

1 **Lithological mapping of the Troodos ophiolite, Cyprus, using airborne LiDAR topographic**
2 **data**

3
4 Stephen Grebby^{a,*}, Dickson Cunningham^a, Jonathan Naden^b, Kevin Tansey^c

5
6 ^aDepartment of Geology, University of Leicester, University Road, Leicester LE1 7RH, UK

7 ^bBritish Geological Survey, Keyworth, Nottingham NG12 5GG, UK

8 ^cDepartment of Geography, University of Leicester, University Road, Leicester LE1 7RH, UK

9
10 *Corresponding author. Tel: +44 (0)116 252 3922; Fax: +44 (0)116 252 3918

11 E-mail address: srg11@le.ac.uk

12

13

14

15

16

17

18

19

20

21 **Keywords:** Airborne LiDAR; lithological mapping; morphometry; Troodos ophiolite; Self-
22 **Organizing Map**

23

24 **Abstract**

25 Traditional field-based lithological mapping can be a time-consuming, costly and challenging
26 endeavour when large areas need to be investigated, where terrain is remote and difficult to
27 access and where the geology is highly variable over short distances. Consequently, rock units
28 are often mapped at coarse-scales, resulting in lithological maps that have generalised contacts
29 which in many cases are inaccurately located. Remote sensing data, such as aerial photographs
30 and satellite imagery are commonly incorporated into geological mapping programmes to obtain
31 geological information that is best revealed by overhead perspectives. However, spatial and
32 spectral limitations of the imagery and dense vegetation cover can limit the utility of traditional
33 remote sensing products. The advent of Airborne Light Detection And Ranging (LiDAR) as a
34 remote sensing tool offers the potential to provide a novel solution to these problems because
35 accurate and high-resolution topographic data can be acquired in either forested or non-forested
36 terrain, allowing discrimination of individual rock types that typically have distinct topographic
37 characteristics. This study assesses the efficacy of airborne LiDAR as a tool for detailed
38 lithological mapping in the upper section of the Troodos ophiolite, Cyprus. Morphometric
39 variables (including slope, curvature and surface roughness) were derived from a 4 m digital
40 terrain model in order to quantify the topographic characteristics of four principal lithologies
41 found in the area. An artificial neural network (the Kohonen Self-Organizing Map) was then
42 employed to classify the lithological units based upon these variables. The algorithm presented
43 here was used to generate a detailed lithological map which defines lithological contacts much
44 more accurately than the best existing geological map. In addition, a separate map of
45 classification uncertainty highlights potential follow-up targets for ground-based verification.
46 The results of this study demonstrate the significant potential of airborne LiDAR for lithological

47 discrimination and rapid generation of detailed lithological maps, as a contribution to
48 conventional geological mapping programmes.

49

50 **1. Introduction**

51 Geological mapping is traditionally carried out by employing field strategies that are best
52 suited to a specific area, including following azimuthal traverses, cross-strike transects, stream
53 sections, ridgetops, bedrock contacts, or moving between individual isolated outcrops (Barnes &
54 Lisle, 2004). However, field mapping in complex and poorly accessible terrain can be
55 challenging, time-consuming and costly (Gad & Kusky, 2007; Grunsky et al., 2009; Rogge et al.,
56 2009). As a consequence, lithologies are often mapped coarsely at reconnaissance (e.g.,
57 1:250,000) or more local scales (e.g., 1:40,000), potentially resulting in geological
58 simplifications and inaccuracies (Roy et al., 2009).

59 Remote sensing data including aerial photographs, and multi- and hyperspectral imagery
60 are also used for lithological mapping (e.g., Drury, 1987; Rothery, 1987; Van der Meer et al.,
61 1997; Rowan & Mars, 2003; Bedini, 2009; Roy et al., 2009). One of the primary benefits of
62 using remote sensing data for lithological mapping is the ability to map areas that are poorly
63 accessible in the field. Although high-resolution aerial photographs can be manually interpreted
64 to help produce detailed lithological maps, the visual discrimination and mapping of surface
65 materials can be subjective, difficult and time-consuming (Crouvi et al., 2006). Multi- and
66 hyperspectral imagery can be automatically classified to rapidly generate lithological maps over
67 large areas, but spatial and spectral limitations of the data may affect the ability to resolve small
68 outcrops or discriminate units with similar spectral properties (Rowan & Mars, 2003; Dong &
69 Leblon, 2004). Dense vegetation cover, such as forests, can also be a hindrance to both field and
70 remote sensing mapping techniques. Whilst making field mapping logistically difficult, dense
71 vegetation also obscures the ground surface and conceals some of the terrain attributes required

72 for photogeological mapping. Additionally, dense vegetation may also obstruct or completely
73 mask the spectral signature of the underlying substrate (Carranza & Hale, 2002).

74 Airborne Light Detection And Ranging (LiDAR) is an emerging active remote sensing
75 technique. It offers a potential solution for overcoming the obscuring effects that dense
76 vegetation has on discrimination of ground materials, as it has the capability of acquiring
77 accurate and high-resolution (ca. 1–4 m) topographic data, even through forest cover (Kraus &
78 Pfeifer, 1998). This is important because individual rock and soil types respond differently to
79 surface processes, such as weathering and erosion, based on their combined mineralogical,
80 petrological and textural characteristics, and thus they typically have distinct topographic
81 characteristics (Kühni & Pfiffner, 2001; Belt & Paxton, 2005). Laser reflections (or returns) from
82 the ground can be separated from vegetation returns to virtually deforest the terrain, enabling the
83 generation of digital terrain models (DTMs; Haugerud & Harding, 2001). The ability to identify
84 subtle topographic features in high-resolution DTMs makes LiDAR an important tool for
85 geosciences research in both vegetated and non-vegetated terrain. Previous geological
86 applications of airborne LiDAR include fault mapping (Harding & Berghoff, 2000; Haugerud et
87 al., 2003; Prentice et al., 2003; Cunningham et al., 2006), mapping and characterisation of
88 landslide morphology (McKean & Roering, 2004; Glenn et al., 2006;) and the characterisation
89 of alluvial fan morphology (Staley et al., 2006; Frankel & Dolan, 2007).

