping the effects of VAIA storm using
10 653 Sentinel 2 data. *Geophysical Research Abstracts*, vol. 21.

11
12 654 Woodhouse, I. H. (2005). *Introduction to microwave remote sensing*. CRC press.

13
14 655 Wu, Y., Ianakiev, K., & Govindaraju, V. (2002) Improved k-nearest neighbor classification.
15 656 *Pattern recognition*, 35(10), 2311-2318.

16
17 657

18
19 658

20
21 659

22
23 660

24
25 661

26
27 662

28
29 663

30
31 664

32
33 665

34
35 666

36
37 667

38
39 668

40
41 669

42
43 670

44
45 671

46
47 672

48
49 673

50
51 674

52
53 675

54
55 676

1
2
3 677 **Table and Figure captions**
4

5 678 **Figure 1** Study area, Northern Italy; in red the 209 polygons used for calibration, cross
6 validation and testing; in orange the rectangles of Sentinel 2 tiles. Image prepared using
7 Google Earth Pro Landsat/Copernicus @2020 GeoBasis-DE/BKG US Dept of State
8 Geographer @2020 Google.
9 681

10
11 682 **Figure 2** Classification results using three algorithms (BGLM = Bayesian Generalized Linear
12 Model; KNN = K-Nearest Neighbor; RF = Random Forests) with four sets of predictors as
13 input (Set 1 = Sentinel 2 bands; Set 2 = Sentinel 2 vegetation indices; Set 3 = Sentinel 1 post
14 event data; Set 4 = Sentinel 1 pre-post event difference data). The models were validated with
15 a 9-fold cross validation approach.
16 686
17
18
19 687
20
21 688
22
23 689
24
25 690
26
27 691
28
29 692
30
31 693
32
33 694
34
35 695
36
37 696
38
39 697
40
41 698
42
43 699
44
45 700
46
47 701
48
49 702
50
51 703
52
53 704
54
55 705
56
57 706
58
59
60

707 **Table 1.** List of vegetation indices used for tests.

	Vis	Index name	Bands
1	NDVI_idx	Normalized Difference Veg index	$(b8-b4)/(b8+b4)$
2	NBr_idx	Normalized Burn Ratio	$(b8 - b12)/(b8 + b12)$
3	NDVI_2	Normalized Difference Veg index 2	$(b12-b8)/(b12+b8)$
4	SR	Simple Ratio	$b8/b4$
5	ARI1	Anthocyanin Reflectance Index 1 (ARI1)	$1/b3-1/b5$
6	EVI	Enhanced Vegetation Index	$2.5*(b8-b4)/(b8+6*b4-7.5*b2)+1000$
7	NDMI	Normal difference moisture index	$(b8-b11)/(b8+b11)$
8	MSI	Moisture soil index	$b11/b8$
9	BAI	Burn Area Index	$1/(0.1-b4)^2+(0.06-b8)^2$
10	DVI	Difference Veg Index	$b8-b4$
11	GDVI	Green Difference Vegetation Index	$b8 - b3$
12	GARI	Green Atmospherically Resistant Index	$b8-(b3- (b2-b4)/b8+(b3- (b2-b4)$
13	GRVI	Green Ratio Vegetation Index	$b8/b3$
14	IPVI	Infrared Percentage Vegetation Index	$b8/b8+b4$

708

709

710

711

712

713

714

715

716 **Table 2.** Set of S2 and S1 predictors used in classification models.

Set	Predictors	Image date
S2_Set1	Sentinel 2 (post event) bands	28/06/2019
S2_Set2	Sentinel 2 (post event) Vegetation Indices	28/06/2019
S1_Set3	Sentinel 1 (post event) bands VH, VV	07-15/12/2018
	Sentinel 1 (post event) band ratios VV/VH, VH/VV	07-15/12/2018
	Sentinel 1 (post event) normalized difference VV-VH, VH-VV	07-15/12/2018
S1_Set4	Sentinel 1 (pre-post event difference) bands VH, VV	26/09- 03/10/2018
		07-15/12/2018
	Sentinel 1 (post event difference) band ratios VV/VH, VH/VV	26/09- 03/10/2018
		07-15/12/2018
	Sentinel 1 (pre-post event difference) normalized difference VV-VH, VH-VV	26/09- 03/10/2018
		07-15/12/2018

717

718

719

720

721

722

723

724

725

726

727

728

729 **Table 3.** Overall accuracy for classification models validated with 9-fold cross validation
 730 (OA_{cv}), with related standard deviation *sd*(OA_{cv}) and best hyperparameters combination.
 731 The highest Accuracy for each model is shown in bold.

		Best hyperparameters		9-fold-cross validation	
BGLM	<i>Predictors set</i>			OA _{cv}	<i>sd</i> (OA _{cv})
	S2_Set1			0.80	0.086
	S2_Set2			0.82	0.073
	S1_Set3			0.68	0.072
	S1_Set4			0.67	0.096
KNN		<i>kmax</i>	<i>Distance</i>		
	S2_Set1	15	2	0.82	0.081
	S2_Set2	20	1	0.85	0.102
	S1_Set3	53	2	0.71	0.138
	S1_Set4	12	2	0.66	0.085
RF		<i>Max. features</i>	<i>Max. depth</i>		
	S2_Set1	4	11	0.83	0.070
	S2_Set2	8	27	0.84	0.064
	S1_Set3	3	40	0.66	0.075
	S1_Set4	4	35	0.66	0.089

732

733

734

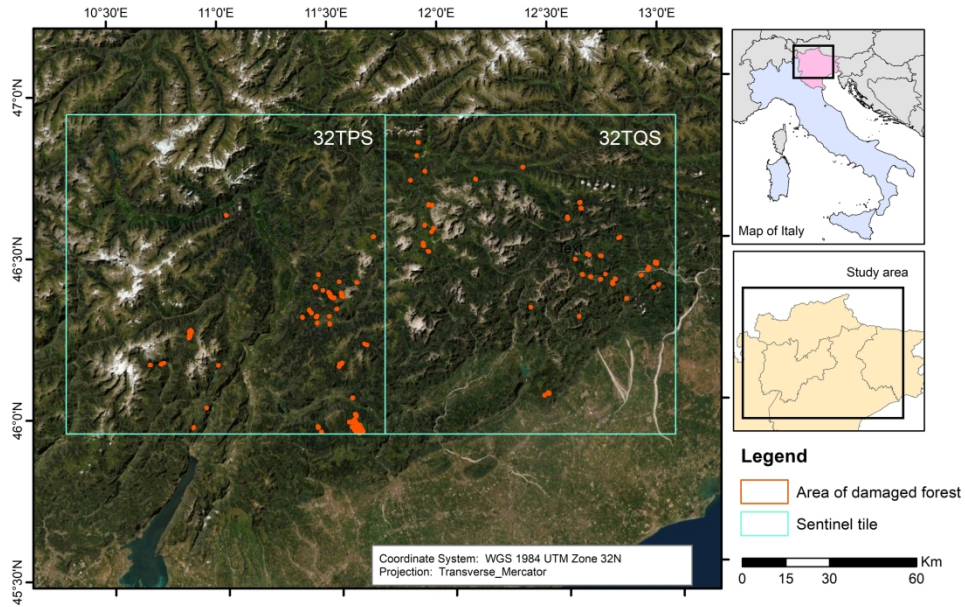
735

736

737 **Table 4.** Accuracy statistics for the three classification models and the four set of predictors
 738 obtained on the Test set (10% of samples). In bold the highest OA obtained.

