

1 Development of aerial photos and LIDAR data approaches to map spatial and temporal evolution of
2 ditch networks in peat dominated catchments

3 Joy Bhattacharjee^{1*}, Hannu Marttila², Ali Torabi Haghighi³, Miia Saarimaa⁴, Anne Tolvanen⁵, Ahti
4 Lepistö⁶, Martyn N Futter⁷, Bjørn Kløve⁸

5 ¹PhD Candidate, Water, Energy and Environmental Engineering Research Unit, PO Box 4300, FI-
6 90014 University of Oulu, Finland. joy.bhattacharjee@oulu.fi

7 ²Professor (Assistant), Water, Energy and Environmental Engineering Research Unit, PO Box 4300,
8 FI-90014 University of Oulu, Finland. hannu.marttila@oulu.fi

9 ³Professor (Associate), Water, Energy and Environmental Engineering Research Unit, PO Box 4300,
10 FI-90014 University of Oulu, Finland. ali.torabihaghighi@oulu.fi

11 ⁴Leading expert in natural science, Finnish Forest Center, FI-90400 Oulu, Finland.
12 miia.saarimaa@metsakeskus.fi

13 ⁵Professor, Natural Resources Institute Finland (Luke), P.O. Box 413, FI-90014 University of Oulu,
14 Finland. anne.tolvanen@luke.fi

15 ⁶Senior Research Scientist, Finnish Environment Institute (SYKE), P.O. Box 140, FI-00251 Helsinki,
16 Finland. ahti.lepisto@ymparisto.fi

17 ⁷Professor (Associate), Department of Aquatic Sciences and Assessment, Swedish University of
18 Agricultural Sciences SLU, P.O. Box 7050, SE-75007 Uppsala, Sweden. martyn.futter@slu.se

19 ⁸Professor, Water, Energy and Environmental Engineering Research Unit, PO Box 4300, FI-90014
20 University of Oulu, Finland. bjorn.klove@oulu.fi

21 *Corresponding author at: Water, Energy and Environmental Engineering Research Unit, PO Box
22 4300, 90014 University of Oulu, Finland. E-mail address: joy.bhattacharjee@oulu.fi

23 **Abstract**

24 Spatiotemporal information on historical peatland drainage is needed to relate past land use to
25 observed changes in catchment hydrology. Comprehensive knowledge of historical development of
26 peatland management is largely unknown at catchment scale. Aerial photos and LIDAR data enlarge
27 the possibilities for identifying past peatland drainage patterns. Here, our objectives are: (1) to
28 develop techniques for semi-automatically mapping the location of ditch networks in peat-dominated
29 catchments by using aerial photos and LIDAR data, and (2) to generate time series of drainage
30 networks. Our approaches provide open-access techniques to systematically map ditches in peat-
31 dominated catchments through time. We focused on the algorithm in such a way that we can identify
32 the ditch networks from raw aerial images and LIDAR data based on the modification of multiple
33 filters and number of threshold values. Such data are needed to relate spatiotemporal drainage patterns
34 to observed changes in many northern rivers. We demonstrate our approach using data from the
35 Simojoki river catchment (3160 km²) in northern Finland. The catchment is dominated by forests and
36 peatlands that were almost all drained after 1960. For two representative locations in cultivated
37 peatland (downstream) and peatland forest (upstream) areas of the catchment; we found total ditch
38 length density (km/km²) estimated from aerial images and LIDAR data based on our proposed
39 algorithm varied from 2% to 50% compared against the monitored ditch length available from
40 National Land survey of Finland (NLSF) in 2018. A different pattern of source variation in ditch
41 network density was also observed for whole catchment estimates and for available drained peatland
42 database from Natural Resources Institute Finland (LUKE). Despite such differences no significant
43 differences were found using the non-parametric Mann-Whitney U-test with 0.05 significance level
44 based on the samples of pixel based identified ditches between (i) aerial images & NLSF vector files
45 and (ii) LIDAR data & NLSF vector files.

46 **Keywords:** Ditch length density, Drainage, Aerial Images, LIDAR, Peatlands, Forestry, Simojoki.

47 **1 Introduction**

48 In boreal and temperate zones, around 15 million hectares of peatlands have been drained for forestry
49 and other land uses since the 1950s (Paavilainen and Päivänen 1995). In Finland, peatland forestry is
50 important as it produces 25% of the annual forest growth. Recent finding stresses the importance not
51 only of evaluating current status but also understanding the historical evolution of ditch networks in
52 peatlands (Nieminen et al. 2017). Moreover, to model land use impacts on water quality and water
53 quantity, we need a good understanding of the past to calibrate and validate numerical models.
54 However, we currently lack adequate methodologies to quantify historical variation of ditch networks
55 at a catchment scale. Historical data about catchment drainage can be obtained from forestry statistics,
56 and more recently from GIS data. However, we lack adequate information on the timing and spatial
57 extent of ditch network development. Gathering spatial information on past drainage, aerial images
58 and satellite imagery are useful for large-scale quantitative assessment (Davis et al. 1978; Lambin
59 1997; Schneider and Gil Pontius 2001; Verburg et al. 1999).

60 Aerial images have been used to illustrate temporal changes in land use and ditches in peat dominated
61 (Linderholm and Leine 2004; Torabi Haghghi et al. 2018) and cultivated catchments (Passalacqua et
62 al. 2012). Automated ditch network identification from aerial images can generate linear features
63 through the application of a Hough line transformation with a specific counting mechanism (Karnieli
64 et al. 1996). Niu et al. (2007) proposed an algorithm to extract linear features from remote sensing
65 images. While the process of detecting features can be automated to certain extent (Artz et al., 2017;
66 Pirasteh et al., 2013), every step is resolution and image specific. As aerial images contain limited
67 ranges of spectral information, ditch identification is highly dependent on texture, pattern and context
68 of the images (Fox et al. 1995).

69 Using high-resolution digital elevation models (DEM) from LIDAR, Roelens et al., (2018) proposed
70 a method to extract vector data representing ditch networks based on local morphological features.
71 Their process identifies possible connections in the ditch network by calculating probability of

72 connectivity based on used logistic regression where the predictor variables are characteristics of the
73 ditch center lines derived from DEM. Passalacqua et al. (2012) and Sangireddy et al. (2016) also
74 developed algorithms that combine nonlinear filtering to remove noise during data pre-processing
75 with cost minimization principles for feature extraction.

76 The objective of this study is to develop methodologies to quantify the spatial and temporal
77 development of past ditching at a catchment scale. Specifically, we will focus on (1) developing a
78 semi-automatic algorithm to detect ditch networks from aerial images along with LIDAR data and
79 (2) use the same data sources and the workflow developed here to document spatial and temporal
80 patterns in the ditch network of a large, peat-dominated Finnish catchment. Documentation of ditch
81 network development is needed to quantify the environmental impacts of land management. Our
82 study focuses on a catchment where the peatlands were drained for forestry, peat extraction and
83 agriculture during the past decades. To our knowledge, this is the first study presenting methods for
84 quantifying past drainage history in catchments dominated by peatland forestry.

85 **2 Materials and Methods**

86 **2.1 Study area**

87 The Simojoki catchment (3160 km²) is located in the southern part of Lapland province in Finland
88 (Figure 1). The river is unregulated and runs 193 km from its headwaters in Lake Simojärvi to the
89 Bothnian Bay, dropping 176 m on its course to the sea. Forests on mineral soils (39%), and peatland
90 forests (53%) dominate (Lepistö et al. 2014) while agriculture covers about 3% of the catchment area
91 mainly near the catchment outflow (Perkkiö et al. 1995). Peatland drainage was carried out
92 throughout the catchment to support forest production and peat extraction. Drainage operations
93 started around 1950 and intensive drainage peaked during the 1970s in Simojoki. However, detailed
94 information about spatial development of drainage throughout the catchment is lacking. We also
95 considered two small representative areas for cultivated peatland (red circle in Figure 1a) and peatland

96 forest (green circle in Figure 1a) in the Simojoki catchment to portray how identification of drainage
97 systems varied spatially at different timescales.

98 **2.2 Aerial Images, LIDAR data and available ditch network database**

99 We used aerial images from National Land survey of Finland (NLSF 2020). Table A.1 (in
100 supplemental materials) illustrates the number of images for corresponding years that were available
101 for the analysis. All available images contain a single band except the images from 2018 which are
102 RGB images. Each image has 0.5 m x 0.5 m pixel resolution with pixel areas varying from 20 x 20
103 m² to 22.5 x 22.5 m².

104 Additionally, we used LIDAR data that contains 3D point clouds measured from an airplane with
105 laser scanning from Paituli, the open source spatial data download portal of the Ministry of Education
106 and Culture, Finland (Paituli 2019). The point density is about 1 point per 2 m². Horizontal accuracy
107 is about 60 cm and vertical 15 cm. We collected 442 LIDAR blocks (each block covers around 9 km²)
108 for the Simojoki catchment which were flown between 2008 and 2018. The data was arbitrarily
109 available for different portions of the catchment for different years. Therefore, we decided to use all
110 the images to cover the catchment from 2008-2018 (Figure A.1 in supplemental materials).

111 We also used the vector database of known base map ditches (hereafter called as “NLSF vector files”)
112 from 2005 to 2018 from National Land Survey of Finland (NLSF 2019) to support our analyses. Ditch
113 locations in the NLSF vector files are mainly derived from the manual addition of field surveys by
114 the National Land Survey of Finland. The accuracy of NLSF vector files vary from a scale of 1:5000
115 to 1:10000 with the EPSG: 2393 coordinate system (Datum: Kartastokoordinaattijärjestelmä (1966)).
116 The NLSF vector file database also includes irrigation channels and other types of ditches. In some
117 cases, elevation contours were misclassified as a ditch network in national map sources. We only
118 considered peatland ditches based on the available attribute information.