90 Lithological mapping using topographic data is highly dependent upon the recognition of
91 differences in the topographic characteristics between lithologies. Despite its potential for
92 detecting subtle topographic features in vegetated terrain, few studies have assessed the use of
93 airborne LiDAR for lithological mapping. Webster et al. (2006a, 2006b) visually identified
94 subtle topographic differences in a LiDAR-derived DTM and used these to help map three basalt

95 flow units in Nova Scotia, Canada. In comparison to other sources of topographic data, only the
96 LiDAR DTM had the resolution required to identify the subtle contacts between the units.
97 Wallace (2005) quantitatively discriminated three distinct lithological units in the Sudbury Basin,
98 Ontario, Canada, using elevation and morphometric variables of slope and plan, profile,
99 minimum and maximum curvatures derived from a LiDAR DTM. Several lithological maps
100 were also generated through the classification of elevation and slope using a number of
101 conventional classifiers, including the Maximum Likelihood Classification algorithm. In the
102 same study area, Wallace et al. (2006) used fractal dimension analysis to discriminate three
103 lithological units according to differences in topographic roughness. These studies demonstrate
104 the potential of airborne LiDAR for both qualitative and quantitative lithological discrimination
105 and mapping in areas with relatively simple lithological distributions. The use of airborne
106 LiDAR for mapping in more geologically complex terrain, where the spatial distribution of
107 lithologies is more heterogeneous and distinction of different rock units is potentially
108 problematic in itself, has not been demonstrated.

109 The aim of this study is to assess the efficacy of airborne LiDAR for the detailed
110 lithological mapping of a section of the Troodos ophiolite, Cyprus. Given the lithological
111 heterogeneity of the study area, the intention was to develop a semi-automated algorithm to
112 increase the speed and objectivity of the mapping process in comparison to traditional field
113 surveys and visual image interpretation. The algorithm is based on the identification and
114 classification of an optimal set of morphometric variables that were chosen for their ability to
115 discriminate four principal lithological units within the study area. The mapping performance of
116 this algorithm is assessed using conventional classification accuracy statistics and is spatially
117 revealed by mapping the classification uncertainty.

118

119 **2. Study area**

120 The Troodos ophiolite has long been recognised as an uplifted slice of oceanic crust and
121 mantle that was created through sea-floor spreading (Gass, 1968; Moores & Vine, 1971).
122 Forming the central region of the eastern Mediterranean island of Cyprus, the ophiolite displays
123 a dome-like structure centred on Mt Olympus (1,952 m; Fig. 1). The ophiolite stratigraphy
124 includes a mantle sequence consisting of harzburgites, dunites and a serpentinite diapir exposed
125 at the highest elevations. Along the north slope of the range, the mantle sequence is
126 stratigraphically overlain by a largely gabbroic plutonic complex, a sheeted dyke complex,
127 extrusive lavas and oceanic sediments (Varga & Moores, 1985).

128 The study area is located on the northern flank of the Troodos ophiolite (Fig. 1) and
129 comprises a 16 km² area with topographic relief on the order of 200 m. The area has a complex
130 landscape in terms of geology and both natural and anthropogenic influences on topography. The
131 area consists of four main lithological units — the Basal Group lavas and dykes, pillow lavas
132 (Upper and Lower), Lefkara Formation chalky marls and alluvium–colluvium. Conventional
133 field and photogeological mapping, together with some ambiguity in defining the units, is
134 apparently responsible for some considerable differences between the two existing geological
135 maps of this study area (Fig. 2). Despite having a coarser scale, the 1:250,000-scale map is the
136 most recent version and considered to be the most geologically accurate.

137 Stratigraphically, the Basal Group is the lowest unit in the study area. This unit represents
138 a transition from the underlying sheeted dyke complex (100% dykes) to the overlying pillow
139 lavas. Consisted of both dykes and screens of pillow lavas, the definition of the Basal Group is
140 somewhat subjective. In general it contains at least 50% dykes, but more commonly has a dyke

141 abundance of 80–90% dykes (Bear, 1960). Typical Basal Group outcrops can usually be
142 identified in the field according to their relatively high topography and steep relief (Fig. 3a).

143 The pillow lavas are divided into the Upper Pillow Lavas and the Lower Pillow Lavas
144 according to mineralogy, colour and dyke abundance (Wilson 1959; Gass, 1960). However, this
145 division is difficult to apply in the field (Govett & Pantazis, 1971) and an unconformable or
146 transitional boundary between the two lava units has led to uncertainty over this division (Boyle
147 & Robertson, 1984). Due to this ambiguity, the pillow lavas are treated as one unit in this study.
148 In the field, pillow lava terrain is characterised by undulating, hummocky topography (Fig. 3b).
149 Accurate mapping of this unit is crucial to volcanogenic massive sulphide (VMS) mineral
150 exploration on Cyprus, as the Troodos VMS deposits are predominantly confined to the pillow
151 lavas (Constantinou, 1980).

152 Two types of sedimentary cover are present within the study area: the Lefkara Formation
153 and alluvium–colluvium. The Lefkara Formation represents part of the early oceanic
154 sedimentation that was deposited during the late Cretaceous to early Miocene (Kähler & Stow,
155 1998). This formation, which comprises marls, chalks and cherts, directly overlays pillow lavas
156 to form gently rolling hills (Fig. 3c). Alluvium–colluvium refers to Quaternary sediments, such
157 as sand, silts, soils and gravels that were deposited fluvially or through erosion. Alluvial–
158 colluvial cover is characterised by its relatively flat and smooth topography (Fig. 3d), which
159 regularly fills depressions in pillow lava terrain. Alluvial–colluvial cover is frequently exploited
160 for agricultural purposes throughout the study area.

161 Major anthropogenic features are quite scarce and include the Mathiati VMS mine with
162 spoil tips and the village of Agia Varvara Lefkosias in the north. Land disturbances due to
163 agricultural activity are confined to alluvial–colluvial areas and although these occur throughout

164 the study area, they are most commonly found in the north-west. The study area has a semi-arid
165 environment and vegetation cover is relatively dense and widespread, resulting in only small
166 areas of completely exposed rock outcrops. Vegetation cover consists of crops, patchy forests,
167 shrubbery, grasses and lichen. The combination of variable geology, vegetation cover and land-
168 use makes this a particularly complex area for evaluating the application of airborne LiDAR to
169 lithological mapping.