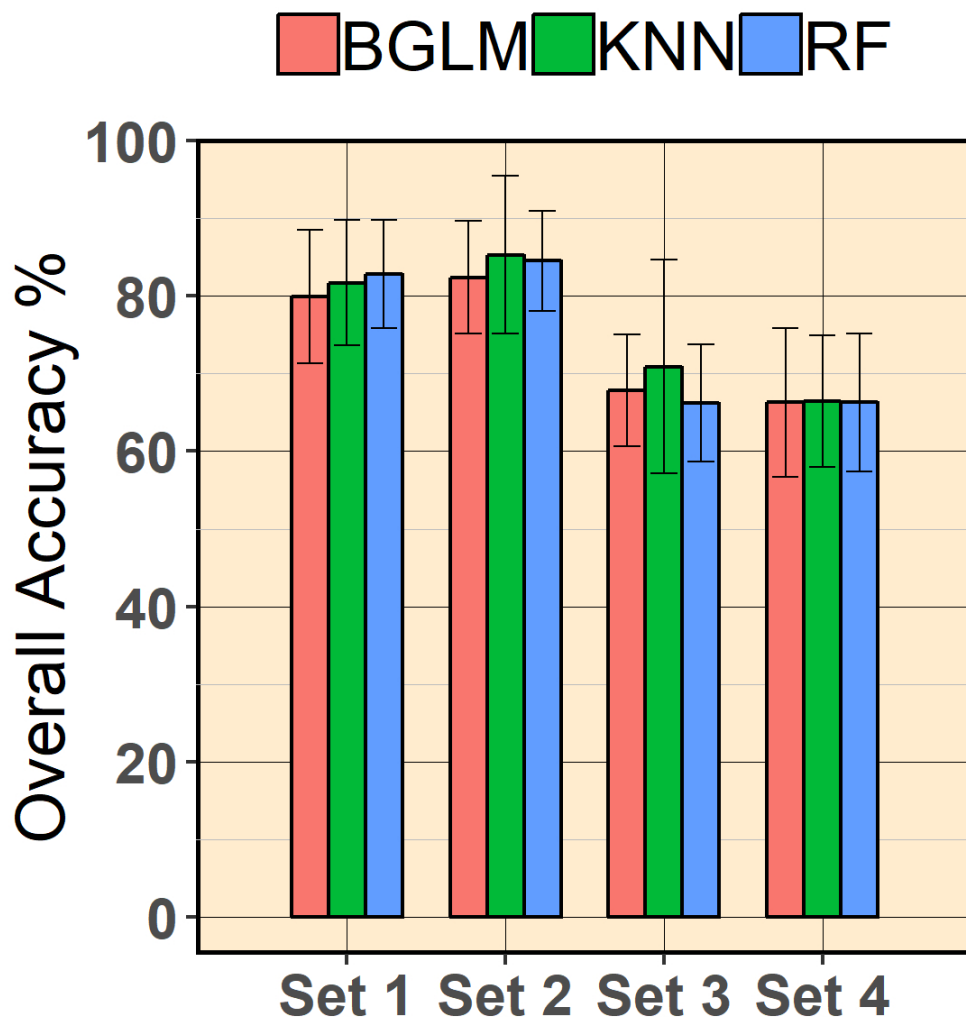
	Predictor set	Overall accuracy	Producers accuracy Healthy forest %	Producers accuracy Damaged areas %	Users accuracy Healthy forest %	Users accuracy Damaged areas %
BGLM	S2_Set1	0.77	87.50	71.43	63.64	90.91
	S2_Set2	0.82	100.00	73.33	63.64	100.00
	S1_Set3	0.50	50.00	50.00	27.27	72.73
	S1_Set4	0.55	54.55	54.55	54.55	54.55
KNN	S2_Set1	0.86	100.00	78.57	72.73	100.00
	S2_Set2	0.82	88.89	76.92	72.73	90.91
	S1_Set3	0.50	50.00	50.00	45.45	54.55
	S1_Set4	0.64	61.54	66.67	72.73	54.55
RF	S2_Set1	0.82	88.89	76.92	72.73	90.91
	S2_Set2	0.86	90.00	83.33	81.82	90.91
	S1_Set3	0.64	63.64	63.64	63.64	63.64
	S1_Set4	0.68	64.29	75.00	81.82	54.55

739



Study area, Northern Italy; in red the 209 polygons used for calibration, cross validation and testing; in orange the rectangles of Sentinel 2 tiles. Image prepared using Google Earth Pro Landsat/Copernicus @2020 GeoBasis-DE/BKG US Dept of State Geographer @2020 Google.

248x175mm (300 x 300 DPI)



Classification results using three algorithms (BGLM = Bayesian Generalized Linear Model; KNN = K-Nearest Neighbor; RF = Random Forests) with four sets of predictors as input (Set 1 = Sentinel 2 bands; Set 2 = Sentinel 2 vegetation indices; Set 3 = Sentinel 1 post event data; Set 4 = Sentinel 1 pre-post event difference data). The models were validated with a 9-fold cross validation approach.

1
2
3 **Response to reviewer #2**
4

5 Many thanks for the positive evaluation of this revised version. We implemented
6 the minor suggested changes as follows:

7
8 L 137-156: please revise: Revision has been made

9
10 L 358-360: please revise: Revision has been made

11 Figure 1: The figure has been improved and the polygons are now well visible.
12
13
14
15
16
17
18
19
20
21
22
23
24
25
26
27
28
29
30
31
32
33
34
35
36
37
38
39
40
41
42
43
44
45
46
47
48
49
50
51
52
53
54
55
56
57
58
59
60

For Review Only

2
3
4
5
6
7
8
9
10
11
12
13
14
15
16
17
18
19
20
21
22
23
24
25
26
27
28
29
30
31
32
33
34
35
36
37
38
39
40
41
42
43
44
45
46
47
48
49
50
51
52
53
54
55
56
57
58
59
60
**Satellite open data to monitor forest damage caused by extreme
climate-induced events: a case study of the Vaia storm in
Northern Italy**

**Gaia Vaglio Laurin^{1*}, Saverio Francini^{2,5}, Tania Luti³, Gherardo Chirici²,
Francesco Pirotti⁴, Dario Papale¹**

¹*Department for Innovation in Biological, Agro-Food and Forest Systems, University of
Tuscia, Viterbo, 01100, Italy*

²*Department of Agricultural, Food and Forestry Systems, Università degli Studi di Firenze,
Firenze, 50145, Italy*

³*Department of Earth Science, Università degli Studi di Firenze, Firenze, 50121, Italy*

⁴*Interdepartmental Research Center of Geomatics (CIRGEO), University of Padova,
Legnaro, 35020, Italy*

⁵*Dipartimento di Bioscienze e Territorio, Università degli Studi del Molise, Pesche, Isernia,
Italy*

*Corresponding author: Tel: +39 0761 357394; Fax: +39 0761 357389; Email: gaia.vl@unitus.it

The ~~occurrence~~ frequency of ~~strong storm~~ extreme storm events has significantly increased in the past decades, causing significant damage to European forests. To mitigate the impacts of extreme events a rapid assessment of forest damage is crucial, and satellite data are an optimal candidate for this task. The integration of satellite data in the operational phase of monitoring ~~of~~ forest damage can be ~~promoted~~ exploiting exploit the complementarity of optical and Synthetic Aperture Radar open datasets from the Copernicus ~~free of cost datasets~~ programme. This study illustrates the testing of Sentinel 1 and Sentinel 2 data for the detection of areas impacted by the Vaia storm in Northern Italy. The use of multispectral Sentinel 2 provided the best performance, with classification Overall Accuracy values up to 86%; however optical data use are seriously hampered by cloud cover that can persist for months after the event and in most cases cannot be considered an appropriate tool if a fast response is required. The results obtained using Synthetic Aperture Radar Sentinel 1 were slightly less accurate (Overall Accuracy up to 68%), but the method

1
2
3 36 was able to provide valuable information rapidly, mainly because the
4 37 acquisition of this dataset is weather independent. Overall, for a fast
5 38 assessment Sentinel 1 is the better of the two methods where multispectral
6 39 and ground data are able to further refine the initial SAR-based assessment.
7
8
9
10 40

11 41 **Introduction**

12
13
14 42 In recent years, extreme climate-induced events have caused significant damage to European
15 43 forests. The occurrence of strong storms has significantly increased in the past decades
16 44 (Usbeck et al. 2010), and the frequency of these events is expected to increase ~~even~~
17 45 ~~more further~~ in future years due to ~~the~~ changing climate dynamics (Saadet al. 2017; Seidl et
18 46 al. 2014). ~~Windthrown is~~ Windthrow has a major ~~driver of impact on~~ forest dynamics. Forests
19 47 affected by repeated damage may not have enough time to recover and are ~~exposed more~~
20 48 ~~vulnerable~~ to ~~additional impacts other threats~~. The ~~regrowth forest regeneration~~ in areas
21 49 damaged by storms can alter the overall ecosystem succession, with consequences on
22 50 biodiversity (Ellison et al. 2005). The way in which the windthrown timber is managed ~~also~~
23 51 has an impact on biodiversity (Duelli et al. 2019). After a storm, the fungal infections over
24 52 deadwood can increase and expand, promoting stand degradation (McCarthy et al. 2012);
25 53 similarly, the expansion of outbreaks of bark beetles ~~outbreaks~~ from damaged to healthy
26 54 stands, that commonly occur from one to three years after windthrow, is known to
27 55 additionally impact conifer forests (Havašová et al. 2017). Civil security issues can also be
28 56 very relevant in windthrow areas (Gardiner et al. 2010).

29
30
31
32
33
34 57 ~~For operational planning, fast intervention, Rapid assessment of forest damage is crucial for~~
35 58 ~~decision support regarding actions to be taken to prevent further damage~~ and to mitigate the
36 59 impacts of future extreme events, ~~a rapid assessment of forest damage is crucial~~. Field
37 60 operations ~~conducted in a timely way by~~ are commonly the first response from forest
38 61 authorities ~~are fundamental~~ for human security, ~~for~~ timber management, and ~~for~~ ecosystem
39 62 conservation. The planning and execution of forest operations ~~needs to take~~ takes advantage
40 63 ~~off from~~ the availability of rapid information, ~~locating the most regarding spatial~~
41 64 ~~characteristics of the strongly~~ impacted sites and also regarding the extent and severity of
42 65 damage, ~~also considering the difficulty to work in~~. Accessibility of remote forest areas is also
43 66 a key factor that influences efforts that are required to collect this type of data. Remote
44 67 sensing (RS) has frequently been employed to monitor different forest hazards and it has
45 68 been previously used also in the case of ~~withdrawn damage~~ the detection of storm-damaged
46 69 trees. Various RS data can be used for post-event forest damage assessment, each having
47 70 specific advantages and disadvantages, with their selection driven by site characteristics,
48 71 imagery and resources availability, and the question being answered (Schwarz et al. 2003).
49 72 Several case studies are needed, considering the variety of instruments and environmental
50 73 conditions, to understand how to better support the integration of remote sensing tools in
51 74 forest management practice.