119 We also considered a regional database of drained peatland only from 1950 onwards from Finnish
120 statistical yearbooks developed by Natural Resources Institute Finland (LUKE 2018). This is
121 questionnaire based database without spatial information and represent province wise temporal
122 changes of drained peatland in Northern Finland, which was further downscaled and available for
123 Simojoki catchment.

124 **2.3 Image processing**

125 The image processing workflow presented here consists of four phases to prepare raw data and extract
126 drainage network features (Figure 2). All code is written in Python 2.7.

127 2.3.1 Phase-01: Raw aerial Image to Mosaic Image

128 To process raw aerial images, we wrote Python scripts (Appendix-B in supplemental materials) to
129 locate all the images at exact location based on a specified projection system (EPSG: 2393; Datum:
130 Kartastokoordinaattijärjestelmä (1966)). At first, we selected one image from each year to identify
131 the exact location of that image manually within the catchment. Using horizontal and vertical distance
132 of the selected image from a specific location, we developed an algorithm to estimate and save all
133 four coordinates (left, top, right and bottom) of other projected images (Step-01 to Step-03 in
134 Appendix-B in supplemental materials). Next, we georeferenced all available images for each specific
135 year based on their estimated coordinates. Results were visualized in ArcGIS 10.7. At this point, there
136 were still some errors for exact positioning of the image. These errors showed the gap between the
137 original location and the georeferenced location. To fix these errors we used features from maps
138 (NLSF 2019) beneath the catchment shape (Step-04 & Step-05 in Appendix-B in supplemental
139 materials) and performed manual quality control corrections for the projections. It should be noted
140 that projection errors did not affect the ditch detection procedure.

141 As numerous images were available for each year (Table A.1 in supplemental materials),
142 computationally it was not worth merging all the images. Instead, we used the radiometry
143 computation method (O'Connell et al. 2013) to generate seamlines from spectral patterns of features

144 within available images in the following manner. First, we created a blank geodatabase file containing
145 mosaic data for each year. It was important to ensure the same projection system and cell size for
146 mosaic data added to the geodatabase. After mosaic formation, we built footprints which showed the
147 outline of each image. By examining the values and patterns in the intersecting area and by computing
148 a path between the intersecting points, this path was then merged with the footprint to create seamlines
149 for each image in the mosaic data.

150 We used this seamline mosaic method to define the line along which images in the mosaic data are
151 connected. We did blending of the images instead of merging because this option depends on the
152 distance from the pixel to the edge (Böhner et al. 2006) to determine the value of overlapping pixels.
153 The option is computationally intensive for mosaicking. The final step of this mosaicking process
154 was to optimize and built the overview of created mosaicked image.

155 2.3.2 Phase-02: Raster preparation from LIDAR

156 LIDAR data (Paituli 2019) were already filtered and available in .lasz format (Heideman 2014). To
157 use these data, first we used ArcGIS 10.7 to convert all .lasz format LIDAR point cloud files to .lasd
158 format. Using elevation from .lasd files, we created a 1m Digital Elevation Model (DEM) which
159 represented the raster used from 2008-2018 for identifying ditch networks (as in section 2.3.3). We
160 used natural neighbor interpolation technique to determine the cell value of the newly created DEM.

161 2.3.3 Phase-03: Edge detection

162 We identified linear features in aerial images and LIDAR data. Phase-03 in Figure 2 represented the
163 steps after having the mosaicked image of corresponding year. We used the OpenCV Python library
164 (Bradski 2000; Pulli et al. 2012) to process mosaicked image for identifying linear ditch features.
165 This library covers associated functions that were used to detect edges from the images. OpenCV
166 cannot function smoothly if the image window contains too many pixels (e.g. greater than 2000).
167 Thus, mosaic images were split using a threshold number between 500 and 2000 pixels. We split

168 rasters to have same number of pixels in a sub raster from the mosaic image. Next, we identified and
169 extracted pixels representing waterbodies identified using the NLSF vector files (NLSF 2019). For
170 each split raster, we applied same technique and prepared the raster for masking. At this stage, we
171 masked the raster with extracted waterbody pixels (as in Figure 3a) to separate pixels which may
172 contain representative pixels of ditches.

173 After extraction of water body pixels, we next identified linear features through an initial application
174 of the Canny edge detection algorithm (Canny 1986) as implemented in OpenCV library (Bradski
175 2000). Edge detection is susceptible to noise in the raster; so prior to applying the detection algorithm
176 we used a Gaussian filter followed by Sobel kernels in both horizontal and vertical directions.
177 Gradient direction is always perpendicular to edges. After identifying gradients, we scanned the raster
178 to remove all unwanted pixels if the pixel has a position of local maximum in the direction of gradient
179 (Appendix-C in supplemental materials). At the final stage of Canny edge detection algorithm, two
180 thresholds such as minimum and maximum values were considered to determine whether the pixels
181 were part of ditches or not. The pixels which existed between these values based on their connectivity
182 and gradients, form the edges. One of the major issues we encountered was the choice of these
183 thresholds for each individual split raster (Bradski 2000).

184 After identifying edge pixels, the next step filtered the corresponding pixels. To filter pixels, we
185 applied a morphological transformation based on the raster shape to remove noise using a two by two
186 kernel filter. Next, we applied opening method by decreasing and then increasing the thickness of all
187 linear features in the raster layer. This stage generated an intermediate raster which contained some
188 pixels that did not represent ditch networks and therefore needed to be removed (Figure 3c). To
189 remove these pixels from the intermediate raster, we selected another threshold based on previously
190 selected thickness of features to apply opening method (Bradski 2000).

191 Next, we used a Hough line transform to identify the possible ditch network (Karnieli et al. 1996). It
192 works with a polar coordinate system (ρ, θ) . A line can be formed based on (ρ, θ) values which may
193 be a possible ditch line where ρ is measured in pixels and θ is measured in radians. For an initial point
194 on a line (with a value of 0), Hough line transform accumulated a pair of (ρ, θ) values from the
195 location of that point (x, y) within the raster. For the second point, it incremented the value in the
196 cells based on the preceding pair. The processes continued for every edge point and for each point,
197 the final incremented value represented the strength of the positive vote in the cell. Finally, the
198 transformation evaluated the maximum vote of the accumulator based on the origin and angle of the
199 image. The maximum vote referred the count of edge points that was used to form the line. We used
200 a Probabilistic Hough transform as it took only random subset of points and applied iteration with
201 different thresholds of minimum line segment and maximum allowed gap which determined how
202 well a line can form from the edge pixels.

203 2.3.4 Phase-04: Ditch network Identification and Validation

204 Ditch network identification (Figure 3d) involved multiple steps. At first, pixels identified using the
205 Probabilistic Hough Transformation method in the preceding step were grouped for each split raster
206 to represent lines (Figure 3d; Appendix-C in supplemental materials). Next, all identified ditch rasters
207 from all available aerial images and LIDAR data (after initial processing in section 2.3.2) were
208 processed by using OpenCV library. Later, polylines were identified and images further processed to
209 generate final ditches for each split raster in a single polyline shape file. Finally, the resultant ditch
210 features were used to build time series of ditch networks from 1952-2018 for aerial images and from
211 2008-2018 for LIDAR data.

212 To explain spatial and temporal distribution of ditch networks, we considered dataset from two
213 different locations (Figure 1a) as representative of different landscape types in the Simojoki
214 catchment from available images for each different year. To identify ditches from cultivated peatland
215 (mixed land cover) in downstream Simojoki we used threshold values and the ranges of values were

216 same for similar land cover areas in other parts of the catchment. We used the same approach for
217 peatland forest areas in other parts of the catchment. In results section, for spatial distribution of
218 ditches (in section 3.2) initially we showed two representative subsets of the catchment in this
219 manuscript. For each representative subset, we calculated the total ditch length (km) per sq. km for
220 each available year.

221 By following the same approach as stated above, we estimated total ditch length for each available
222 year from all available split images. At this stage, we also knew the percentage of aerial image
223 coverage of the entire catchment for each available year. So, we estimated total ditch length per area
224 of image coverage for entire catchment for each year. Later, the temporal changes of this estimation
225 was analyzed based on all available resources (aerial images, LIDAR data, NLSF vector files and
226 LUKE database).

227 Next, we tested statistical significance of ditch network estimations based on identified ditch lines
228 from the aerial images and available ditch network from NLSF vector files using the non-parametric
229 Mann-Whitney U-test with 0.05 significance level. Due to data unavailability from all available data
230 sources, to analyze statistical Mann-Whitney U-test there was only one common period at year 2018.
231 We tested the hypothesis that the median of sample of identified ditches from aerial images is equal
232 to the median of ditches available from NLSF vector files. We applied same hypothesis for Mann-
233 Whitney U-test for identified ditches from LIDAR data and ditches from NLSF vector files. For each
234 representative location as shown in Figure 1a, we intersected the identified ditches and NLSF vector
235 files with each pixel boundary for 2018. Then, for both cultivated peatland and peatland forest we
236 prepared samples of peatland ditches based on aerial images, LIDAR data and NLSF vector files. A
237 sample of peatland ditches basically represented the number of identified ditches (or features) from
238 each site for each available data source to apply statistical approach. Finally we applied Mann-
239 Whitney U-test for the prepared sample between (i) aerial images & NLSF vector files and (ii) LIDAR
240 data & NLSF vector files.

241 Later, we included an additional analysis based on the percentage of drained peatland (LUKE 2018)
242 to portray how ditch length would be distributed as drained peatland areas have increased in the
243 Simojoki catchment over the last 50 years. Inclusion of this analysis do not include peatlands which
244 are currently undrained. From LUKE database, we used the percent area drained in the Simojoki
245 catchment as a time series to compare with the ditch network densities obtained from drained peatland
246 portion of the Simojoki catchment by using the aerial images and LIDAR data.