170

171 **3. Airborne LiDAR data and pre-processing**

172 ***3.1 Data acquisition***

173 Airborne LiDAR data were acquired on the 14th May, 2005 by the Natural Environment
174 Research Council Airborne Research and Survey Facility (NERC ARSF). The survey was
175 undertaken at an average flying altitude of 2550 m above sea level, using a Dornier aircraft
176 mounted with an Optech ALTM-3033 system. The aircraft–ground distance ranged between
177 2100–2300 m due to topographic relief within the study area. Operating with a laser pulse
178 repetition rate of 33 kHz and half-scan angle of $\pm 19.4^\circ$ either side of nadir, approximately
179 7,600,000 points were acquired for the study area with an average point density of 0.48 m^{-2} . The
180 dataset contains point data from five overlapping flight lines, each with a swath width of 1400–
181 1500 m and an overlap of 20%–50% between adjacent swaths.

182 Initial data processing was undertaken by the Unit for Landscape Modelling at the
183 University of Cambridge, UK. This involved combining Global Positioning System (GPS) data
184 with the aircraft orientation—recorded using an Inertial Navigation System (INS)—to determine
185 the 3-dimensional coordinates of each laser return (Wehr & Lohr, 1999). The LiDAR point data
186 were delivered as ASCII files containing the x-y-z coordinates and intensity values of all first

187 and last returns in the WGS84 Universal Transverse Mercator (UTM) zone 36-North coordinate
188 system. Information regarding the absolute accuracy of the processed point data was not
189 provided, however the relative vertical accuracy was found to be less than 8 cm as determined
190 from the standard deviation of returns from a flat water surface (Glenn et al., 2006).

191

192 ***3.2. Digital terrain model (DTM) generation***

193 The LiDAR dataset originally contained returns from both ground and non-ground
194 objects, such as trees and buildings. In order to generate a DTM it is necessary to remove all
195 non-ground features from the dataset. Point data were classified as either ground or non-ground
196 returns using a triangulated irregular network (TIN) densification algorithm (Axelsson, 2000),
197 implemented in the TerraScan software (www.terrasolid.fi/en). This algorithm iteratively
198 classifies returns as either ground or non-ground according to angle and distance thresholds
199 applied to TIN facets. Due to the relatively high degree of topographic variability within the
200 study area, the data in individual flight lines were classified separately. In each case the
201 classification parameters and threshold were determined experimentally. The maximum terrain
202 angle and iteration distance threshold were kept constant throughout, at 88° and 1.40 m,
203 respectively. The appropriate maximum building size and iteration angle threshold were found to
204 be more scene-dependent. In general, the maximum building size and iteration angle varied from
205 20 m and 14° for flight lines dominated by relatively high relief, to 60 m and 6° for flight lines
206 acquired over relatively flat terrain. To verify the results of the classification process, several
207 cross-sections were extracted from each flight line and inspected to ensure the point data were
208 assigned to the correct return class. Wherever necessary, misclassified points were manually re-

209 assigned to the correct class. Following classification, non-ground returns were discarded, while
210 points classified as ground returns were used in the generation of the DTM.

211 The accuracy of gridded LiDAR data products is affected by the choice of interpolation
212 algorithm and spatial resolution (Smith et al., 2005; Palamara et al., 2007; Bater & Coops, 2009).
213 It is therefore important to select an appropriate algorithm and resolution in order to avoid errors
214 in the DTM having a significant effect on subsequent morphometric analysis. To determine the
215 most appropriate algorithm and resolution, DTMs were generated at 1, 2, 3, 4 and 5 m
216 resolutions using a range of popular interpolation algorithms. The interpolation algorithms
217 evaluated were inverse distance weighted, block kriging, nearest neighbour, cubic polynomial,
218 modified Shepard's and triangulation with linear interpolation. Interpolation errors associated
219 with each algorithm and resolution were assessed quantitatively using statistics generated
220 through split-sample validation (Smith et al., 2005). This involved the random selection and
221 omission of approximately 9% of the ground returns, while the remaining 91% were used to
222 generate DTMs. The residuals between all omitted data points and their predicted values in the
223 DTM were calculated and used to generate interpolation error statistics, such as the mean error
224 (indicating the magnitude and direction of any bias) and mean absolute error (Bater & Coops,
225 2009). The DTMs were also visually inspected for interpolation artefacts (e.g., null and spurious
226 elevations) using shaded relief images with varying illumination directions and vertical
227 exaggeration. The DTM generation, along with both visual and quantitative interpolation
228 analysis were all undertaken using Surfer 8.0 (Golden Software, Inc.).

229 The split-sample validation results showed that all of the interpolation algorithms tended
230 to underestimate the actual elevation (mean errors ranging from -0.10 m to -0.12 m), with the
231 exception of the triangulation with linear interpolation which slightly overestimated elevation

232 (mean errors ranging from 0.01 m to 0.04 m). Mean absolute errors were generally consistent
233 between the interpolation algorithms and spatial resolutions (ranging from 0.23 m to 0.28 m),
234 except for the triangulation with linear interpolation algorithm for which mean absolute error
235 increased significantly with increasing spatial resolution (from 0.23 m at 1 m resolution to 0.49
236 m at 5 m).