1
2
3 75 With airborne or UAVs surveys, very detailed information up to single tree level can be
4 76 produced using digital cameras (Duan et al. 2017; Hamdi et al. 2019; Honkavaara et al.
5 77 2013; Mokroš et al. 2017; Pirotti et al. 2016), or laser scanning instruments (Marchi et al.
6 78 2017; Chirici et al. 2018), or commercial multispectral sensors (Jackson et al. 2000). But on-
7 79 demand airborne or UAVs surveys are costly, suited for areas of limited extent, and flights
8 80 can be hampered for weeks after the event by bad weather conditions, such as heavy rain or
9 81 post-fire smoke. ~~As natural~~Natural hazards ~~are usually phenomena affecting~~often affect large
10 82 areas and ~~causing~~cause diffuse impacts, consequently the mapping and monitoring of
11 83 ~~damage~~the area can take advantage of the use of data from satellite platforms, that can cover
12 84 broad extents repeatedly in time (Poursanidis and Chrysoulakis 2017). For this purpose, on-
13 85 demand multispectral satellite images at very high spatial resolution (< 5 m) were previously
14 86 successfully used in Russian (Kislov et al. 2020), and in German forests (Einzmann et al.
15 87 2017; Schwarz et al. 2003). Several forest damage assessments were conducted using
16 88 medium spatial resolution multispectral data, particularly the Landsat open archive that offers
17 89 free data at 30 m pixel size. As an example, with Landsat data an old diffusive windthrow
18 90 caused by a storm was detected in a French forest by Haidu et al. (2019); disturbance due to
19 91 intensive harvesting and strong windthrow was mapped in Western Siberia forests by
20 92 Dyukarev et al. (2011); damage caused by a wind storm were assessed in Lithuania forests by
21 93 Jonikavičius and Mozgeris (2013); windthrow disturbance was mapped in the temperate
22 94 forest zone of European Russia and the southern boreal forest zone of the United States by
23 95 Baumann et al. (2014); ~~and recently~~. Recently, a Pan-European mapping of windthrow was
24 96 generated through a model based on Landsat images, plus ancillary forest data from other
25 97 satellites and national inventory data (Pecchi et al. 2019). Overall, optical-based studies have
26 98 demonstrated the feasibility of detecting windthrow in forests using satellite images and that
27 99 the accuracy of results depends mainly on the spatial and spectral resolutions of the datasets.
28 100 However, the use of optical data for the rapid assessment of forest windthrows is not
29 101 encouraged due to different factors, including the purchase cost and time in the case of on-
30 102 demand images, and, importantly, the presence of clouds that in most cases persists for
31 103 weeks after weather-related events and hamper the use of satellite data.

32
33
34
35
36
37
38
39
40
41
42 104 To cope with these limits, the use of Synthetic Aperture Radar (SAR) satellite data is~~are~~
43 105 recommended for a fast~~faster~~ response, ~~it is daylight and weather~~. SAR data are independent
44 106 data, that in principle are from solar illumination and weather. They are therefore available
45 107 soon~~right~~ after the event only depending on the revisit time of the carrier, even when adverse
46 108 weather conditions persist. Different SAR missions are available at present, each with
47 109 specific configurations in terms of the frequency of the active signal, the polarization, and the
48 110 spatial resolution. In forests, the energy backscattered by SAR systems at higher frequencies
49 111 (e.g. X and C-band) mainly comes from the crowns and the upper forest strata, while at lower
50 112 frequencies (e.g. L and P-bands) the contribution from branches and trunks increases,
51 113 together with the signal penetration through the canopy (Solimini et al. 2016). The SAR
52 114 backscatter is also influenced by the water content and the geometric features of the target
53 115 object, thus in the case of forests by the moisture levels (in vegetation and soil) and the
54 116 vegetation structure, including stem, branches and leaf characteristics and architecture
55 117 (Woodhouse 2005). In severely damaged forest areas the geometric features suddenly

1
2
3 118 change, as well as the surface roughness, making SAR data potentially suitable for forest
4 119 damage assessment, and specifically for detection of windthrow (Eriksson et al. 2012). SAR
5 120 datasets can bring additional and complementary information with respect to optical data (e.g.
6 121 on canopy roughness, water content, and volume) (Green 1998). Few studies proved the
7 122 value of SAR data in the context of detection of forest windthrows including: a multisensor
8 123 based research conducted by Schwarz et al. (2003), who compared the results obtained with
9 124 SAR data against those from optical data; the detection of areas affected by wind and insect
10 125 outbreaks performed with L-band data (Tanase et al. 2018); and the detection of windthrow
11 126 in Germany and Switzerland based on Sentinel 1 C-band data (Rüetschi et al. 2019).

12
13
14
15
16 127 Considering the different characteristics of satellite data, their integration into operational
17 128 forest monitoring after extreme events seeks to exploit the complementary features of optical
18 129 and SAR data. At present, this is feasible using the Copernicus European satellite
19 130 ~~facilities~~missions, specifically the Sentinel 1 C-band SAR data and the Sentinel 2 optical
20 131 multispectral data (Drusch et al. 2012; Torres et al. 2012). More relevant, these datasets are
21 132 also available as preprocessed products in Google Earth Engine (GEE), an integrated
22 133 platform designed to empower not only traditional remote sensing scientists but also a wider
23 134 audience with limited technical image processing skills (Gorelick et al. 2017). With its dense
24 135 time series of optical and SAR data provided free, already preprocessed, the Copernicus
25 136 datasets represent an optimal tool for the rapid assessment of land processes, including large
26 137 scale forest damage and windthrow, as in this specific case study.

27
28
29
30
31 138 The Vaia storm hit the North-Eastern part of Italy on the 29th October 2018; with winds
32 139 exceeding 200 km/h and strong rainfall it caused extensive forest damage in 494
33 140 municipalities, destroying or severely damaging forests of about 42,500 ha, with an
34 141 estimated stock of fallen trees of 85 million of cubic metres (Chirici et al. 2019). The
35 142 Copernicus Emergency Mapping system reports only about 4000 ha of damaged areas, about
36 143 10% of the affected area, due to cloud cover presence in the optical images used for mapping
37 144 (data available at: <https://emergency.copernicus.eu/mapping/list-of-components/EMSR334>).
38 145 Following the Vaia storm, the impacted regions assessed the forest damage by means of the
39 146 integration of aerial photographs or very high-resolution optical satellite images with data
40 147 from field surveys.

41
42
43
44
45 148 The present research ~~examines—the testing of tests~~ Sentinel 1 and Sentinel 2 data for the
46 149 detection of areas impacted by the Vaia storm. ~~Different algorithms were exploited for the~~
47 150 ~~classification of~~To classify healthy and damaged areas, ~~different algorithms were evaluated~~,
48 151 including a Bayesian Generalized Linear Model, a k-Nearest Neighbors approach, and
49 152 Random Forest, using ground ~~truthing data~~ provided by the regional authorities for model
50 153 calibration and validation. ~~With respect to sensors~~Change detection approaches based on pre
51 154 ~~and post event image differencing were frequently~~ used in previous ~~studies, research~~
52 155 ~~(Dalponte et al. 2020; Ruetschi et al. 2019; Tanase et al. 2018). In this work we evaluated the~~
53 156 ~~impact of algorithm selection on results, to support the selection of proper methods in~~
54 157 ~~operational forest monitoring. The present research expands on the classical~~common
55 158 ~~monitoring of~~ forest windthrow ~~monitoring~~-based on optical data, which is ineffective in case
56 159 of adverse atmospheric conditions, ~~and it introduces~~ testing ~~of~~ Sentinel 1 SAR C-band;

1
2
3
4 160 ~~furthermore,~~ Sentinel 2 optical data are [here](#) tested for the first time, according to our
5 161 knowledge, [in the context of the detection of forest damages by storms](#). Even if the
6 162 sensitivity of SAR signal to forest ~~damaged damages~~ was previously illustrated by various
7 163 authors (Eriksson et al. 2012; Thiele et al. 2012; Ulander et al. 2005), only very few studies
8 164 exploited SAR for windthrow mapping (Ruetschi et al. 2019; Tanase et al. 2018), possibly
9 165 due to data complexities and limited access to user-friendly processing tools. Thus, the
10 166 present study can be of help to understand how SAR can support forestry practice, also
11 167 considering that these data and related tools are increasingly available by different space
12 168 agencies, and preprocessed ~~S1 datasets are delivered by GEE. With respect to algorithms,~~
13 169 ~~change detection approaches based on pre and post event image differencing were frequently~~
14 170 ~~used in previous research (Dalponte et al. 2020; Ruetschi et al. 2019; Tanase et al. 2018).~~
15 171 ~~Here we experimented with three different supervised classification algorithms, evaluating~~
16 172 ~~the impact of algorithm choice on results, to provided additional knowledge that can guide~~
17 173 ~~the selection of proper methods in operational forest monitoring.~~ Sentinel 1 datasets are
18 174 [delivered by GEE](#).

19 175 The aim of this research is to contribute to forestry practice, developing knowledge useful to
20 176 operational management to exploit ~~freely available~~ satellite [open](#)-data, and defining a
21 177 strategy for the rapid detection of forest damage and the further refinement of information.
22 178 Remote sensing has great potential to cost-efficiently map storm-affected regions, but
23 179 previous research has been somewhat limited, as Sentinel 2 imagery was not previously
24 180 exploited with this aim; and Sentinel 1 was only partially examined. Thus, ~~more studies are~~
25 181 ~~needed~~ [further investigation](#) to ~~translate~~ [assess](#) the ~~remote sensing~~ potential [of integrating](#)
26 182 [satellite open-data](#) into ~~actual practice~~. [forest practical workflows is needed](#). Here the focus is
27 183 on [open-data from the Copernicus free data, preprocessed in programme, exploiting the GEE](#)
28 184 [platform](#) for ~~easiness of application, fast processing, demonstrating that demonstrates to be an~~
29 185 ~~optimal choice for the~~ [this approach can significantly support decision makers with](#) remote
30 186 sensing-based assessment of windthrow damaged areas.