247 **3 Results**

248 **3.1 Semi-Automatic georeferenced aerial images**

249 The techniques presented here enabled us to geo-reference all available images semi-automatically
250 (Figure 4). For all years as listed in Table A.1 in supplemental materials, application of mosaicking
251 process as described in section 2.3.1 resulted nine final mosaicked aerial images for year 1952, 1958,
252 1963, 1978, 1987, 1994, 1998, 2002 and 2018 (Figure 4) which we used to identify ditches (see results
253 in section 3.2 and section 3.3). The coverage of aerial images over the entire catchment varied from
254 5.60% in 1952 to 100% in 2018. We found the aerial images covered less than 60% of the catchment
255 area before 1987 and it increased from onwards.

256 **3.2 Ditch network identification**

257 Our methodology proved successful to identify ditches from aerial images and LIDAR data. The
258 methodology worked both in single band and RGB images, and at agricultural open areas as well as
259 in canopy covered peatland forest sites. With the methodology we were able to document the
260 historical development of drainage networks in the Simojoki river catchment (Figure 5 and Figure 6).
261 In both cases, our applied methodology recognized plausible existing ditches for all available years
262 for these different locations in Simojoki catchment.

263 Based on identified ditches from aerial images, intensive drainage had already started in downstream
264 parts of the catchment (red circle in Figure 1a) by 1952 (11.63 km/km²), and continued to 1978 (18.93

265 km/km²) (Figure 5a, Figure 5b). In 1994 (Figure 5c) and 1998 (Figure 5d), the results were almost
266 same although more ditches (~0.03 km/km² difference) were identified in 1994 because of better
267 image quality. The algorithm also underestimated the number of ditches in 2002 (11.81 km/km²),
268 especially in agricultural areas as shown in Figure 5e. Figure 5f also showed spatial variation of
269 ditches from NLSF vector files in 2018 where we found 16.68 km/km² ditch length which was 25%
270 higher (12.51 km/km²) than identified aerial image ditches. But from our statistical analysis as
271 described in section 2.3.4, the output did not show any statistically significant difference (Mann-
272 Whitney U = 5614, sample size-01= sample size-02 = 111, p=0.2534). This indicated a similar
273 distribution for two sample dataset of ditch networks from aerial image and NLSF vector files for this
274 specific block in Figure 5f.

275 The upstream, forest dominated parts of the catchment (green circle in Figure 1a) had a different
276 spatiotemporal development of ditch networks (Figure 6) identified from aerial images. In 1958, there
277 was no ditch network at all. In the 1970s initial drainage started in this region (Figure 6b). For almost
278 all other years from 1987 onwards, there was very slight change after initial drainage (around 15.5
279 Km/km²) in this area. In 2018 we found 16 km/km² identified ditches from aerial images which were
280 less than NLSF vector files (21.92 km/km²) by 37% in Figure 6f though we found p=0.3274 (Mann-
281 Whitney U = 2087, sample size-01= sample size-02 = 68) for U-test for the samples of identified
282 ditches between upstream aerial images & NLSF vector files.

283 The ditch network density identified using Lidar data was higher with the ditch network density found
284 by using the aerial imagery and NLSF vector files (Figure 7) in two representative locations shown
285 in Figure 1a. The ditches identified from LIDAR missions flown between 2008 and 2018 detected
286 some pixels which presented plausible ditches not identified in the aerial images, and which were not
287 connected to the total ditch network of that specific block. The parameters such as minimum line
288 segment and maximum allowed gap, used to identify ditches from LIDAR, are the main reasons for
289 the differences in ditch network extent identified in these versus aerial images. In 2018, for

290 downstream Simojoki (Figure 7a) ditch network length derived from LIDAR data (40 km/km^2) was
291 more than 50% higher than the length derived from NLSF vector files (around 21 km/km^2) whereas
292 it was 2% lower (30.23 km/km^2) from NLSF vector files (30.77 km/km^2) for upstream location of the
293 catchment in Figure 7b. In spite of having spatial differences for identified ditches between LIDAR
294 data and NLSF vector files, we found p value (0.2491 for downstream (Mann-Whitney $U = 5609$,
295 sample size-01= sample size-02 = 111) & 0.6229 for upstream (Mann-Whitney $U = 2199$, sample
296 size-01= sample size-02 = 68)) > 0.05 . This also did not indicate statistically significant difference
297 for Mann-Whitney U test for all the samples for both upstream and downstream locations of the
298 catchment.

299 **3.3 Temporal variation of ditch network**

300 There were different temporal patterns in the development of ditch network density in the downstream
301 (Figure 8a) and upstream (Figure 8b) study sites. Based on results from the aerial images from the
302 downstream site (for same location as in Figure 5), ditch network density (km/km^2) increased by
303 approximately 63% between 1952 and 1978 and then declined (Figure 8a). At this site in 2018, ditch
304 network density obtained from the NLSF vector files (NLSF 2019) was 16.68 km/km^2 whereas it was
305 estimated to be 12.51 km/km^2 based on aerial images. At the upstream forest site (for same location
306 as in Figure 6); the change of ditch network density occurred from 1958 to 1987, was around 15.5
307 km/km^2 . Estimated ditch network density started to decrease again after 1994 but not abruptly as
308 happened in the downstream site (Figure 8b).

309 LIDAR estimated ditch network densities (40 km/km^2) were higher than estimates derived from aerial
310 images (12.51 km/km^2) or NLSF vector files (around 21 km/km^2) in the downstream site (Figure 8a).
311 At the upstream forest site, LIDAR and NLSF vector files derived estimates of ditch network density
312 were approximately similar (30.77 km/km^2), but were higher than the densities estimated from aerial
313 images (16 km/km^2) (Figure 8b).

314 A different pattern of source variation in ditch network density was observed for whole catchment
315 estimates (Figure 9) as compared to estimates for the small scale sites (Figure 8). When making
316 estimates for the entire catchment, we primarily found an inverse relationship between ditch network
317 density and percentage of aerial image coverage. As the percentage of available aerial images
318 changed, ditch network density also varied throughout the study period (Figure 9a). For example, in
319 1998 area image coverage was 97.61% where we estimated a ditch network density of around 34
320 km/km². Whereas in 2002 though the percentage coverage was lower than 1998 (58.27%), ditch
321 network density increased (44.26 km/km²). This finding was confirmed by examining NLSF vector
322 files from 2018, that covered entire Simojoki catchment (100%) but only contained 18 km/km² ditch
323 network density (green color in Figure 9a). Ditch network densities obtained from NLSF vector files
324 were around 56% lower than the ditch network density estimated from 2018 aerial images (41
325 km/km²) (Figure 9a). We found slight changes of ditch length (km/km²), around 8% from 2005 to
326 2018 for available NLSF vector files (Paituli 2019). Whereas from 2002 to 2018 aerial images,
327 without considering percentage of coverage area we found the change was around 7.5% for ditches
328 identified from aerial images.

329 For the whole catchment (Figure 9a), the ditch network density (34.43 km/km²) estimated from
330 LIDAR data was greater than that estimated from NLSF vector files (18 km/km²) and less than that
331 estimated from aerial images (41 km/km²). The statement is also applicable in case of drained
332 peatlands (Figure 9b) for ditch network density (21.22 km/km²) estimated from LIDAR data.

333 When we only considered temporal patterns of the percentage of drained peatland in the Simojoki
334 catchment obtained from forest drainage database (LUKE 2018) (Figure 9b), the situation was bit
335 different for aerial images than that for the whole catchment (Figure 9a). The percentage of drained
336 peatland increased from 31% in 1967 to 61.63% in 2018. Ditch network density in drained peatlands
337 estimated from aerial images showed a slight a decreasing tendency throughout the same period,
338 especially in 2018 when we found 25.38 km/km². The decreasing pattern was more visible from 2005

339 (27 km/km²) in Figure 9b when there was slight change for NLSF vector files (16.56 km/km² to 17.96
340 km/km²) for the same period.

341 **4 Discussion**

342 **4.1 Development of ditch network maps**

343 Knowledge of ditch network extent and history is essential for sustainable and efficient forest
344 management, especially to understand environmental changes at the catchment scale (Ecke, 2009).
345 Previously, identification of ditch networks was based on time consuming field observations or
346 manual digitalization from aerial images. The approach we developed can automatically identify
347 peatland ditch networks from aerial images and LIDAR data where ditches are identified as linear
348 feature based on parameter value thresholds. This methodology is also suitable for open peatlands or
349 cultivated fields.

350 For small scale zones, there are some shallow ditches in downstream parts of the catchment which
351 were identified by LIDAR were not always detected using NLSF vector files (Figure 8). Thus ditch
352 network density identified from LIDAR data is higher than NLSF vector files in downstream
353 (cultivated peatland) parts of the catchment and is closer to the NLSF vector files in upstream peatland
354 forest site.

355 Conversely at the catchment scale, the methodology identified more plausible ditches from aerial
356 images than LIDAR, especially from 2008-2018 (Figure 9). The major change in ditch network
357 detection occurred in forested areas of the catchment, probably as a result of canopy cover masking
358 ditches on the ground. However, sometimes the method also identified ditches from LIDAR which
359 are not artificial (natural depression) as addressed by (Duke et al. 2006; Roelens et al. 2018).

360 **4.2 Reliability of proposed techniques to estimate historical ditch network use practices**

361 Processing raw aerial images which lack projection information involves multiple steps from geo-
362 referencing to ditch identification. The technique we present here provides generic tools to correctly
363 geo-reference large numbers of aerial images and to reliably classify pixels based on threshold values.
364 For instance, when there are values that exist above or below plausible ditch pixel threshold values;
365 our proposed method showed some noise, indicative of undrained areas in the image. The ditch
366 networks identified from both aerial images and LIDAR data sources were slightly different than
367 those identified from NLSF vector files (NLSF 2020) (Figure 9a).