237 During visual inspection, a “ridge and trough” pattern was observed in all DTMs at the
238 extreme edges of areas where adjacent flight lines overlap. Cross-sectional profiles extracted
239 from the flight lines revealed that elevation exhibited an upward concavity error with increasing
240 scan angle towards the edges of swaths — a phenomenon often referred to as “smiley face error”
241 (Lohani & Mason, 2005). Such parabolic vertical error has been attributed to vertical beam
242 misalignment or systematic range errors (Latypov, 2005). The observed DTM artefact is
243 generated when data from multiple flight lines are merged and measurements from large scan
244 angles do not coincide with corresponding measurements from smaller scan angles. The effect of
245 “ridge and trough” artefact on the quantitative analysis was isolated by recalculating the split-
246 sample error statistics using only a subset of residuals selected from outside the areas of overlap
247 (corresponding to ~3% of the total ground returns). As a result, mean errors were reduced to
248 underestimations of between 0.01 m and 0.03 m for all interpolation algorithms except
249 triangulation with linear interpolation, for which the overestimation increased to between 0.02 m
250 and 0.09 m. Also, the choice of interpolation algorithm was found to have a greater effect on
251 mean absolute errors than the spatial resolution, again with the exception of triangulation with
252 linear interpolation. Nevertheless, the mean absolute error showed a significant decrease in all
253 cases when calculated using residuals from outside the areas of overlap. Kriging, modified
254 Shepard’s and cubic polynomial interpolation resulted in the smallest mean absolute errors

255 (ranging from 0.09 m to 0.13 m for all resolutions), followed by the inverse distance weighted
256 and nearest neighbour algorithms (0.15 m to 0.17 m). Triangulation with linear interpolation was
257 the worst performing algorithm, with mean absolute error increasing from 0.12 m at 1 m
258 resolution to 0.43 m at 5 m.

259 As the “ridge and trough” pattern was solely confined to the areas of overlap where the
260 point density is greater, it was possible to almost completely eradicate this artefact from the
261 DTMs using a simple point spacing based filter prior to interpolation. The filter discarded the
262 point with the highest elevation (i.e., the point most affected by “smiley face error”) when
263 multiple ground returns were present within a given radius. The size of the radius was chosen so
264 that the filter only operated on data points within the areas of overlap (in this case a point spacing
265 ≤ 2 m). In addition to removing this artefact, the filter also generates a dataset with a globally
266 uniform point density. The most appropriate interpolation algorithm and spatial resolution for the
267 final DTM was selected as that which minimised the mean and mean absolute errors, and the
268 appearance of interpolation artefacts in the DTM. Consequently, 100% of the ground returns
269 were used to generate the final DTM at a spatial resolution of 4 m, by applying the point-spacing
270 filter prior to interpolation with the kriging algorithm.

271

272

273 4. Methods

274 The efficacy of airborne LiDAR topographic data for detailed lithological mapping is
275 assessed using the methodological approach presented in Fig. 4. Following the generation of the
276 DTM, the method consists of five major steps, which are discussed in the following section.

277

278 ***4.1. Training and validation data***

279 Two independent sets of pixels were selected for the purpose of training and validating
280 the results of the algorithm developed herein. Using knowledge of the study area, QuickBird
281 imagery (0.70 m resolution) and the existing geological maps, four training areas (i.e., regions of
282 interest; ROIs) were carefully selected in ENVI 4.3 (Research Systems, Inc.) to represent the
283 four lithological classes. All pixels located within these four training areas were included in the
284 training dataset. The validation pixels were selected using a random stratified sampling protocol
285 to ensure that each class was represented proportionately and to avoid spatial autocorrelation
286 within the dataset (Chini et al., 2008; Pacifici et al., 2009). To do this, several ROIs were
287 identified for each lithological class in the same way as that used to identify the training areas.
288 Validation pixels were then randomly sampled from these according to the total area of the ROIs
289 associated with each lithological class. Table 1 shows the number of pixels, the equivalent area
290 and the proportion of the study area selected for each lithological class for use in training and
291 validation. In order to determine their effect on the mapping performance, it was decided not to
292 mask-out or treat anthropogenic features as a separate class.

293

294 ***4.2. Morphometric variables***

295 The correlation between lithology and topography that is apparent in the field is also
296 clearly evident in the 4 m DTM of the study area (Fig. 5). In order to automatically classify and
297 map lithology using LiDAR data, it is first necessary to numerically quantify the topographic
298 characteristics of the lithologies using variables that enable adequate discrimination. After
299 considering the observed topographic characteristics, seven candidate morphometric variables
300 were derived from the DTM for this purpose (Table 2).

301 Morphometric variables like slope, plan and profile curvature are typical examples of
302 basic first and second order derivatives of elevation. These three variables were derived using a
303 standard routine in ENVI 4.3, which calculates the derivatives from a quadratic surface fitted to
304 elevations within a moving window (or kernel) that is passed over the DTM (Wood, 1996).
305 Absolute values of plan and profile curvature were used to avoid an alternating pattern of
306 convexity and concavity in highly undulating such as that of the pillow lavas. Morphometric
307 variables such as these are scale-dependent; therefore, in order to identify the most suitable
308 scales for maximum lithological discrimination, each variable was derived using fifteen different
309 moving window sizes ranging from 3×3 pixels (12 m \times 12 m) to 31×31 pixels (124 m \times 124
310 m). Moving window sizes were limited to 31×31 pixels as larger windows were found to reflect
311 more regional-scale topographic information, rather than the local-scale information which is
312 more relevant to detailed lithological discrimination.

313 Relief, hypsometric integral and the two LiDAR-derived measures of surface roughness
314 were derived in Surfer 8.0. Hypsometry describes the elevation distribution within a given area
315 (Strahler, 1952) and can be estimated using the hypsometric integral (Pike & Wilson, 1971). The
316 hypsometric integral (HI) is calculated as:

$$\text{HI} = \frac{h_{\text{mean}} - h_{\text{min}}}{h_{\text{max}} - h_{\text{min}}} \quad (1)$$

317 where h_{mean} , h_{min} and h_{max} are the average, minimum and maximum elevations within a moving
318 window, respectively. This hypsometric integral variable was also derived at multiple scales
319 using the same set of fifteen moving window sizes detailed above.

320 Surface roughness can be measured using the standard deviation of slope within a
321 moving window (Frankel & Dolan, 2007). This variable — referred to here as slope roughness —
322 was derived at multiple scales by first determining slope within a 3×3 pixel window (i.e., 12 m

323 $\times 12$ m) and then calculating the standard deviation of slope within each of the fifteen moving
324 windows. The second measure of surface roughness (known here as residual roughness) is
325 defined as the standard deviation of residual topography (Cavalli et al., 2008). First, a 100 m
326 mean DTM was created by smoothing the 4 m DTM using a 25×25 pixel moving average filter.
327 A residual topographic surface was then calculated by subtracting the 100 m mean DTM from
328 the 4 m DTM. Finally, the standard deviation of this residual topographic surface was calculated
329 within each of the fifteen different sized moving windows.