31 187 **Methods**

32 188 *Study area and ground data*

33 189
34 190 The research was conducted in Northern Italy, in areas affected by the windthrow and
35 191 included in two selected Sentinel-2 tiles for which ground truth data were made available by
36 192 local administrations encompassing three regions: Friuli Venezia Giulia, Trentino Alto
37 193 Adige, and Veneto (fig. 1). These regions host important forest resources, and different local
38 194 agencies in charge of their census and management were involved in assessing the Vaia
39 195 impacts. [These data can also be found in open databases \(Forzieri et al. 2020\)](#).

40 196 *Insert fig.1*

41 197 For the Trentino Alto Adige region, ground data for the Trento Autonomous Province were
42 198 provided by the local forest service, and for the Autonomous Province of Bozen by the
43 199 Province authority; in both cases the area of damaged forest were detected on the basis of

1
2
3
4 200 photointerpretation of aerial orthophotos, and integrated with data from field surveys.
5 201 Overall, in Trentino Alto Adige there were 1463 discrete areas, covering 5913 ha, and with
6 202 a mean area of 4 ha. For the Friuli Venezia Giulia region, ground data on forest damage were
7 203 provided by the local forest service using aerial orthophotos and ground surveys; there were
8 204 499 damaged areas for this region, covering 3693 ha, with mean area of 7.4 ha. For the
9 205 Veneto region, the ground data were provided by the Veneto Agency for agriculture
10 206 payments (AVEPA), who are responsible to provide economic help in case of natural
11 207 disasters. AVEPA provided a shapefile of the affected areas based on photointerpretation of
12 208 very high resolution orthophotos (20 cm spatial resolution) and SPOT 6/7 satellite pre and
13 209 post event images at 1.5 m spatial resolution. The dataset included information on
14 210 ~~perimeters~~area borders and estimation of percentage of damaged trees in each area. In total,
15 211 there were 1588 damaged areas detected in Veneto, covering 4020 ha, and having a mean
16 212 surface of 2.5 ha.

17
18
19
20
21 213 The data provided for these Italian regions ~~were~~included 3550 polygons that identify any area
22 214 affected by the windthrow ~~with an area greater than X.X ha~~. These polygons were filtered out
23 215 to create a subset for testing and validation purposes, according to the following inclusion
24 216 criteria: (i) polygons >2 ha, to include areas compatible with the spatial resolution and the
25 217 detection capability of the remote sensing data used in this study; (ii) polygons in which the
26 218 average terrain slope was below 20% in at least 85% of the surface, to exclude areas of
27 219 unreliable SAR signal, according to distortion; (iii) for the Veneto region only, polygons in
28 220 which the amount of damaged trees resulted > 80% ~~(representing >%, that were more than~~
29 221 40% of the total Veneto polygons); (iv) polygons included in two Sentinel tiles, to test the
30 222 methods in the most affected area (3218 polygons). For classification purposes, polygons in
31 223 forest not impacted by the Vaia storm were also drawn in proximity of the impacted
32 224 polygons, by on-screen photointerpretation of post-event Google ~~Earth~~ imagery.

33
34
35
36
37
38 225 In particular, the first criterion was guided by the imagery spatial resolution and allowed to
39 226 retain the larger damaged areas. The second criterion was derived after exploring the SAR
40 227 distortion masks based on local incidence angle. These maps are a by-product of SAR data
41 228 processing, that was additionally performed as these layers are not included in the GEE
42 229 available datasets; ~~the~~ The maps indicated frequent distortions above the 20% slope; ~~to~~
43 230 facilitate the analysis using GEE data, this single threshold was selected. The third criterion
44 231 was applied in Veneto, and was introduced due to the different ways the Italian regions
45 232 assessed damage. In fact, it was noted that ~~in~~ Trentino and Friuli Venezia Giulia ~~assessed~~
46 233 reported only areas ~~were the mostly highly impacted where damage was very significant,~~
47 234 while ~~in~~ Veneto Region also digitized areas that were partially impacted ~~were included in the~~
48 235 census, that reported the damage percentage. Retaining. The dataset was standardized by
49 236 keeping only the polygons with damage >degree above 80% in Veneto ~~allowed to standardize~~
50 237 ~~the dataset among regions, according to visual inspection of forest cover in Google Earth~~
51 238 imagery. The application of the mentioned criteria resulted in a standardized dataset
52 239 including the larger and most affected areas, where SAR data ~~were robust and had the higher~~
53 240 signal to noise ratio and the forest impacts were similar. The ~~damaged~~damaged and
54 241 undamaged non-damaged datasets included a total of 209 polygons, 104 from healthy forest
55
56
57
58
59
60

stands, and 105 from damaged forest areas; the corresponding pixels were extracted from the imagery and averaged at polygon level. In total 90% of the 209 polygons (~~187, namely the calibration dataset~~) were used to calibrate and validate with the k-fold approach classification algorithms. The remaining 22 polygons were used as independent test set for ~~double~~ further evaluation of the overall accuracy. The total area used for calibration, validation and testing the methods was considerable: in Trentino Alto Adige it was equal to 622 ha; in Veneto to 533 ha; and in Friuli 237 ha, representing the different forest types and environmental conditions occurring in the area of interest.

250 *Copernicus Sentinel data*

251 The Sentinel 2 (S2) multispectral images were downloaded from Google Earth Engine as
252 Level-2A orthorectified atmospherically corrected surface reflectance. The S2 Multispectral
253 Instrument (MSI) samples 13 spectral bands: visible and NIR at 10 meters, red edge and
254 SWIR at 20 meters, and atmospheric bands at 60 meters spatial resolution. Only bands at 10 -
255 20 m spatial resolution were used for tests (bands # 2, 3, 4, 5, 6, 7, 8, 8A, 11, 12), resampling
256 at 10 m the 20 m bands with a nearest neighbor approach. The vegetation indices included in
257 Table 1 were also computed.

258 *Insert Table 1*

259 To evaluate the hypothesis ~~concerning the ability of that~~ Sentinel 2 data ~~to correctly can~~ detect
260 the damaged forest areas with significant accuracy, post-event Sentinel 2 images were used.
261 The possibility to use also a pre-damage image and focus the analysis on the variations in
262 reflectance was also evaluated, but confounding factors such as day-specific atmospheric
263 conditions, including cloud cover, and plant phenology stage at different dates were
264 considered ~~important~~ relevant causes of increased uncertainty in results and leading to
265 uncertainty. Therefore, we preferred to work only on post-damage optical images using a
266 binary classification approach (healthy forest/damaged areas).

267 The first available post-event Sentinel 2 imagery is dated June 2019 (7 months after the
268 event). The predictor set named S2_Set1 (set of image bands) was used to evaluate the
269 contribution of bandseach single band, and the S2_Set2 (vegetation indices) ~~contributed to~~
270 evaluate the contribution of the derived vegetation indices, ~~that. The latter combination is~~
271 preferable in case ~~of usage of multiple different~~ images is preferred are used (predictor sets in
272 Table 2).

273 The Sentinel 1 SAR images were downloaded from Google Earth Engine as Ground Range
274 Detected (GRD) scenes, already pre-processed using the Sentinel-1 Toolbox to generate a
275 calibrated, ortho-corrected product at 10 m spatial resolution in dual-band cross polarization
276 mode (VV – VH). Preprocessing included thermal noise removal, radiometric calibration, and
277 terrain correction using a digital terrain model (SRTM 30 m). Five pre-event scenes were
278 collected from the period 26 September - 3 October 2018 (pre-event period without frost or
279 snow), and 5 post-event scenes were from the period 7 - 15 December 2018. The pre and post
280 event scenes were averaged at pixel level, and band ratios (VV/VH, VH/VV), and band
281 normalized differences were also computed (VV-VH, VH -VV). In fact, with respect to the

1
2
3 282 optical images, the SAR data are less affected by atmospheric condition and vegetation
4 283 phenology and for this reason the use of differences between pre and post event was also
5 284 evaluated to detect forest damaged areas.
6
7

8 285 Thus, the set of predictors named S1_Set3 (based only on post-event bands), and the S1_Set4
9 286 (based on pre-post event scenes differences) were used in tests (predictor sets in Table 2).

10
11 287 *Insert Table 2*

12 13 14 288 **Classification approaches**

15
16 289 Three different approaches were tested for the classification task: a generalized linear
17 290 Bayesian model and two machine learning models, the k-Nearest Neighbors and Random
18 291 Forest. Using the three models and the four sets of available predictors (Table 2), a total of 12
19 292 models-predictors combinations were developed. The tests were conducted using the
20 293 RFTrainer, KNN, and Bayesglm R packages in R environment ([R Core Team 2013](#)).

21
22
23 294 Bayesian inference is a method of statistical inference in which Bayes' theorem is used to
24 295 update the probability for a hypothesis as more evidence or information becomes available. It
25 296 facilitates representing and taking full account of the uncertainties related to models and
26 297 parameter values. The Bayesian generalized linear model (BGLM) is based on Bayesian
27 298 functions that finds an approximate posterior mode and variance using extensions of the
28 299 classical generalized linear model computations. The Bayesian function allows the user to
29 300 specify independent prior distributions for the coefficients in the t family, with the default
30 301 being Cauchy distributions with center 0 and scale set to 10 (for the regression intercept), 2.5
31 302 (for binary predictors), or $2.5/(2 \cdot sd)$, where sd is the standard deviation of the predictor in
32 303 the data (for other numerical predictors) (Berrett and Calder 2016; Gelman et al. 2008)

33
34
35
36
37 304 The k-Nearest Neighbors (KNN) technique is a popular method for producing spatially
38 305 contiguous predictions of forest attributes by combining field and remotely sensed data. KNN
39 306 are appealing as they can be used for both univariate and multivariate prediction, no
40 307 assumptions regarding the distributions of response or auxiliary variables are necessary and
41 308 they can be used with a wide variety of datasets (Chirici et al 2016). The k nearest vectors,
42 309 used to perform the classification, are found according to Minkowski distance and the
43 310 classification is performed by means of the maximum of summed kernel densities; both
44 311 ordinal and continuous variables can be predicted (Wu et al. 2002).