368 The approach presented here to identify ditch networks functioned better in areas identified as drained
369 peatlands (Figure 9b) than for the catchment as a whole (Figure 9a). This occurred because the gap
370 between the ditch length identified from aerial images for the entire catchment and NLSF vector files
371 (Figure 9a) is higher than the ditch length identified from aerial images for drained peatlands and
372 NLSF vector files (Figure 9b).

373 The approach presented here relies on user-supplied parameter ranges (Appendix-C in supplemental
374 materials) to determine whether a linear feature should be treated as a ditch line or not. While this
375 study used a temporal series of aerial datasets to identify ditch age and extent, the proposed approach
376 can be used with any other high resolution satellite image so long as the limitations of these data
377 sources recognized. For example, it is recommended to consider only high spatial resolution (0.5 m -
378 2 m) data sources to adequate identification of ditch networks.

379 **4.3 Uncertainties involved with ditch identification processes**

380 In this study; coarse pixel sizes, multiple bands in images, variable gray colors and image sharpness
381 were prominent factors that caused errors in estimation of ditch network length. Most of the
382 uncertainties at the catchment scale occurred due to poor quality of old aerial photos.

383 Spectral variation among mosaicked images was also an issue in this study. We addressed this
384 concern and developed a recursive process to split the mosaic so that all the identified pixels within
385 that image boundary could drop in a certain range. Differences in spectral information among all
386 mosaic images, even after preprocessing can also affect parameter ranges for Canny edge detection
387 (Bradski 2000) and Hough line transformation (Karnieli et al. 1996). To overcome this issue,
388 automated, iterative techniques could be developed based on image spectral information. In this
389 study, we could not apply same threshold values throughout as spectral information of each image is
390 unique. Thus, we applied kernel filtering after splitting each image into small rectangles so that the
391 parameter values can be effective only for the relevant image section.

392 Unavailability of aerial images was another major issue at the catchment scale. While it was possible
393 to create one single mosaic image for each specific year, catchment scale ditch network quantification
394 was challenging as the mosaicked image did not cover the entire catchment. In some cases we noticed
395 some pixels were recognized as part of ditch network which did not represent the original field
396 situation. Thus, ditch length for the Simojoki catchment was represented as ditch network density,
397 i.e., the total length of identified ditches per area coverage of images for each year, recognizing that
398 this might result in uncertainty for catchment scale.

399 Therefore, for the entire catchment our approach generates tentative estimates of ditch network
400 density based on image coverage for each year. This effect is especially relevant before 1987 when
401 image coverage was less than 60%. Additionally, we also found a sharp rise of ditch network density
402 in 1998 (Figure 9) even though the spatial extent of image coverage did not change abruptly from
403 1994. One major reason for the change in estimated density was ditch network maintenance
404 operations which had been conducted in 1990s. By maintaining ditches, old vegetated ditch sections
405 became more visible in the aerial images. For example, Figure 4h shows a mosaic image from 2002
406 that contains multiple segments of the catchment with different drainage patterns, instead of one

407 combined zone. So, the spatial distribution of ditches based on the proposed technique is always
408 relative to the specific portion of the image.

409 Discrepancies were noted when validating results against the available ditch network database for the
410 Simojoki catchment. Due to lack of spatial information, the local survey-based LUKE drained
411 peatland database (Figure 9b) may be less accurate for local future analysis compared to our approach.
412 There are limitations to all of available ditch network databases as the data quality is not known and
413 the contents are often not up-to-date. Data quality in monitored ditch network databases and threshold
414 selection are important aspects that need to be addressed in a more detailed way to improve the
415 identification process of the algorithm presented here. Furthermore, a sensitivity analysis is needed
416 in the future to address the possible errors that can affect accuracy of ditch density estimates.

417 **4.4 Can the proposed approaches be beneficial for future management of peat dominated** 418 **catchments?**

419 From a land manager's perspective, it is always important to understand the temporal distribution of
420 ditch networks, especially in catchments containing a high percentage of peat soil. Some
421 environmental changes, especially in soils, occur slowly (Araújo and Rahbek 2006) and thus
422 historical land use activities may show impacts for years or even decades. For example, (Nieminen
423 et al. 2018) found timing of initial drainage can explain decadal-scale variation in nutrient
424 concentrations in runoff from drained peatland. Vegetation at aapamire complexes have also been
425 observed in their study to react over decades to small scale drainage. Catchment-scale identification
426 of timing for initial drainage is needed to better estimate catchment scale consequences of peatland
427 drainage. By considering ditch network density as an independent variable, our approach can also be
428 used to estimate water quality variations and biodiversity status of peatlands, especially in spatially
429 distributed hydrological models. Aspects of catchment-scale hydrology which were not possible to
430 address previously can be modelled now with our proposed approach to understand the effects of all
431 transported elements. This can be beneficial in the future to take actions from managerial and

432 ecological point of view. However, including other relevant factors such as peat depth in the ditch
433 identification process will require further attention.

434 The approaches presented here can also be implemented to estimate hydrological fluxes and
435 emissions from drained peatlands (Ojanen et al. 2019) to guide future potential peatland restoration
436 actions (Laine et al. 2019). Our approach can be applied for any peat dominated catchment to identify
437 existing ditch networks and to consider future actions.

438 **5 Conclusions**

439 Analysis of historical aerial images leads to new insights about temporal and spatial distribution of
440 catchment scale peatland drainage. Our approach was successfully applied in a peatland forestry
441 dominated catchment where increasing canopy cover made ditch identification challenging.
442 Introduction of multiple filters increased the possibility of accuracy to identify the ditch networks. If
443 a user follows our approaches, to identify ditch network from scratch, all small steps associated in
444 this study can be applied sequentially or part by part. For two representative locations in cultivated
445 peatland (downstream) and peatland forest (upstream) areas of the catchment, we found total ditch
446 length density (km/km^2) estimated based on our proposed algorithm from aerial images and LIDAR
447 data varied from 2% to 50% compared against the monitored ditch length available from National
448 Land survey of Finland (NLSF) in 2018. A different pattern of source variation in ditch network
449 density was also observed for whole catchment estimates and for available drained peatland database
450 from Natural Resources Institute Finland (LUKE). We found no statistically significant differences
451 between ditch network length identified from (i) aerial images & NLSF vector files and (ii) LIDAR
452 data & NLSF vector files. Our analysis showed the potential of old aerial images and LIDAR data to
453 identify drainage network changes for sustainable catchment scale management of future land uses
454 and restoration actions. In future, it may be possible to develop more fragmented algorithmic section
455 to further improve the approaches of identifying ditches. Our open access methodology and database

456 can be a source to develop future conceptual models and to apply them in different forest based
457 managerial and ecological scenarios.

458 **Data Availability Statement**

459 Some or all data, models, or code that support the findings of this study are available from the
460 corresponding author upon reasonable request. (All codes are available in Supplementary Materials)

461 **Acknowledgements**

462 We would like to thank python group, stack overflow group and arcpy developers for discussion. This
463 work was part of the Nordic Centre of Excellence BIOWATER, funded by Nordforsk under project
464 number 82263.

465 **References**

- 466 Araújo, M. B., and Rahbek, C. (2006). “How does climate change affect biodiversity?” *Science*,
467 313(5792), 1396 LP – 1397.
- 468 Artz, R. R.E., Donaldson-Selby, G., Poggio, L., Donnelly, D., and Aitkenhead, M. J. (2017).
469 *Comparison of remote sensing approaches for detection of peatland drainage in Scotland*.
- 470 Böhner, J., Selige, T., and Ringeler, A. (2006). “Image segmentation using representativeness
471 analysis and region growing.” *Göttinger Geographische Abhandlungen*, 115, 29–38.
- 472 Bradski, G. (2000). “OpenCV Library.” *Dr. Dobb’s Journal of Software Tools*.
- 473 Canny, J. (1986). “A Computational Approach to Edge Detection.” *IEEE Transactions on Pattern
474 Analysis and Machine Intelligence*, Morgan Kaufmann, PAMI-8(6), 679–698.
- 475 Davis, S. M., Landgrebe, D. A., Phillips, T. L., Swain, P. H., Hoffer, R. M., Lindenlaub, J. C., and
476 Silva, L. F. (1978). “Remote sensing: The quantitative approach.” *New York, McGraw-Hill
477 International Book Co., 1978. 405 p.*
- 478 Duke, G. D., Kienzle, S. W., Johnson, D. L., and Byrne, J. M. (2006). “Incorporating ancillary data
479 to refine anthropogenically modified overland flow paths.” *Hydrological Processes*, 20(8),
480 1827–1843.
- 481 ECKE, F. (2009). “Drainage ditching at the catchment scale affects water quality and macrophyte
482 occurrence in Swedish lakes.” *Freshwater Biology*, John Wiley & Sons, Ltd (10.1111), 54(1),
483 119–126.
- 484 Fox, J., Krummel, J., Yarnasarn, S., Ekasingh, M., and Podger, N. (1995). “Land use and landscape
485 dynamics in northern Thailand: assessing change in three upland watersheds.” *Ambio*, 24(6),
486 328–334.