330 In general, good discrimination and classification performance relies upon homogeneity
331 within classes and dissimilarity between classes (Li et al., 2009). The morphometric
332 homogeneity of the lithologies can be maximised by identifying the optimal scale for each
333 candidate variable. The optimal scales can be determined statistically by identifying the moving
334 windows size which minimises the spread of morphometric data within the training areas (Prima
335 et al., 2006). Here, using the standard deviation of each training area as a measure of its spread,
336 the most suitable moving window size for each candidate variable was defined as that which
337 minimised the average data spread within the training areas. More specifically, for each of the
338 fifteen moving window sizes, the standard deviations within each of the four training areas were
339 calculated and then averaged. The moving window size resulting in the smallest average was
340 deemed to represent the most suitable scale for that variable. This procedure was applied
341 separately to each candidate variable, thus enabling multi-scale topographic information to be
342 utilised. The optimal moving window size for each candidate variable is shown in Table 2.

343

344 *4.3. Variable selection*

345 Classification using all available variables might not necessarily produce the highest
346 mapping accuracy. Some of these variables may be highly correlated, noisy, redundant or
347 irrelevant (Pacifci et al., 2009). Better classification results may be achieved when such input
348 variables are discarded and classification is performed using a smaller set of informative
349 variables (Kavzoglu & Mather, 2002; Verikas & Bacauskiene, 2002). An optimal set of variables
350 can be determined independently of the classification algorithm, based on statistical criteria such
351 as class separability (the filter approach), or in conjunction with the chosen classifier (the
352 wrapper approach). Despite using a non-parametric classifier, a filter approach was adopted as
353 this enabled an exhaustive evaluation of all possible variable combinations to be conducted more
354 efficiently than with a wrapper approach.

355 The number of candidate variables was initially reduced by identifying and discarding
356 linearly correlated and therefore redundant variables through the calculation of Pearson's
357 Product Moment Correlation Coefficients. The optimal set of variables for lithological
358 discrimination was then determined from the remaining candidates through class separability
359 analysis (Dong & Leblon, 2004). To do this, the morphometric separability between pairs of
360 lithological classes (i.e., training areas) was calculated for every combination of two or more
361 variables using the Jeffries-Matusita (JM) distance (Richards, 1994). For four lithologies, there
362 are six possible pairs of classes and therefore six JM distances for each combination of variables.
363 The JM distance ranges from 0–2, with pairs classes being inseparable for JM distances of 0 but
364 completely separable for distances close to 2. The combination of variables resulting in both the
365 largest minimum and largest average JM distances is selected as the optimum for lithological
366 discrimination.

367

368 **4.4. Classification**

369 A lithological map was generated using the optimal set of morphometric variables as
370 inputs to a topologically preserving artificial neural network classifier; the Kohonen Self-
371 Organizing Map (SOM) (Kohonen, 1982, 2001). Artificial neural networks possess many
372 advantages over conventional statistical classifiers, since they are non-parametric, robust in
373 handling noisy data and can learn complex patterns (Ji, 2000). Applications of the SOM to
374 remote sensing data include land-use classification (Ji, 2000; Bagan et al., 2005; Jianwen &
375 Bagan, 2005), lithological mapping (Mather et al., 1998; Bedini, 2009) and geomorphometric
376 feature analysis (Ehsani & Quiel, 2008a, 2008b).

377 The SOM network consists of an input layer and an output layer. The input layer contains
378 one neuron for each of the input variables, whereas the output layer is a two-dimensional array of
379 neurons. Neurons in the output layer are connected to those in the input layer via synaptic
380 weights. Random synaptic weights, ranging from 0 to 1, are initially assigned to the output
381 neurons. These weights are then adjusted during learning to best describe patterns in the input
382 data (Mather et al., 1998). Network learning is an iterative process and involves two stages:
383 unsupervised coarse tuning and supervised fine tuning. The SOM algorithm in IDRISI Andes
384 was used in this study (Li & Eastman, 2006).

385 An input vector (a pixel in morphometric space) is represented by the vector $\mathbf{x} = \{x_1,$
386 $x_2, \dots, x_n\}$, where n is the number of input variables (and input neurons) used in the classification.
387 During coarse tuning, input vectors are presented to the network and in each case the output
388 neuron with the minimum Euclidean distance between its weight vector and the input vector is
389 selected as the winner:

$$\text{winner} = \text{arg min}_j \left(\sqrt{\sum_{i=1}^n (x_i(t) - w_{ji}(t))^2} \right) \quad (2)$$

390 where $x_i(t)$ is the input to neuron i at iteration t and $w_{ji}(t)$ is the synaptic weight connecting output
 391 neuron j to the input neuron i at iteration t . The weight vector of the winner and output neurons
 392 within a neighbourhood of radius γ of the winner are then adjusted in the direction of the input
 393 vector:

$$w_{ji}(t+1) = w_{ji}(t) + \alpha(t)[x_i(t) - w_{ji}(t)] \quad (3)$$

394 where $w_{ji}(t+1)$ is the adjusted weight vector and $\alpha(t)$ is the learning rate at iteration t . The
 395 weights of neurons outside the neighbourhood remain unadjusted. The learning rate decreases
 396 gradually during the coarse tuning stage from an initial learning rate (α_{\max}) to a final learning rate
 397 (α_{\min}), after the total number of iterations (t_{\max}):

$$\alpha(t) = \alpha_{\max} \left(\frac{\alpha_{\min}}{\alpha_{\max}} \right)^{\frac{t}{t_{\max}}} \quad (4)$$

398 Similarly, the radius of the neighbourhood also decreases steadily during the coarse
 399 tuning stage:

$$\gamma(t) = \gamma_{\max} \left(\frac{\gamma_{\min}}{\gamma_{\max}} \right)^{\frac{t}{t_{\max}}} \quad (5)$$

400 A large initial neighbourhood is usually chosen, resulting in widespread adjustments to the
 401 weight vectors of neurons in the output layer. As learning progresses, γ decreases until the
 402 weight of only the winning neuron is adjusted.