45
46
47
48
49 312 Random Forest (RF) is an ensemble of decision trees that learns through a supervised
50 313 approach and produces multiple models that are aggregated, using a bootstrap aggregating
51 314 procedure, to produce the result. The models are built using different training subsets,
52 315 generated by bootstrapping, that are used to build the “forest”. RF is able to reduce the output
53 316 variance and the overfitting problem with respect to other machine learning approaches,
54 317 improving model stability and accuracy (Breiman 2001).

55
56
57 318 When a model is ~~fit~~trained with data there is ~~always~~ the risk of overfitting, i.e. that the
58 319 parameters are estimated to reproduce closely the example training data used, losing the
59 320 capacity to generalize outside the calibration examples. To avoid overfitting one of the most

useful method is k-fold cross validation (k-fold CV) that ~~further~~ splits the training set into K number of subsets, called folds: the models are then iteratively fitted K times each time training the data on data from K-1 of the folds and evaluating the performances on data from the Kth fold. At the end of calibration, the performance on each of the K folds are evaluated in term of Overall Accuracy (OA), i.e. the percentage of cases where the classification as damaged or not damaged was correct. The Overall Accuracy from Cross Validation Overall Accuracy, OA_{cv} and the relative standard deviation $sd(OA_{cv})$ are finally computed averaging the K folds and calculating their standard deviation. This provides more information over how stable the model is by testing it over multiple sets of data.

RF and kNN models require the calibration of hyperparameters, ~~here performed~~. The hyperparameters are calculated using the training polygons datasets. BGLM instead does not require hyperparameters calibration. A procedure based on a random search grid was used for the optimization (Bergstra and Bengio, 2012). The procedure defines a grid of hyperparameter ranges, as those defined above; ~~100~~. One hundred combinations were randomly sampled from the grid and for each combination a k-Fold CV was performed. For both KNN and RF the optimal hyperparameters combination with the greater OA_{cv} was finally selected. In the RF case, the hyperparameters that were tuned ~~includes~~ include the maximum depth of each tree (max_depth) in the forest and the number of features (max_features) considered by each tree when splitting a node. The number of trees in the forest was set equal to 400, while the minimum number of samples required to split an internal node was set equal to 1.

In the kNN case the three hyperparameters that were optimized are: the number of neighbours considered (k), the Minkosky distance, and the kernel to use. The max_features ranged between 2 and n, where n equals the number of predictors used in input; the max_depth ranges between 1 and 40. k ranged between 1 and 60, the Minkosky distance was equal to the Euclidean and Manhattan distances, and the kernel to use were Unweighted, Weighted, Inverse, Reciprocal.

The models, once optimized and validated with k-fold approach, were further independently evaluated in terms of accuracy using the test set that includes 10% of the polygons never used during the optimization procedure. For the evaluation of the classification results, different statistics are reported, including overall accuracy, users accuracy, producers accuracy, and the percentage of omission and commission errors. User's accuracy represents how reliable the classification is in terms of actually finding damage in the real world over an area that was classified as "damaged" in the map. Producer's accuracy reports how often a damage that is found in the real world is reported in the final classified map (Cohen 1968; Congalton 1991).

Results

In table 3 we present for each model and for each of the four sets of predictors the averaged OA_{cv} obtained with a 9-fold cross validation procedure and the related standard deviation

1
2
3 362 obtained averaging the different iterations. For the KNN and RF models the best
4 363 hyperparameters combination, identified using a RandomSearch algorithm are also presented.

5
6
7 364 *Insert Table 3*

8
9 365 The results in Table 3, also graphically shown in fig. 2, show consistency in different models,
10 366 with negligible differences among the considered approaches and limited variance from
11 367 different iterations.

12
13 368 The best results are obtained with S2 images, with OA_{cv} always > 0.8 and included in the 0.8-
14 369 0.85 range for either bands or vegetation indices, with the latter reaching a slightly higher
15 370 accuracy. The standard deviation values resulted were always small with a maximum value of
16 371 0.102. This confirms the ability of S2 to detect impacted forest areas.

17
18
19
20 372 Lower accuracy results –in the 0.6-0.7 OA_{cv} range- are obtained when using SAR data with a
21 373 slightly better scores obtained when using the S1_Set3, that includes only data from post-
22 374 event scene.

23
24
25 375 *Insert fig. 2*

26
27 376 The three models were also applied to the independent test set ($n = 22$) to evaluate their final
28 377 performance on new and unseen data and the results are shown in Table 4.

29
30 378 *Insert Table 4*

31
32 379 The results obtained using the independent test set are similar to those obtained with 9-fold
33 380 CV but span, as expected, over a slightly higher range, ~~as expected~~ considering the limited
34 381 number of samples in the test set ($n=22$).

35
36
37 382 For sets 1 and 2, based on S2 data, the accuracy is included in the 0.77-0.86 range, and is
38 383 similar across the three different models. For sets 3 and 4, based on S1 data, the accuracy
39 384 range is 0.5-0.68, with higher results obtained using Random Forest model.

40
41
42 385 Overall, results in Table 4 confirm the higher accuracies obtained with the use of Sentinel 2
43 386 data with respect to those obtained with SAR data.

44
45 387

46
47 388

48
49 389

50 51 52 390 **Discussion**

53
54 391 The availability of satellite open-and-data that is also partly preprocessed ~~satellite data is~~
55 392 erueial-for-adds significant value to the procedure of the assessment of forest damages, from
56 393 windthrow, as well as of or other forestsources of damage, allowing- that change the
57 394 landscape. The final product is a classified damage map that supports rapid responses in
58 395 terms of forest management ~~response. Overall, the. The~~ results ~~shown~~ indicate that data from

the Copernicus Sentinel 1 and 2 ~~datasets~~missions are suited for the detection of damaged forest areas. SAR is especially useful for a ~~preliminary~~fast evaluation ~~aimed at fast intervention and, providing useful information for~~ immediate ~~/short-term response actions for~~ risk mitigation. Sentinel 2 can be used to refine the SAR initial information unless post event data are immediately available. The use of cloud-based platforms like Google Earth Engine ~~resulted an optimal choice, allowing the user to avoid helps to reduce the long and often complex time that operators need for image download and standard pre-processing of.~~ A pre-defined workflow over ready-to-use imagery, and providing data can avoid requiring highly skilled operators for processing imagery. The workflow can be partly automatic, providing maps useful to multiple end-users, even those less familiar with image processing techniques.

~~Overall, focusing~~Focusing on applications, the present research suggests that the sequential use of GEE Sentinel 1 and 2 data for better windthrow information provision is an optimal combination. Specifically, the testing of the different predictors from S1 and S2 data provided useful insights on the advantages and limits of these datasets.

The best detection of the forest areas impacted by the Vaia storm is always obtained using Sentinel 2 images. Using a 9-fold cross validation approach and either S2 bands or vegetation indices as input, the obtained overall accuracies were > 80%, with limited differences among modeling approaches, low variance from iterations, and results included in the 80-85% range. The use of vegetation indices with KNN and RF approaches provided the higher OA_{CV} values, equal to 85 and 84 %, respectively.

Very similar results are obtained when the parameterized models were validated against the independent test set, represented by 22 samples not used to calibrate the models. The obtained results, although the number of test samples is relatively low, are in a very similar OA range (77- 86%) compared with those reported for 9-fold cross validation. The highest OA score (86%) is obtained either using S2 bands with KNN model or using VIs with Random Forest. ~~From the user's perspective, thus considering how often a given class predicted by the model will actually be present on the ground, over 90% of the damaged areas were correctly classified in the different S2-based tests. From the producer's perspective, thus considering how often the real features on the ground are correctly classified by the model, damaged areas accuracy is in the 71-83% range, with higher scores obtained with RF and VIs. User accuracies were over 90% whereas producer's accuracies were in the 71-83% range, with higher scores obtained with RF and VIs. This indicates that commission errors were lower than omission errors, in other words some damaged areas where not correctly detected by the classifier, thus leaving out some areas from the final map, but most of the areas classified as damaged were really damaged. It might be due to canopy of felled trees still significantly showing in the image, or also reflectance from water vapor, that commonly rises in the morning in mountainous areas, mixing with the reflectance values from the tree trunks.~~

It should be considered that the post-event S2 imagery employed in the study were from a single date and about 7 months after the Vaia event. The time gap between the storm and the

1
2
3 437 S2 imagery allows the greening of the ground in damaged areas, from herbal and shrubs
4 438 vegetation regrowth that can produce a change in reflectance values and a consequent
5 439 negative impact in the classification accuracy. The results are however in line with what has
6 440 been already found in the past with Landsat data, thus at 30 m spatial resolution (30 m): an
7 441 OA equal to 86% was reached in the detection of windthrows in Voges mountains in France
8 442 (Haidu et al. 2019); in European Russia and United States the OA was about 75%, with more
9 443 accurate results reported for larger areas (Baumann et al. 2014); and with an automatic
10 444 algorithm based on Landsat time series, historical disturbance from windthrow and logging
11 445 was detected in United States forest with an accuracy about 80%. According to our
12 446 knowledge, there are no studies based on the use of S2 data for forest windthrow detection,
13 447 except two abstracts where the accuracy of the obtained results is not reported (Cenci et al.
14 448 2019; Valt et al. 2019).