- 487 Heideman, H. K. (2014). "LIDAR base specification (ver. 1.2)." *U.S. Geological Survey Techniques*
488 *and Methods 11-B4*, 41.
- 489 Karnieli, A., Meisels, A., Fisher, L., and Arkin, Y. (1996). "Automatic extraction and evaluation of
490 geological linear features from digital remote sensing data using a Hough Transform."
491 *Photogrammetric Engineering & Remote Sensing*, 62(5), 525–531.
- 492 Laine, A. M., Mehtätalo, L., Tolvanen, A., Frolking, S., and Tuittila, E. S. (2019). "Impacts of
493 drainage, restoration and warming on boreal wetland greenhouse gas fluxes." *Science of the*
494 *Total Environment*, Elsevier B.V., 647, 169–181.
- 495 Lambin, E. F. (1997). "Modelling and monitoring land-cover change processes in tropical regions."
496 *Progress in Physical Geography*, 21(3), 375–393.
- 497 Lepistö, A., Futter, M. N., and Kortelainen, P. (2014). "Almost 50 years of monitoring shows that
498 climate, not forestry, controls long-term organic carbon fluxes in a large boreal watershed."
499 *Global Change Biology*, 20(4), 1225–1237.
- 500 Linderholm, H. W., and Leine, M. (2004). "An assessment of twentieth century tree-cover changes
501 on a southern Swedish peatland combining dendrochronology and aerial photograph analysis."
502 *Wetlands*, Springer Netherlands, 24(2), 357.
- 503 LUKE. (2018). "Drainage status of forestry land." <<http://statdb.luke.fi/PXWeb/pxweb/fi/LUKE/>>
504 (Oct. 28, 2019).
- 505 Nieminen, M., Palviainen, M., Sarkkola, S., Laurén, A., Marttila, H., and Finér, L. (2018). "A
506 synthesis of the impacts of ditch network maintenance on the quantity and quality of runoff from
507 drained boreal peatland forests." *Ambio*, Springer Netherlands.
- 508 Nieminen, M., Sallantausta, T., Ukonmaanaho, L., Nieminen, T. M., and Sarkkola, S. (2017). "Nitrogen
509 and phosphorus concentrations in discharge from drained peatland forests are increasing."
510 *Science of the Total Environment*, Elsevier B.V., 609, 974–981.
- 511 Niu, R., Mei, X., Zhang, L., and Li, P. (2007). "Linear Features Extraction From Remote Sensing
512 Image Based on Wedgelet Decomposition." *Fourth International Conference on Image and*
513 *Graphics (ICIG 2007)*, IEEE, 508–512.
- 514 NLSF. (2019). "Laser scanning data | National Land Survey of Finland."
515 <[https://www.maanmittauslaitos.fi/en/maps-and-spatial-data/expert-users/product-](https://www.maanmittauslaitos.fi/en/maps-and-spatial-data/expert-users/product-descriptions/laser-scanning-data)
516 [descriptions/laser-scanning-data](https://www.maanmittauslaitos.fi/en/maps-and-spatial-data/expert-users/product-descriptions/laser-scanning-data)> (Mar. 29, 2019).
- 517 NLSF. (2020). "Aerial photographs | National Land Survey of Finland."
518 <[http://www.maanmittauslaitos.fi/en/maps-and-spatial-data/expert-users/topographic-data-](http://www.maanmittauslaitos.fi/en/maps-and-spatial-data/expert-users/topographic-data-and-how-acquire-it)
519 [and-how-acquire-it](http://www.maanmittauslaitos.fi/en/maps-and-spatial-data/expert-users/topographic-data-and-how-acquire-it)>.
- 520 O'Connell, J., Connolly, J., Vermote, E. F., and Holden, N. M. (2013). "Radiometric normalization
521 for change detection in peatlands: a modified temporal invariant cluster approach." *International*
522 *Journal of Remote Sensing*, Taylor & Francis, 34(8), 2905–2924.
- 523 Ojanen, P., Penttilä, T., Tolvanen, A., Hotanen, J. P., Saarimaa, M., Nousiainen, H., and Minkkinen,
524 K. (2019). "Long-term effect of fertilization on the greenhouse gas exchange of low-productive
525 peatland forests." *Forest Ecology and Management*, Elsevier B.V., 432, 786–798.
- 526 Paavilainen, E., and Päivänen, J. (Juhani). (1995). *Peatland forestry: ecology and principles*.
527 Springer-Verlag.

- 528 Paituli. (2019). "AVAA - Paituli spatial data service." <[https://avaa.tdata.fi/web/avaa/-/paituli-](https://avaa.tdata.fi/web/avaa/-/paituli-paikkatietopalvelu)
529 [paikkatietopalvelu](https://avaa.tdata.fi/web/avaa/-/paituli-paikkatietopalvelu)> (Mar. 29, 2019).
- 530 Passalacqua, P., Belmont, P., and Foufoula-Georgiou, E. (2012). "Automatic geomorphic feature
531 extraction from lidar in flat and engineered landscapes." *Water Resources Research*, 48(3), 1–
532 18.
- 533 Perkkiö, S., Huttula, E., and Nenonen, M. (1995). "Water protection plan for the Simojoki river
534 basin." *Publ. Water and Environment Administration-series A*, 200, 1–102.
- 535 Pirasteh, S., Pradhan, B., Safari, H. O., and Ramli, M. F. (2013). "Coupling of DEM and remote-
536 sensing-based approaches for semi-automated detection of regional geostructural features in
537 Zagros mountain, Iran." *Arabian Journal of Geosciences*, Springer-Verlag, 6(1), 91–99.
- 538 Pulli, K., Baksheev, A., Korniyakov, K., and Eruhimov, V. (2012). "Real-time computer vision with
539 OpenCV." *Communications of the ACM*, 55(6), 61–69.
- 540 Roelens, J., Höfle, B., Dondeyne, S., Van Orshoven, J., and Diels, J. (2018). "Drainage ditch
541 extraction from airborne LiDAR point clouds." *ISPRS Journal of Photogrammetry and Remote*
542 *Sensing*, 146, 409–420.
- 543 Sangireddy, H., Stark, C. P., Kladzyk, A., and Passalacqua, P. (2016). "GeoNet: An open source
544 software for the automatic and objective extraction of channel heads, channel network, and
545 channel morphology from high resolution topography data." *Environmental Modelling &*
546 *Software*, 83, 58–73.
- 547 Schneider, L. C., and Gil Pontius, R. (2001). "Modeling land-use change in the Ipswich watershed,
548 Massachusetts, USA." *Agriculture, Ecosystems & Environment*, Elsevier, 85(1–3), 83–94.
- 549 Torabi Haghighi, A., Menberu, M. W., Darabi, H., Akanegbu, J., and Kløve, B. (2018). "Use of
550 remote sensing to analyse peatland changes after drainage for peat extraction." *Land*
551 *Degradation and Development*, John Wiley & Sons, Ltd, 29(10), 3479–3488.
- 552 Verburg, P. H., de Koning, G. H. J., Kok, K., Veldkamp, A., and Bouma, J. (1999). "A spatial explicit
553 allocation procedure for modelling the pattern of land use change based upon actual land use."
554 *Ecological Modelling*, Elsevier, 116(1), 45–61.
- 555

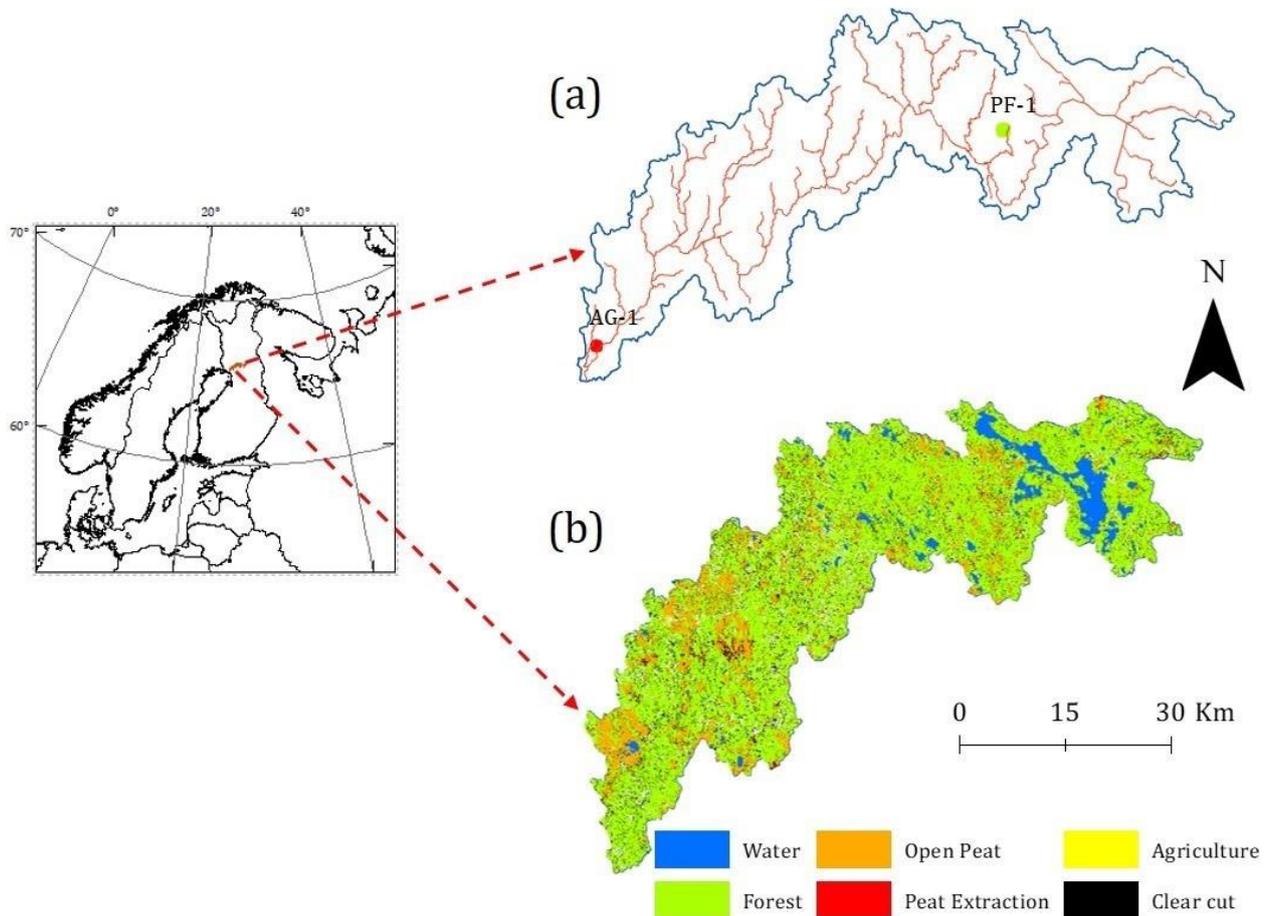


Figure 1: Location of the Simojoki catchment in Finland: (a) catchment boundary with rivers and two specific sites where red circle in downstream part of Simojoki presents cultivated peatland (mixed land cover types, mainly agriculture and forest) and green circle in upstream Simojoki shows peatland forest , these sites were chosen in this manuscript to present spatial and temporal variability of the ditches because of their land cover variations based on (b) existing land cover from Corine 2018 data (“adapted from (a) National Land Survey of Finland 2018 and (b) Copernicus land portal 2013”).