403 The SOM network parameters used in this study are based on experimentation guided
 404 using the existing literature (e.g., Ji, 2000; Jianwen & Bagan, 2005; Bedini, 2009). An output
 405 layer consisting of 10×10 neurons was chosen, with $\alpha_{\max} = 0.05$, $\alpha_{\min} = 0.01$ and $\gamma_{\max} = 12$.
 406 Coarse tuning was performed using all input vectors, therefore t_{\max} was equal to the number of
 407 pixels in each input variable image (i.e., 1,012,841 iterations). Prior to learning, the input
 408 variables were normalised to the range 0–1 using a logistic (softmax) function. This function
 409 performs a nearly linear transformation on most of the data whilst also acting to reduce the
 410 influence of any outliers in each variable (Priddy & Keller, 2005). Normalisation increases the
 411 learning efficiency and also ensures that the input variable with the largest range does not
 412 dominate the calculation of the Euclidean distances and the organisation of the output layer
 413 (Ehsani & Quiel, 2008a).

414 Before fine tuning commences, neurons in the output layer must be preliminarily labelled
 415 using input vectors with known class identities. To achieve this, pixels from the training areas
 416 were presented to the coarsely tuned network and in each case the output neuron with the closest
 417 matching weights was triggered. Output neurons were labelled according to the training pixel
 418 class they were triggered by most frequently — a procedure known as majority voting.

419 Fine tuning was performed using the type-one Learning Vector Quantization (LVQ1)
 420 algorithm (Kohonen, 1990). The aim of fine tuning is to improve the classification accuracy by
 421 defining the class boundaries in the output layer more precisely. Pixels within the training areas
 422 were again presented to the SOM and the output neuron with the minimum Euclidean distance
 423 between a training pixel and its weight vector was selected as the Best Matching Unit (BMU).
 424 The weights of the BMU were adjusted accordingly:

$$w_c(t+1) = w_c(t) + \delta(t)[x_i(t) - w_c(t)], \quad \text{if } \mathbf{x} \text{ is correctly labelled} \quad (6)$$

$$425 \quad w_c(t+1) = w_c(t) - \delta(t)[x_i(t) - w_c(t)], \quad \text{if } \mathbf{x} \text{ is incorrectly labelled} \quad (7)$$

$$426 \quad w_i(t+1) = w_i(t), \quad \text{if } i \neq c \quad (8)$$

427 where w_c is the weight vector of the BMU, $w_c(t+1)$ is the adjusted BMU weight vector and $\delta(t)$
 428 is a scalar gain term, which decreases with each iteration like the learning rate during coarse
 429 tuning. Consequently, if the class identity of a training pixel matches the label of its BMU, the
 430 weight vector of the BMU is adjusted in the direction of the training vector, but is moved away if
 431 not. Fine tuning was performed using $\delta_{\max} = 0.005$, which decreases to $\delta_{\min} = 0.001$ after 200
 432 iterations. Output neurons were re-labelled following fine tuning. In order to classify lithology,
 433 all input vectors were presented again to the trained network and assigned the class identity of
 434 their corresponding BMU.

435

436 **4.5. Accuracy assessment**

437 The classification accuracy was assessed by determining the overall (OA), user's (UA)
 438 and producer's (PA) accuracies and the Kappa coefficient (K) from a confusion matrix
 439 (Congalton, 1991). The OA is the percentage of validation data correctly classified, whereas the
 440 UA and PA detail the commission and omission errors, respectively. The K is considered a more
 441 reliable measure of classification accuracy because, unlike the OA, it takes into account the
 442 possibility of agreements occurring by chance in a random classification (Brown et al., 1998;
 443 Pignatti, 2009).

444 In addition to the lithological map, a second map was generated to analyse the spatial
 445 context of classification uncertainties. To do this, the degree of commitment that each pixel has
 446 to its assigned lithological class was determined using the SOM Commitment (SOM-C) (Li &

447 Eastman, in press). Calculated from the triggering proportion of classes on output neurons during
448 labelling, SOM-C essentially provides an indication of classification uncertainty. Values range
449 from 0 to 1, with SOM-C values close to 1 indicating little uncertainty in the class identity of a
450 pixel, whereas values close to 0 indicate high classification uncertainty.

451

452 **5. Results and discussion**

453 *5.1. Variable selection for lithological discrimination*

454 The Pearson's Product Moment Correlation Coefficients revealed that the relief variable
455 was highly linearly correlated ($r > 0.80$) with both the slope and the residual roughness variables.
456 Also, slope roughness showed moderate-to-high positive correlation ($r > 0.54$) with almost all
457 candidate variables. Consequently, the relief and slope roughness variables were deemed to be
458 redundant and discarded, reducing the number of candidate variables from seven to five.

459 Minimum and average JM distances for pairs of lithological classes were computed for
460 all twenty-six combinations of two or more of the five remaining candidate variables (Fig. 6).
461 The minimum and average JM distances are generally smallest when separability is calculated
462 using only pairs of variables and increases when additional variables are included. The slope
463 variable appears to have the greatest influence on the separability, since its exclusion results in at
464 least a 20% and 50% decrease in the minimum and average JM distances, respectively. In terms
465 of the pair-wise class separability, the Lefkara Formation and pillow lavas were consistently the
466 least separable lithological units and were responsible for the minimum JM distance for almost
467 all variable combinations. The lack of morphometric separability between these two units can be
468 attributed to their stratigraphic relationship, where the Lefkara Formation has been deposited
469 directly on top of the pillow lavas. This results in the Lefkara Formation displaying some

470 topographic characteristics of the subdued pillow lava terrain that it drapes. Conversely, the
471 Basal Group and alluvium–colluvium were consistently the most separable units with JM
472 distances typically exceeding 1.90. Such separability is expected due to their contrasting
473 topographic characteristics. Large JM distances were also usually observed between alluvium–
474 colluvium and both the pillow lavas and Lefkara Formation.