15 449 Vegetation indices are especially useful when multiple images are used (as in the case of
16 450 change detection analysis), or when the study area is large and covered by different **image**
17 451 tiles, or even when a mosaic from different dates is composed to mitigate cloud cover issues.
18 452 In fact, VIs are designed to maximize sensitivity to the vegetation characteristics while
19 453 minimizing confounding factors such as soil background reflectance, directional, or
20 454 atmospheric effects, that change among different acquisitions (Fang and Liang 2008).
21 455 According to this and based on the reported classification results, the use of VIs from S2 data,
22 456 **fitting the model with the RF model approach method**, appears the best solution to detect
23 457 damaged forest area in this case study. An improvement in accuracy is expected if optical
24 458 imagery **becomebecomes** available in dates close to the time of the damage event.

25 459 The use of SAR inputs produced lower accuracies compared to S2 inputs with an OA values
26 460 in the range 0.66-0.71 according to 9-fold cross validation and a low standard deviation score
27 461 (< 0.14). Differences among the three models are minor, as well as those when using post-
28 462 event data only or pre-post event backscattering difference. The highest $OA_{cv}OA_{cv}$ score is
29 463 obtained with pre-event data and KNN approach ($OA_{cv}OA_{cv} = 0.71$).

30 464 According to results from the independent test set, the OAs slightly decrease and the
31 465 variability in values increases, being included in the 0.5-0.68 range. The best scores are
32 466 obtained using RF: 64% and 68% in OA with post-event data and pre-post event difference,
33 467 respectively. With RF, the user's accuracy is low (54%) when using S1 pre-post event scenes
34 468 differences, but the producer's accuracy is in line with the one obtained using the S2 data
35 469 (75%). When using post-event data, the user's and producer's accuracies are on the same
36 470 order (63%). The results suggest the use of RF as classification model, and the combination
37 471 of pre- and post-event SAR scenes to better meet the user's needs.

38 472 Previous studies conducted with C-band ERS 1/2 and RADARSAT 1 at 30 m spatial
39 473 resolution were not successful for the detection of forest windthrows (Schwarz et al. 2003;
40 474 Ulander et al. 2005). However, with L-band data - that better penetrates into the forest - OAs
41 475 included in the 69-84% range were obtained in the Bavarian Forest National Park, with
42 476 accuracy values depending on the acquisition date and environmental conditions (Tanase et
43 477 al. 2018). Using C-band S1 and a change detection approach, the producer's accuracy

reached 88% in a German validation site, but the user's accuracy was quite low (21%). ~~%)~~ and limitations consisted in a minimum area of 0.5 ha and the requirement of 10 post-event images (Rüetschi et al. 2019). Positive results were also obtained using X-band data with very high spatial resolution (Thiele et al. 2012).

The results here presented outline the relevance of SAR spatial resolution for forest windthrow detection, and confirm the ability of S1 data to produce fast preliminary information on impacted areas, with the obtained OA and user's accuracy included in a range of values similar to those reported by other studies.

It is important to note ~~that there are the~~ limitations ~~in~~ of the present study: in terms of suitability in certain cases. First, the ~~application of the models to~~ areas tested were all above 2 ha, to cope with the 10 m spatial resolution of both S1 and S2 datasets; however the average size of damaged areas in the three considered regions resulted higher than 2 ha. Then, the ground truth was filtered out to exclude slopes using a low threshold ($> 20^\circ$) where the first SAR distortion effects were observed. Producing the distortion masks to filter out data, instead of a fixed threshold that excluded most slopes for easiness of analysis in GEE, could result in a detailed mapping of unreliable SAR pixels and a larger availability of goodreliable data over slopes. Similarly, the use of temporal series can improve the amount of area with reliable SAR data, as at each pass the acquisition angle may changevary. When SAR data are used, it is also important to ~~evaluate the presence of detect and mask pixels with~~ wet or dry snow over the canopy, as it ~~impactschanges~~ the backscatteringbackscatter values (Koskinen et al. 1997). This implies ~~that an added complexity of the method~~ in areas seasonally covered by ~~wet snow a data quality assessment has to be performed in certain periods~~.

~~This study aimed at a first evaluation of GEE Copernicus S1 and S2 data for windthrow monitoring, to facilitate remote sensing data exploitation in applied forestry.~~ For future operational use, the application of a pre-disturbance forest/non-forest map can help to perform semi-automatic classification. Further tests are also needed to understand the response of satellite data over less impacted forests, where a mixture of healthy and damaged trees is present, and how to minimize the impact of SAR distortion areas.

The combined use of S1 and S2 was not investigated here, as for data integration the optical and radar imagery ~~shallshould~~ be from same period. ~~In~~ The S2 images used in this investigation were dated months after the storm, when the herbal and ~~shrubs~~shrub vegetation ~~regrowthrenovation~~ influence both the optical reflectance and the backscattering in C-band SAR: ~~too many. These are~~ confounding factors would have been present, but a data integration approach is feasible if optical data are available soon after the event. Higher accuracy in classification is known to occur from combining SAR ~~with~~ optical joined use with respect to use single datasets usesensor type (Clerici et al. 2017; Vaglio Laurin et al. 2012), so this might be another strategy to improve the information accuracy.

516 Conclusion

517 This study showed the suitability of ~~GEE~~ Copernicus S1 and S2 data for the detection of
518 areas affected by windthrow. Sentinel 2 provided the best performance for detection of
519 windthrow areas, but its use was seriously hampered by cloud cover. For events occurring in
520 winter, Sentinel 2 data might only be available after several months. In those cases, the use
521 of Sentinel 1 data, being independent with respect to atmospheric condition and with a fast
522 return time, becomes the best option for a first and rapid evaluation of the forest damage, to
523 support field operations and the formation of management response plans.

524 Thus, for operational monitoring, the results suggest a sequential approach, based initially on
525 S1 for fast response. This initial SAR assessment can be refined in later dates, integrating S2
526 imagery when available and data from ground or aerial surveys, for a more accurate mapping
527 also over steep slopes.

528 X

529 Data availability statement

530 Remote sensing data are freely available by the Copernicus facilities. Ground data were
531 provided by local authorities and can be requested directly to them. Part of the data can be
532 found in the open database published by Forzieri et al. (2020).

533 **Funding**

534 This work was supported by internal resources of the participating institutes. In particular F.P.'s work
535 was supported by University of Padova's VAIA-FRONT (VAIA, FROM lessons learnt to future
536 options) project.

537 **Acknowledgements**

538 For the Trentino Autonomous Province, we acknowledge Dr. A. Wolynski from the local forest
539 service who coordinated the acquisition of ground data on the basis of aerial photography and field
540 surveys.

541 For the Alto Adige region, we acknowledge Dr. F. Maistrelli from the Autonomous Province of
542 Bozen who provided the ground data delimited on the basis of aerial photography and field surveys.

543 For the Friuli Venezia Giulia region, we acknowledge Dr. R. Comino from the local forest service
544 who coordinated the acquisition of ground data on the basis of aerial photography and ground
545 surveys.

546 For the Veneto Region we acknowledge the Agenzia Veneta per i Pagamenti in Agricoltura (AVEPA)
547 that provided the ~~ground~~-data.

548 **Conflict of interest statement**

549 'None declared.'

550 **References**

- 1
2
3 551 Baumann, M.; Ozdogan, M.; Wolter, P.T.; Krylov, A.; Vladimirova, N.; Radeloff, V.C.
4 552 (2014) Landsat remote sensing of forest windfall disturbance. *Remote Sens. Environ.*, 143,
5 553 171–179.
- 6
7
8 554 Bergstra, J., & Bengio, Y. (2012). Random search for hyper-parameter optimization. *Journal*
9 555 *of machine learning research*, 13, 281-305.
- 10
11 556 Berrett, C., & Calder, C. A. (2016). Bayesian spatial binary classification. *Spatial Statistics*,
12 557 16, 72-102.
- 13
14
15 558 Breiman, L. (2001). Random forests. *Machine learning*, 45(1), 5-32.
- 16
17 559 Cenci, L., De Giorgi, A., Squicciarino, G., Pulvirenti, L., Moser, G., & Boni, G. (2019).
18 560 Exploiting Sentinel 2 data for mapping windstorm damages in forested areas. Case Study: the
19 561 event of October 2018 occurred in Northeast Italy. *Geophysical Research Abstracts*, vol. 21.
- 20
21
22 562 Chirici, G., Mura, M., McNerney, D., Py, N., Tomppo, E. O., Waser, L. T., ... & McRoberts,
23 563 R. E. (2016). A meta-analysis and review of the literature on the k-Nearest Neighbors
24 564 technique for forestry applications that use remotely sensed data. *Remote Sensing of*
25 565 *Environment*, 176, 282-294.
- 26
27
28 566 Chirici, G., Bottalico, F., Giannetti, F., Del Perugia, B., Travaglini, D., Nocentini, S., ... &
29 567 Fattorini, L. (2018). Assessing forest windthrow damage using single-date, post-event
30 568 airborne laser scanning data. *Forestry: An International Journal of Forest Research*, 91(1), 27-
31 569 37.
- 32
33
34 570 Chirici, G., Giannetti, F., Travaglini, D., Nocentini, S., Francini, S., D'Amico, G., ... &
35 571 Tonner, J. (2019). Stima dei danni della tempesta “Vaia” alle foreste in Italia. *Forest@-*
36 572 *Journal of Silviculture and Forest Ecology*, 16(1), 3.
- 37
38
39 573 Clerici, N., Valbuena Calderón, C. A., & Posada, J. M. (2017). Fusion of Sentinel-1A and
40 574 Sentinel-2A data for land cover mapping: a case study in the lower Magdalena region,
41 575 Colombia. *Journal of Maps*, 13(2), 718-726.
- 42
43
44 576 Cohen, J. (1968) Weighted kappa: Nominal scale agreement with provision for scaled
45 577 disagreement or partial credit. *Psychological Bulletin* 70 (4):213-220.
- 46
47
48 578 Congalton, R.G. (1991) A review of assessing the accuracy of classification of remotely
49 579 sensed data. *Remote Sensing of Environment* 37:35-46
- 50
51 580 Dalponte, M., Marzini, S., Solano-Correa, Y. T., Tonon, G., Vescovo, L., & Gianelle, D.
52 581 (2020). Mapping forest windthrows using high spatial resolution multispectral satellite
53 582 images. *International Journal of Applied Earth Observation and Geoinformation*, 93, 102206.
- 54
55
56 583 Drusch, M., Del Bello, U., Carlier, S., Colin, O., Fernandez, V., Gascon, F., ... & Meygret, A.
57 584 (2012) Sentinel-2: ESA's optical high-resolution mission for GMES operational services.
58 585 *Remote sensing of Environment*, 120, 25-36.
- 59
60