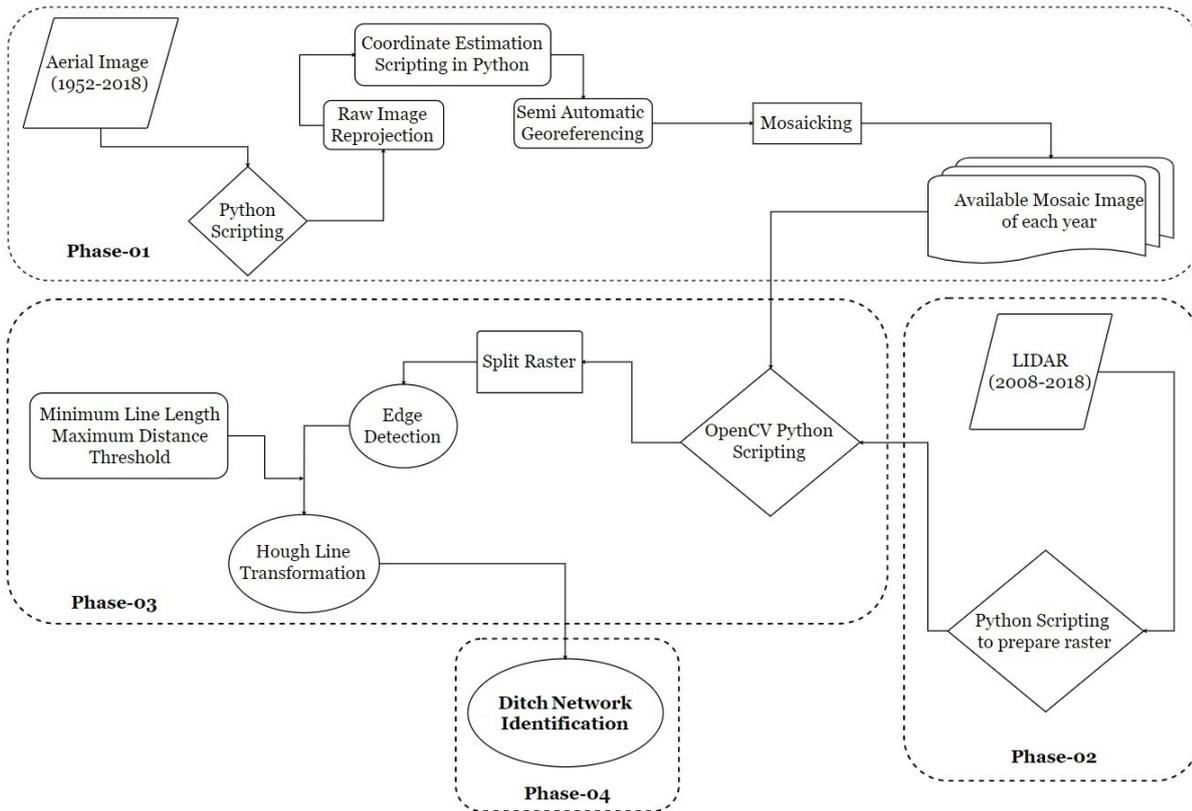


Figure 2: Detailed processes of data analysis to implement in this study.

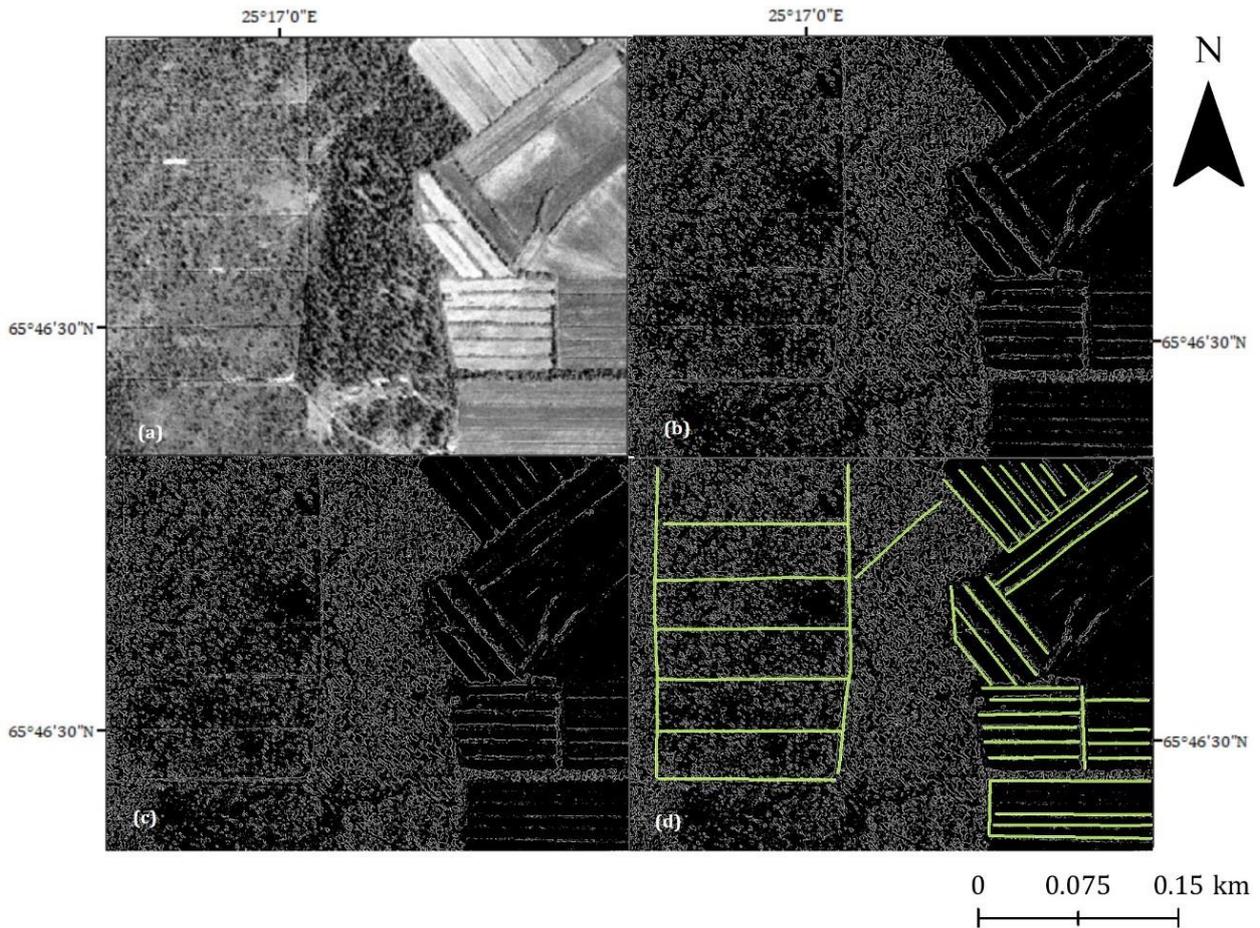


Figure 3: From split image to ditch detection: (a) a sample split image from 1978, (b) white cells represented detected edge pixels after applying Canny edge detection, (c) cells after morphological transformation with probable pixels for line formation and (d) final selection of pixels after Hough transformation where green lines represented finally detected ditches (“adapted from National Land Survey of Finland 2018”).