475 The combination which includes all five remaining candidate variables is the optimum
476 for lithological discrimination, as this combination resulted in both the largest minimum and
477 largest average JM distances (1.20 and 1.69, respectively). Furthermore, this combination of
478 variables results in the largest JM distances for all six pairs of classes. For this optimal
479 combination, the Lefkara Formation and pillow lavas were the least separable lithologies,
480 followed successively by the Lefkara Formation and Basal Group (JM distance of 1.22), pillow
481 lavas and Basal Group (1.70) and alluvium–colluvium versus all other units (all with JM
482 distances of 2.00). The relative importance of each variable to the separability of lithologies was
483 evaluated by examining the decrease in the JM distances after each variable was removed (Table
484 3). Removing the slope variable produced the largest decrease in the JM distances for all six
485 pairs of lithological classes and the minimum and mean JM distances. This suggests that slope
486 contributes most to the separability of the lithologies in the study area. Apparently, absolute plan
487 curvature is also an important variable; particularly for separating the morphometric
488 characteristics of the Lefkara Formation, Basal Group and pillow lavas. The absolute profile
489 curvature variable is arguably the least important as its removal resulted in the smallest decrease
490 in the minimum, mean and the majority of pair-wise JM distances. Removing the residual
491 roughness and hypsometric integral variables produced a similar decrease in all JM distances,
492 suggesting these are of equal importance. This optimal set of morphometric variables — slope,

493 absolute profile curvature, absolute plan curvature, residual roughness and the hypsometric
494 integral (Fig. 7)—was subsequently used in the classification stage.

495

496 ***5.2. Lithological mapping and accuracy assessment***

497 A lithological map displaying the four principal units and a SOM-C map, indicating the
498 classification uncertainty, were generated using the LiDAR-derived topographic data (Fig. 8).
499 Following classification, a small amount of noise in the classified image was reduced using a 3 ×
500 3 mode filter.

501 The accuracy of the lithological map was assessed using the validation pixels and the
502 results were summarised using a confusion matrix (Table 4). The lithological map has an overall
503 accuracy of 65.4% and a K of 0.53. Alluvium–colluvium is the best mapped unit with a
504 producer’s accuracy of 87.9% and a user’s accuracy of 98.8%, while the Lefkara Formation was
505 mapped with the least accuracy. A good producer’s classification accuracy was achieved for the
506 pillow lavas (66.8%), however more than 50% of all validation pixels mapped as pillow lavas
507 actually belong to other classes. Only 50.4% of Basal Group validation pixels were mapped
508 correctly, but with a commission error of just 29.7%. The most classification confusion occurs
509 between the Lefkara Formation, pillow lavas and Basal Group, which corroborates the results of
510 the separability analysis. Although the majority of this confusion can be explained by their
511 stratigraphic relationships or natural deviations from the typical topographic characteristics of
512 each unit, anthropogenic activity is also responsible for a significant component. An obvious
513 example of this can be found proximal to Mathiati mine and spoil tips where the natural
514 topographic characteristics have been destroyed, leading to misclassification (Fig. 8).

515 Through comparison with the QuickBird imagery, it is clear that the algorithm is capable
516 of defining lithological contacts more accurately than the best existing geological map (Fig. 9).
517 Furthermore, the algorithm can be used to generate a more detailed lithological map by
518 identifying lithologies in areas that have not been mapped previously. The SOM-C map is useful
519 for highlighting areas of uncertainty in the lithological map. In general, SOM-C values less than
520 0.75 correspond to areas with a high degree of classification uncertainty, as clearly illustrated by
521 the portion of Lefkara Formation incorrectly classified as pillow lavas (Fig. 9). In this particular
522 case, the confusion is related to the difficulty in detecting the ground beneath some types of low-
523 lying vegetation using airborne LiDAR. The class containing SOM-C values of 0–0.7 consists
524 solely of SOM-C values of 0. These values are due to unlabelled neurons in the output layer
525 which were not triggered by any of the training pixels (Li & Eastman, in press). For the purpose
526 of classification, unlabelled neurons were assigned class labels using a minimum distance
527 auxiliary labelling algorithm (Li & Eastman, 2006), resulting in no unclassified pixels in the
528 lithological map. Pixels in the lithological map with corresponding SOM-C values of 0 do not
529 necessarily possess a higher degree of uncertainty than pixels associated with larger SOM-C
530 values. The uncertainty of pixels classified using the auxiliary labelling algorithm is case
531 specific. Examples where such SOM-C values correspond to both correct and incorrect
532 classification are evident throughout the study area and therefore each case should be considered
533 individually. Frequent misclassifications occurring at the contacts between agricultural
534 alluvium–colluvium and upstanding Lefkara Formation outcrops are highlighted by SOM-C
535 values of 0. Ploughing proximal to the contacts is responsible for pixels with atypical
536 topographic characteristics, which results in them being incorrectly classified as pillow lavas
537 through the auxiliary labelling algorithm.

538 The accuracy of the lithological map produced in this study is higher than the accuracies
539 reported by Wallace (2005) who investigated an area with a simpler lithological outcrop pattern.
540 In contrast to Wallace's (2005) study, our analysis involves a larger number of morphometric
541 variables and a more complex classification algorithm. In addition, the distribution of the pillow
542 lavas, Basal Group and overlying sediments is more complex because they are separated by low-
543 angle contacts and are differentially eroded. Therefore, there is no simple strike-belt pattern.
544 Given the geological complexity and anthropogenic factors affecting the topography in this study
545 area, we consider the results of our algorithm to be good. Additionally, the algorithm was
546 implemented using minimal a priori knowledge of the spatial distribution of each lithological
547 unit. However, higher mapping accuracies can be achieved using more a priori knowledge.
548 Doubling the total number of training pixels (to approximately 2% of the total number of pixels
549 within the study area) increases the overall accuracy to 67.3% and K to 0.56 when the same
550 SOM network parameters are used. The ability to produce good mapping results given limited
551 knowledge regarding the spatial distribution of units makes this algorithm particularly relevant to
552 mapping relatively unexplored terrain.