- 1
2
3 586 Duan, F.; Wan, Y.; Deng, L. (2017) A novel approach for coarse-to-fine windthrown tree
4 587 extraction based on unmanned aerial vehicle images. *Remote Sens.*, 9, 306.
- 6
7 588 Duelli, P., Wermelinger, B., Moretti, M., & Obrist, M. K. (2019). Fire and windthrow in
8 589 forests: Winners and losers in Neuropterida and Mecoptera. *Alpine Entomology*, 3, 39.
- 10
11 590 Dyukarev, E.A.; Pologova, N.N.; Golovatskaya, E.A.; Dyukarev, A.G. (2011) Forest cover
12 591 disturbances in the South Taiga of West Siberia. *Environ. Res. Lett.*, 68.
- 13
14 592 Einzmann, K., Immitzer, M., Böck, S., Bauer, O., Schmitt, A., & Atzberger, C. (2017)
15 593 Windthrow detection in European forests with very high-resolution optical data. *Forests*, 8(1),
16 594 21.
- 18
19 595 Ellison, A. M. et al. (2005) Loss of foundation species: Consequences for the structure and
20 596 dynamics of forested ecosystems. *Front. Ecol. Environ.* 3, 479–486.
- 22
23 597 Eriksson, L.E.B.; Fransson, J.E.S.; Soja, M.J.; Santoro, M. (2012) Backscatter signatures of
24 598 wind-thrown forest in satellite SAR images. In *Proceedings of the International Geoscience
25 599 and Remote Sensing Symposium (IGARSS)*, Munich, Germany, 22–27 July; pp. 6435–6438.
- 26
27 600 Fang H. and S. Liang. (2008) Leaf Area Index Models. *Encyclopedia of Ecology*, pp. 2139-
28 601 2148
- 30
31 602 [Forzieri, G., Pecchi, M., Girardello, M., Mauri, A., Klaus, M., Nikolov, C., ... Beck, P. S. A.](#)
32 603 [\(2020\). A spatially explicit database of wind disturbances in European forests over the period](#)
33 604 [2000 - 2018. *Earth System Science Data*, 12\(1\), 257–276.](#)
- 35
36 605 Gardiner, B., Blennow, K., Carnus, J. M., Fleischer, P., Ingemarsson, F., Landmann, G., ... &
37 606 Peyron, J. L. (2010). Destructive storms in European forests: past and forthcoming impacts.
38 607 European Forest Institute, Efiatlantic.
- 39
40 608 Gelman, A., Jakulin, A., Pittau, M. G., & Su, Y. S. (2008) A weakly informative default prior
41 609 distribution for logistic and other regression models. *The annals of applied statistics*, 2(4),
42 610 1360-1383.
- 44
45 611 Gorelick, N., Hancher, M., Dixon, M., Ilyushchenko, S., Thau, D., & Moore, R. (2017)
46 612 Google Earth Engine: Planetary-scale geospatial analysis for everyone. *Remote sensing of
47 613 Environment*, 202, 18-27.
- 49
50 614 Green, R.M. (1998) The sensitivity of SAR backscatter to forest windthrow gaps. *Int. J.*
51 615 *Remote Sens.*, 19, 2419–2425.
- 52
53 616 Haidu, I., Furtuna, P. R., & Lebaut, S. (2019) Detection of old scattered windthrow using low
54 617 cost resources. The case of Storm Xynthia in the Vosges Mountains, 28 February 2010. *Open
55 618 Geosciences*, 11(1), 492-504.
- 57
58 619 Hamdi, Z. M., Brandmeier, M., & Straub, C. (2019) Forest Damage Assessment Using Deep
59 620 Learning on High Resolution Remote Sensing Data. *Remote Sensing*, 11(17), 1976.

- 1
2
3 621 Havašová, M., Ferencík, J., & Jakuš, R. (2017) Interactions between windthrow, bark beetles
4 622 and forest management in the Tatra national parks. *Forest Ecology and Management*, 391,
5 623 349-361.
- 6
7
8 624 Honkavaara, E.; Litkey, P.; Nurminen, K. (2013) Automatic Storm Damage Detection in
9 625 Forests Using High-Altitude. Photogrammetric Imagery. *Remote Sens.*, 5, 1405–1424.
- 10
11 626 Huang, C., Goward, S. N., Masek, J. G., Thomas, N., Zhu, Z., & Vogelmann, J. E. (2010) An
12 627 automated approach for reconstructing recent forest disturbance history using dense Landsat
13 628 time series stacks. *Remote Sensing of Environment*, 114(1), 183-198.
- 14
15
16 629 Jackson, R. G., Foody, G. M., & Quine, C. P. (2000) Characterizing windthrown gaps from
17 630 fine spatial resolution remotely sensed data. *Forest Ecology and Management*, 135(1-3), 253-
18 631 260.
- 19
20
21 632 Jonikavičius, D., & Mozgeris, G. (2013) Rapid assessment of wind storm-caused forest
22 633 damage using satellite images and stand-wise forest inventory data. *iForest-Biogeosciences*
23 634 *and Forestry*, 6(3), 150.
- 24
25
26 635 Kislov, D. E., & Korznikov, K. A. (2020) Automatic Windthrow Detection Using Very-
27 636 High-Resolution Satellite Imagery and Deep Learning. *Remote Sensing*, 12(7), 1145.
- 28
29
30 637 Koskinen, J.T.; Pulliainen, J.T.; Hallikainen, M.T. (1997) The Use of ERS-1 SAR Data in
31 638 Snow Melt Monitoring. *IEEE Trans. Geosci. Remote Sens.* 35, 601–610.
- 32
33 639 [Marchi, N., Pirotti, F., & Lingua, E. \(2018\). Airborne and Terrestrial Laser Scanning Data for](#)
34 640 [the Assessment of Standing and Lying Deadwood: Current Situation and New Perspectives.](#)
35 641 [Remote Sensing, 10\(9\), 1356.](#)
- 36
37
38 642 McCarthy, J. K., Hood, I. A., Kimberley, M. O., Didham, R. K., Bakys, R., Fleet, K. R., ... &
39 643 Brockerhoff, E. G. (2012) Effects of season and region on sapstain and wood degrade
40 644 following simulated storm damage in *Pinus radiata* plantations. *Forest ecology and*
41 645 *management*, 277, 81-89.
- 42
43
44 646 Mokroš, M.; Výbošťok, J.; Merganič, J.; Hollaus, M.; Barton, I.; Koren, M.; Tomaščík, J.;
45 647 Černá, J. (2017) Early stage forest windthrow estimation based on unmanned aircraft
46 648 system imagery. *Forests* 8, 306.
- 47
48
49 649 Pecchi, M., Forzieri, G., Ceccherini, G., Spinoni, J., Feyen, L., Cescatti, A., & Chirici, G.
50 650 (2019) Pan-European mapping of windthrows. *Geophysical Research Abstracts*, vol. 21.
- 51
52 651 [Pirotti, F., Travaglini, D., Giannetti, F., Kutchartt, E., Bottalico, F., & Chirici, G. \(2016\).](#)
53 652 [Kernel feature cross-correlation for unsupervised quantification of damage from windthrow](#)
54 653 [in forests. ISPRS - International Archives of the Photogrammetry, Remote Sensing and](#)
55 654 [Spatial Information Sciences, XLI-B7, 17–22.](#)
- 56
57
58
59
60