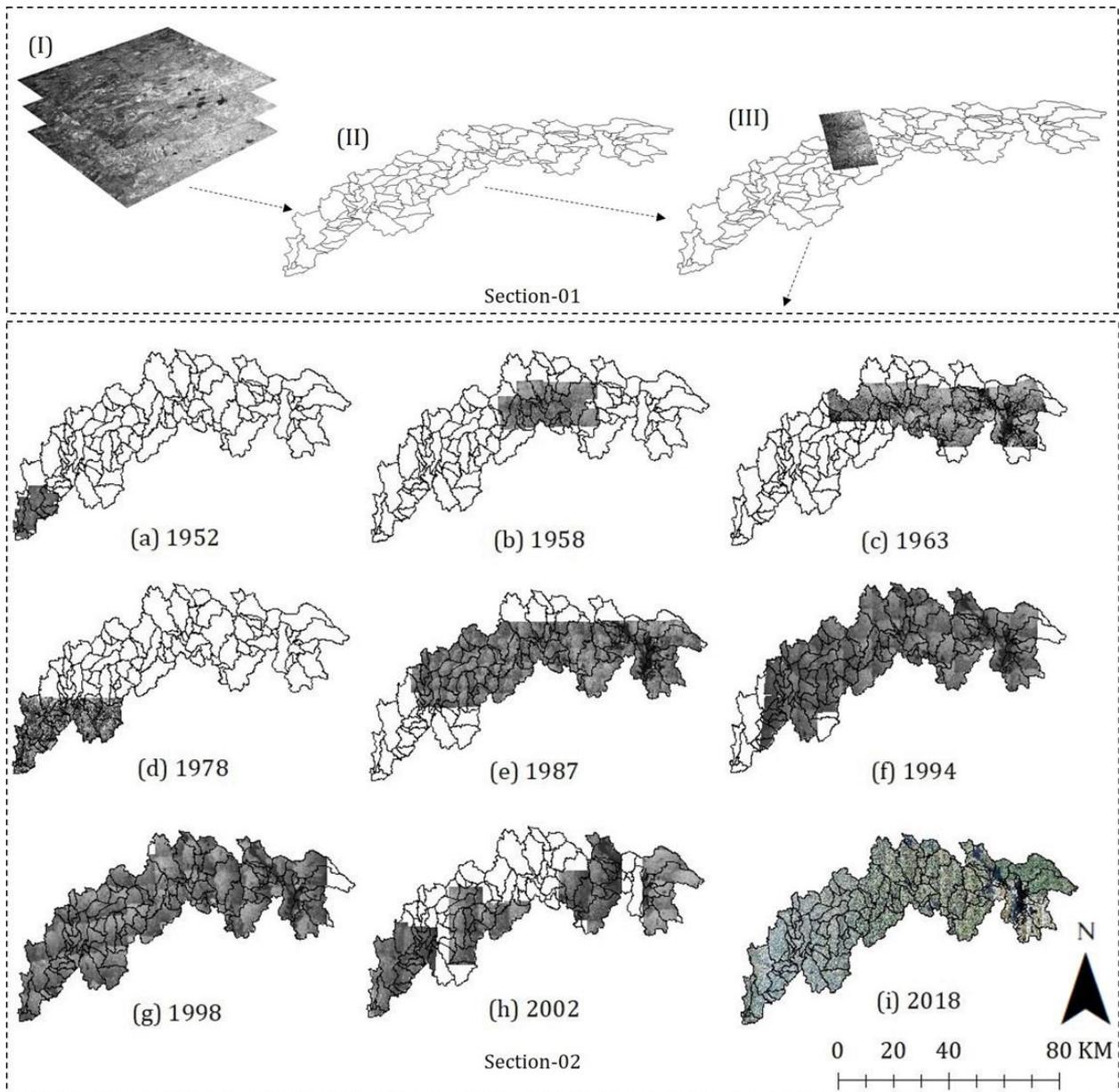


Figure 4: The mosaicking process and temporal changes of aerial images in the Simojoki catchment. The top panel represents section-01 of the methods (I) non-geo referenced images overlapping with each other at same point far from (II) catchment boundary and (III) after applying Python scripts from Appendix-B (in supplemental materials) georeferenced location of the same images; whereas section-02 (a - i) show mosaicked output of available aerial images with different coverage portion of entire catchment for different years from 1952-2018 (“adapted from National Land Survey of Finland 2018”).

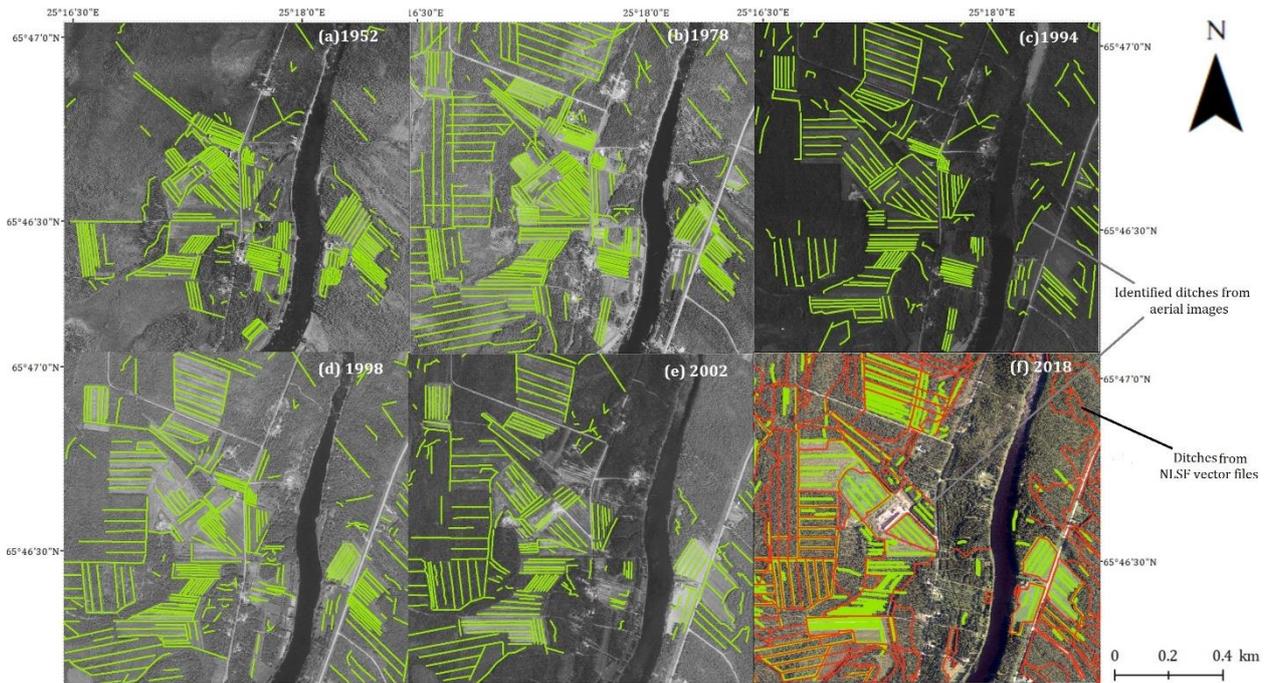


Figure 5: Samples of identified ditch networks for different years of available aerial images: (a) 1952, (b) 1978, (c) 1994, (d) 1998, (e) 2002 and (f) 2018 in downstream Simojoki based on location of red circle in Figure 1a with cultivated peatland (mainly agriculture and forest). Figure f also contains 2018 ditches from NLSF vector files in red color. (“adapted from National Land Survey of Finland 2018”).

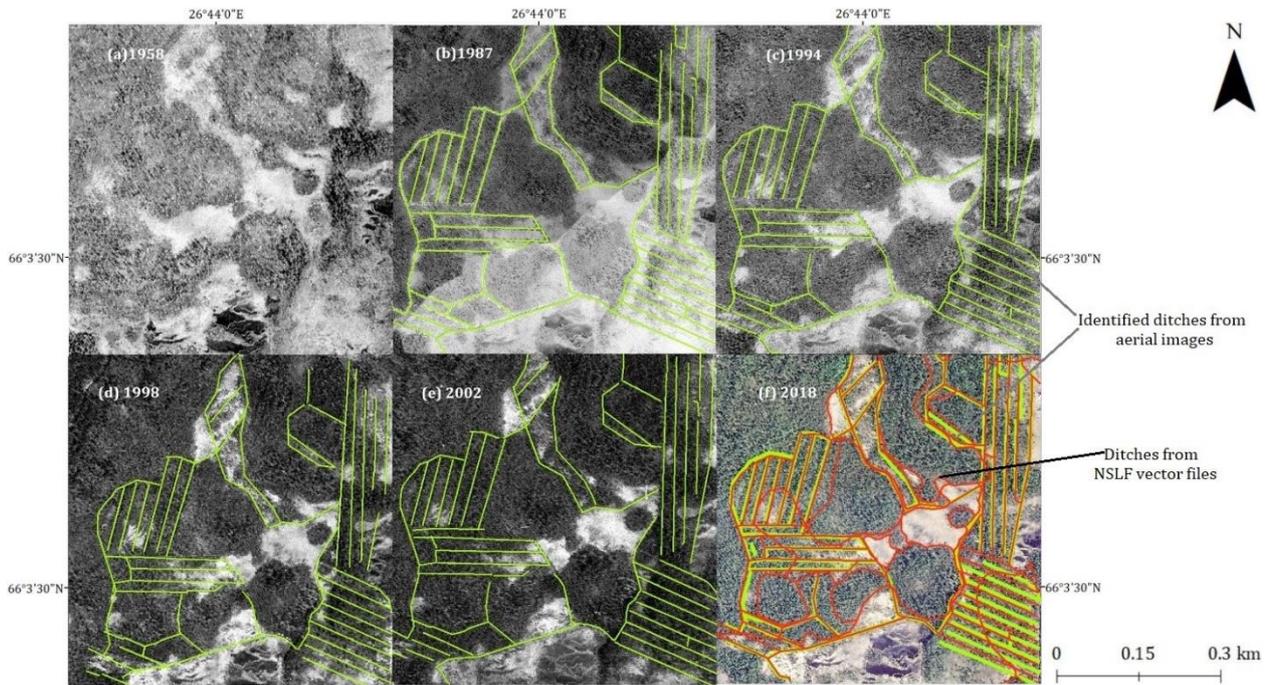


Figure 6: Samples of identified ditch networks for forested areas in upstream Simojoki (green circle in Figure 1a) for different years of available aerial images: (a) 1958, (b) 1987, (c) 1994, (d) 1998, (e) 2002 and (f) 2018. Figure f also contains 2018 ditches from NLSF vector files in red color (“adapted from National Land Survey of Finland 2018”).

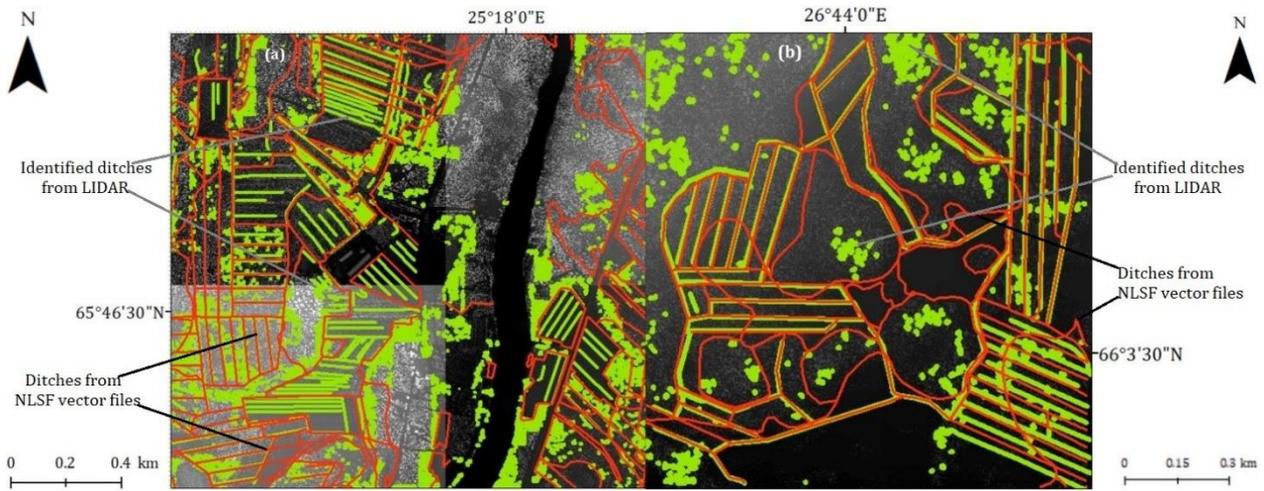


Figure 7: Identified ditch networks in green lines based on processed and combined database of LIDAR (2008-2018) along with ditch networks from NLSF vector files in red lines for (a) cultivated peatland in downstream Simojoki for same location as in Figure 5 and (b) forested areas in upstream Simojoki for same location as in Figure 6 (“adapted from National Land Survey of Finland 2018”).

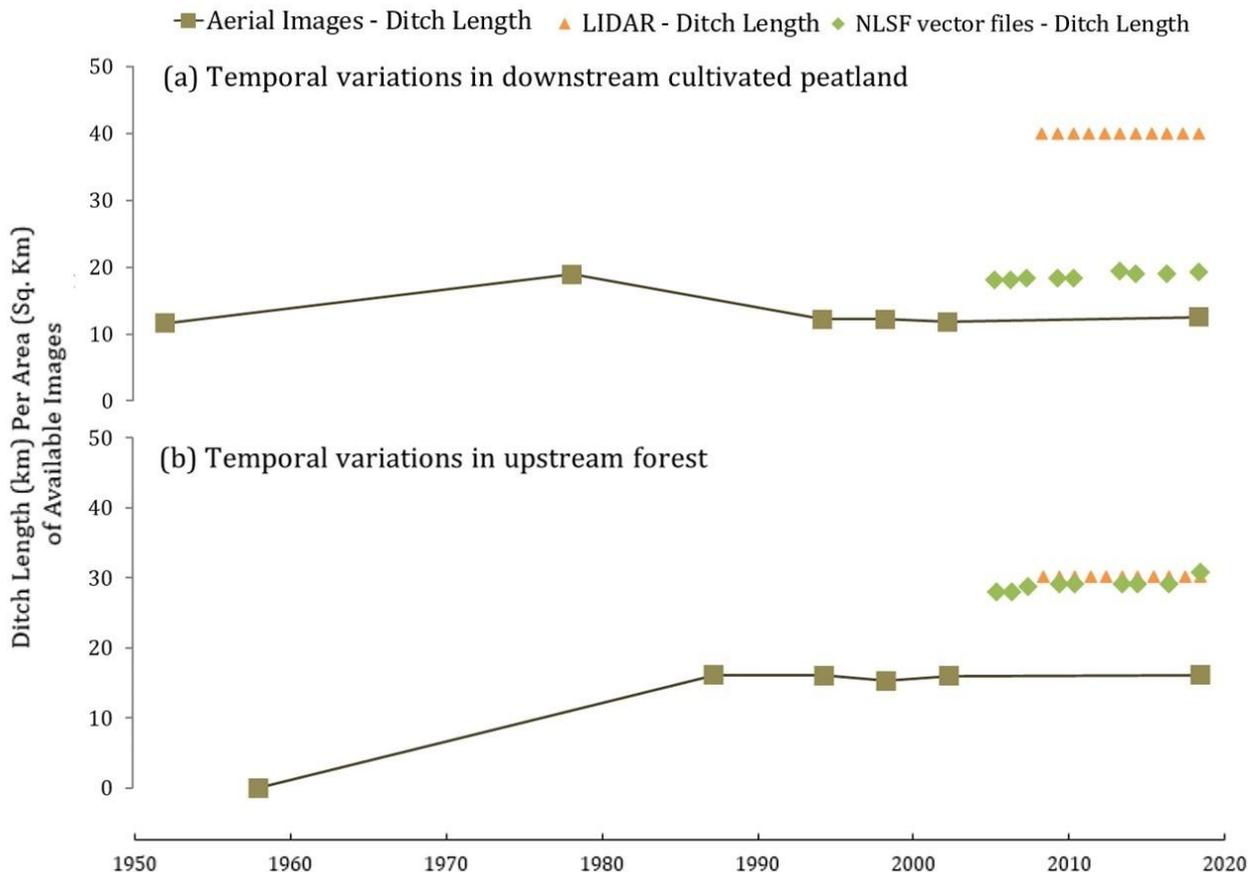


Figure 8: Presents temporal variations of ditch network density (km/km²) estimated for a small scale (1~2 Sq. km) zone in (a) cultivated peatland (downstream) and (b) forested areas (upstream) of the Simojoki catchment. Values shown in the figure are the total ditch length found for each available year based on each available data source.

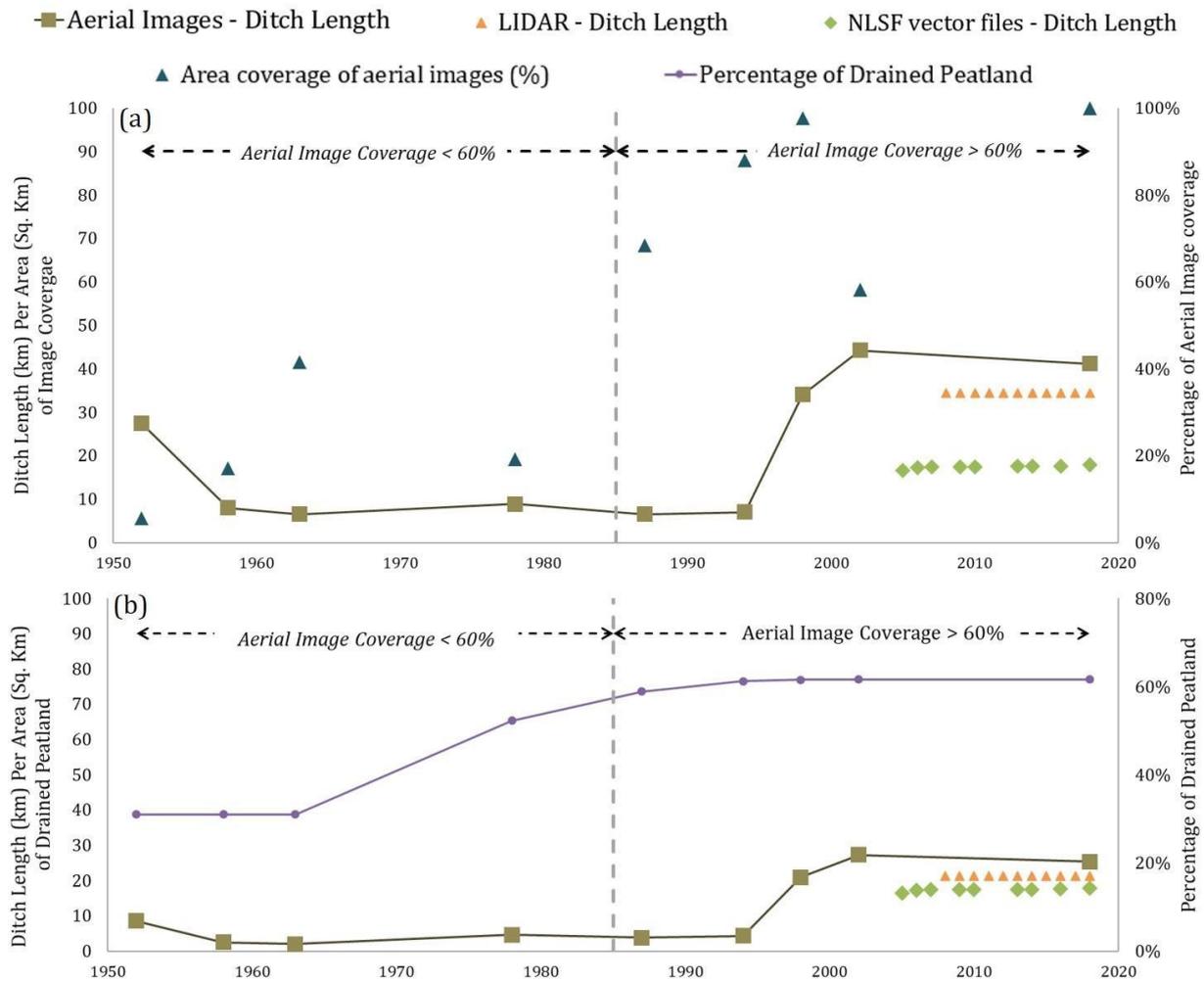


Figure 9: Catchment scale temporal variations of ditch network density (km/km²). Total ditch length for each available year (a) based on available images for the whole catchment and (b) only for drained peatlands.