- 655 Poursanidis, D., & Chrysoulakis, N. (2017) Remote Sensing, natural hazards and the
656 contribution of ESA Sentinels missions. *Remote Sensing Applications: Society and*
657 *Environment*, 6, 25-38.
- 658 [R Core Team \(2013\). R: A language and environment for statistical computing. R Foundation
659 for Statistical Computing, Vienna, Austria. URL <http://www.R-project.org/>.](http://www.R-project.org/)
- 660 Rüetschi, M., Small, D., & Waser, L. T. (2019) Rapid detection of windthrows using
661 sentinel-1 c-band sar data. *Remote Sensing*, 11(2), 115.
- 662 Saad, C., Boulanger, Y., Beaudet, M., Gachon, P., Ruel, J. C., & Gauthier, S. (2017) Potential
663 impact of climate change on the risk of windthrow in eastern Canada's forests. *Climatic*
664 *Change*, 143(3-4), 487-501.
- 665 Schwarz, M., Steinmeier, C., Holecz, F., Stebler, O., & Wagner, H. (2003) Detection of
666 windthrow in mountainous regions with different remote sensing data and classification
667 methods. *Scandinavian journal of forest research*, 18(6), 525-536.
- 668 Seidl, R., Schelhaas, M. J., Rammer, W., & Verkerk, P. J. (2014) Increasing forest
669 disturbances in Europe and their impact on carbon storage. *Nature climate change*, 4(9), 806-
670 810.
- 671 Solimini, D. (2016) *Understanding Earth Observation*; Springer International Publishing:
672 Basel, Switzerland, pp. 1–703, ISBN 978-3-319-25633-7.
- 673 Tanase, M. A., Aponte, C., Mermoz, S., Bouvet, A., Le Toan, T., & Heurich, M. (2018)
674 Detection of windthrows and insect outbreaks by L-band SAR: A case study in the Bavarian
675 Forest National Park. *Remote Sensing of Environment*, 209, 700-711.
- 676 Thiele, A., Boldt, M., & Hinz, S. (2012) Automated detection of storm damage in forest areas
677 by analyzing TerraSAR-X data. In: 2012 IEEE International Geoscience and Remote Sensing
678 Symposium (pp. 1672-1675). IEEE.
- 679 Torres, R.; Snoeij, P.; Geudtner, D.; Bibby, D.; Davidson, M.; Attema, E.; Potin, P.;
680 Rommen, B.Ö.; Floury, N.; Brown, M.; et al. (2012) GMES Sentinel-1 mission. *Remote*
681 *Sens. Environ.*, 120, 9–24.
- 682 Ulander, L.M.H.; Smith, G.; Eriksson, L.; Folkesson, K.; Fransson, J.E.S.; Gustavsson, A.;
683 Hallberg, B.; Joyce, S.; Magnusson, M.; Olsson, H.; et al. (2005) Mapping of wind-thrown
684 forests in southern Sweden using space- and airborne SAR. In *Proceedings of the*
685 *International Geoscience and Remote Sensing Symposium (IGARSS)*, Seoul, South Korea,
686 25–29 July 2005; pp. 3619–3622.
- 687 Usbeck, T., Wohlgemuth, T., Dobbertin, M., Pfister, C., Bürgi, A., & Rebetz, M. (2010).
688 Increasing storm damage to forests in Switzerland from 1858 to 2007. *Agricultural and Forest*
689 *Meteorology*, 150(1), 47-55.

1
2
3 690 Vaglio Laurin, G., Liesenberg, V., Chen, Q., Guerriero, L., Del Frate, F., Bartolini, A., ... &
4 691 Valentini, R. (2013) Optical and SAR sensor synergies for forest and land cover mapping in a
5 692 tropical site in West Africa. *International Journal of Applied Earth Observation and*
6 693 *Geoinformation*, 21, 7-16.

7
8
9 694 Valt, M., Salvatori, R., & Salzano, R. (2019) Mapping the effects of VAIA storm using
10 695 Sentinel 2 data. *Geophysical Research Abstracts*, vol. 21.

11
12
13 696 Woodhouse, I. H. (2005). *Introduction to microwave remote sensing*. CRC press.

14
15 697 Wu, Y., Ianakiev, K., & Govindaraju, V. (2002) Improved k-nearest neighbor classification.
16 698 *Pattern recognition*, 35(10), 2311-2318.

17
18
19 699

20
21 700

22
23 701

24
25 702

26
27 703

28
29 704

30
31 705

32
33 706

34
35 707

36
37 708

38
39 709

40
41 710

42
43 711

44
45 712

46
47 713

48
49 714

50
51 715

52
53 716

54
55 717

56
57 718

1
2
3 **719 Table and Figure captions**
4

5 **720 Figure 1** Study area, Northern Italy; in red the 209 polygons used for calibration, cross
6 validation and testing; in orange the rectangles of Sentinel 2 tiles. Image prepared using
7 Google Earth Pro Landsat/Copernicus @2020 GeoBasis-DE/BKG US Dept of State
8 Geographer @2020 Google.
9
10

11 **724 Figure 2** Classification results using three algorithms (BGLM = Bayesian Generalized Linear
12 Model; KNN = K-Nearest Neighbor; RF = Random Forests) with four sets of predictors as
13 input (Set 1 = Sentinel 2 bands; Set 2 = Sentinel 2 vegetation indices; Set 3 = Sentinel 1 post
14 event data; Set 4 = Sentinel 1 pre-post event difference data). The models were validated with
15 a 9-fold cross validation approach.
16
17
18

19 729

20 730

21 731

22 732

23 733

24 734

25 735

26 736

27 737

28 738

29 739

30 740

31 741

32 742

33 743

34 744

35 745

36 746

37 747

38 748

749 **Table 1.** List of vegetation indices used for tests.

	Vis	Index name	Bands
1	NDVI_idx	Normalized Difference Veg index	$(b8-b4)/(b8+b4)$
2	NBr_idx	Normalized Burn Ratio	$(b8 - b12)/(b8 + b12)$
3	NDVI_2	Normalized Difference Veg index 2	$(b12-b8)/(b12+b8)$
4	SR	Simple Ratio	$b8/b4$
5	ARI1	Anthocyanin Reflectance Index 1 (ARI1)	$1/b3-1/b5$
6	EVI	Enhanced Vegetation Index	$2.5*(b8-b4)/(b8+6*b4-7.5*b2)+1000$
7	NDMI	Normal difference moisture index	$(b8-b11)/(b8+b11)$
8	MSI	Moisture soil index	$b11/b8$
9	BAI	Burn Area Index	$1/(0.1-b4)^2+(0.06-b8)^2$
10	DVI	Difference Veg Index	$b8-b4$
11	GDVI	Green Difference Vegetation Index	$b8 - b3$
12	GARI	Green Atmospherically Resistant Index	$b8-(b3-(b2-b4)/b8+(b3-(b2-b4)$
13	GRVI	Green Ratio Vegetation Index	$b8/b3$
14	IPVI	Infrared Percentage Vegetation Index	$b8/b8+b4$

750

751

752

753

754

755

756

757

758 **Table 2.** Set of S2 and S1 predictors used in classification models.

Set	Predictors	Image date
S2_Set1	Sentinel 2 (post event) bands	28/06/2019
S2_Set2	Sentinel 2 (post event) Vegetation Indices	28/06/2019
S1_Set3	Sentinel 1 (post event) bands VH, VV	07-15/12/2018
	Sentinel 1 (post event) band ratios VV/VH, VH/VV	07-15/12/2018
	Sentinel 1 (post event) normalized difference VV-VH, VH-VV	07-15/12/2018
S1_Set4	Sentinel 1 (pre-post event difference) bands VH, VV	26/09- 03/10/2018
		07-15/12/2018
	Sentinel 1 (post event difference) band ratios VV/VH, VH/VV	26/09- 03/10/2018
		07-15/12/2018
	Sentinel 1 (pre-post event difference) normalized difference VV-VH, VH-VV	26/09- 03/10/2018
		07-15/12/2018

759

760

761

762

763

764

765

766

767

768

769

770

771 **Table 3.** Overall accuracy for classification models validated with 9-fold cross validation
 772 (OA_{cv}), with related standard deviation *sd*(OA_{cv}) and best hyperparameters combination.
 773 The highest Accuracy for each model is shown in bold.

		Best hyperparameters		9-fold-cross validation	
BGLM	<i>Predictors set</i>			OA_{cv}	<i>sd</i> (OA _{cv})
	S2_Set1			0.80	0.086
	S2_Set2			0.82	0.073
	S1_Set3			0.68	0.072
	S1_Set4			0.67	0.096
KNN		<i>kmax</i>	<i>Distance</i>		
	S2_Set1	15	2	0.82	0.081
	S2_Set2	20	1	0.85	0.102
	S1_Set3	53	2	0.71	0.138
	S1_Set4	12	2	0.66	0.085
RF		<i>Max. features</i>	<i>Max. depth</i>		
	S2_Set1	4	11	0.83	0.070
	S2_Set2	8	27	0.84	0.064
	S1_Set3	3	40	0.66	0.075
	S1_Set4	4	35	0.66	0.089

774

775

776

777

778

779 **Table 4.** Accuracy statistics for the three classification models and the four set of predictors
 780 obtained on the Test set (10% of samples). In bold the highest OA obtained.

	Predictor set	Overall accuracy	Producers accuracy Healthy forest %	Producers accuracy Damaged areas %	Users accuracy Healthy forest %	Users accuracy Damaged areas %
BGLM	S2_Set1	0.77	87.50	71.43	63.64	90.91
	S2_Set2	0.82	100.00	73.33	63.64	100.00
	S1_Set3	0.50	50.00	50.00	27.27	72.73
	S1_Set4	0.55	54.55	54.55	54.55	54.55
KNN	S2_Set1	0.86	100.00	78.57	72.73	100.00
	S2_Set2	0.82	88.89	76.92	72.73	90.91
	S1_Set3	0.50	50.00	50.00	45.45	54.55
	S1_Set4	0.64	61.54	66.67	72.73	54.55
RF	S2_Set1	0.82	88.89	76.92	72.73	90.91
	S2_Set2	0.86	90.00	83.33	81.82	90.91
	S1_Set3	0.64	63.64	63.64	63.64	63.64
	S1_Set4	0.68	64.29	75.00	81.82	54.55

781