553

554 **6. Conclusions**

555 This study assesses the efficacy of airborne LiDAR topographic data for detailed
556 lithological mapping of a geologically complex area of the Troodos ophiolite, Cyprus. Typical
557 topographic characteristics associated with each of the lithologies were recognised in a 4 m
558 LiDAR DTM and quantified using a morphometric approach. An optimal set of morphometric
559 variables for lithological discrimination were identified and used in conjunction with a SOM
560 classifier to produce a lithological map. The resulting map achieved an overall accuracy of

561 65.4% and a K of 0.53, which is considered good given the complexity of the study area and the
562 lack of a priori knowledge. The lithological map is more detailed than the best existing
563 geological map and the lithological contacts are more accurately defined. The results of this
564 study demonstrate the significant potential of airborne LiDAR as a tool for generating detailed
565 lithological maps over large areas of either forested or non-forested terrain, where conventional
566 methods are of limited use. Furthermore, the SOM-C map highlights areas with high
567 classification uncertainty, therefore providing information regarding follow-up targets for
568 efficient ground-based verification.

569 Further studies are required to assess whether improvements in the lithological mapping
570 accuracy can be made through the integration of airborne LiDAR data with high-resolution
571 multispectral imagery. It is anticipated that the multispectral imagery will help to reduce
572 misclassification in non-vegetated areas where the natural topographic characteristics of the
573 various rock types have been destroyed by anthropogenic activity.

574 The detailed lithological map generated in this study represents a valuable aid to VMS
575 mineral exploration in the Troodos ophiolite because the mapped distribution of potential host
576 rocks is now much better resolved than on previous maps. In addition, the efficacy of this
577 algorithm extends to other geological settings where lithology and topography are positively
578 correlated, with exciting implications beyond mineral exploration. In particular, the relative ease
579 with which basement rocks and sedimentary cover can be discriminated at high-resolution could
580 be useful in all terrains from open ground to densely forested landscapes for: 1) identifying local
581 areas for groundwater extraction, 2) locating areas with enhanced agricultural potential, and 3)
582 for general infrastructure planning where it is important to know construction site substrates.

583 Thus the methods presented here may have widespread utility for a range of applications,
584 especially in areas of mixed basement and sedimentary cover exposure.

585

586 **Acknowledgements**

587 This work was primarily supported through a NERC CASE Studentship (NE/F00673X/1) in
588 collaboration with the British Geological Survey University Funding Initiative, awarded to SG
589 and NERC ARSF grant MC04/30 awarded to JN. We would like to express our gratitude to the
590 Geological Survey Department of Cyprus (GSD) for providing the digital geological maps and
591 QuickBird imagery, and Dr. Stelios Nicolaides (GSD) and Dr. Simon Jowitt (Monash University,
592 Australia) for invaluable logistical and scientific help in the field. SG is grateful to the
593 Geological Remote Sensing Group for a Student Fieldwork and Travel Award and to Professor
594 Danny Donoghue of Durham University for Terrasolid software support and advice. We also
595 thank the three anonymous reviewers for their comments and suggestions which helped to
596 improve the manuscript.

597

598 **References**

- 599 Axelsson, P. (2000). DEM Generation from Laser Scanner Data Using Adaptive TIN Models.
600 *International Archives of Photogrammetry and Remote Sensing*, 33 (B4/1), 110–117.
- 601 Bagan, H., Wang, Q.X., Watanabe, M., Yang, Y.H., & Jianwen, M. (2005). Land cover
602 classification from MODIS EVI times-series data using SOM neural network.
603 *International Journal of Remote Sensing*, 26, 4999–5012.
- 604 Barnes, J.W., & Lisle, R. J. (2004). *Basic Geological Mapping, Fourth edition*. (pp. 43–48).
605 Chichester: John Wiley & Sons.

- 606 Bater, C.W., & Coops, N.C. (2009). Evaluating error associated with lidar-derived DEM
607 interpolation. *Computers & Geosciences*, 35, 289–300.
- 608 Bear, L.M. (1960). *The geology and mineral resources of the Akaki-Lythrodondha area: Cyprus*
609 Geological Survey Department Memoir 3.
- 610 Bedini, E. (2009). Mapping lithology of the Sarfartoq carbonatite complex, southern West
611 Greenland, using HyMap imaging spectrometer data. *Remote Sensing of Environment*,
612 113, 1208–1219.
- 613 Belt, K., & Paxton, S.T. (2005). GIS as an aid to visualizing and mapping geology and rock
614 properties in regions of subtle topography. *Geological Society of America Bulletin*, 117,
615 149–160.
- 616 Boyle, J.F., & Robertson, A.H.F. (1984). Evolving metallogenesis at the Troodos spreading axis.
617 *Geological Society, London, Special Publications*, 13, 169–181.
- 618 Brown, D.G., Lusch, D.P., & Duda, K.A. (1998). Supervised classification of types of glaciated
619 landscapes using digital elevation data. *Geomorphology*, 21, 233–250.
- 620 Carranza, E.J.M., & Hale, M. (2002). Mineral imaging with Landsat Thematic Mapper data for
621 hydrothermal alteration mapping in heavily vegetated terrane. *International Journal of*
622 *Remote Sensing*, 23, 4827–4852.
- 623 Cavalli, M., Tarolli, P., Marchi, L., & Fontana, G.D. (2008). The effectiveness of airborne
624 LiDAR data in the recognition of channel-bed morphology. *Catena*, 73, 249–260.
- 625 Chini, M., Pacifici, F., Emery, W.J., Pierdicca, N., & Del Frate, F. (2008). Comparing statistical
626 and neural network methods applied to very high resolution satellite images showing
627 changes in man-made structures at rocky flats. *IEEE Transactions on Geoscience and*
628 *Remote Sensing*, 46, 1812–1821.

- 629 Congalton, R.G. (1991). A review of assessing the accuracy of classifications of remotely sensed
630 data. *Remote Sensing of Environment*, 37, 35–46.
- 631 Constantinou, G. (1980). Metallogensis associated with Troodos ophiolite. In A. Panayiotou
632 (Ed.), *Ophiolites: Proceedings of the International Ophiolite Symposium, Cyprus, 1979*
633 (pp. 663–674). Nicosia, Cyprus: Geological Survey Department.
- 634 Crouvi, O., Ben-Dor, E., Beyth, M., Avigad, D., & Amit, R. (2006). Quantitative mapping of
635 arid alluvial fan surfaces using field spectrometer and hyperspectral remote sensing.
636 *Remote Sensing of Environment*, 104, 103–117.
- 637 Cunningham, D., Grebby, S., Tansey, K., Gosar, A., & Kastelic, V. (2006). Application of
638 airborne LiDAR to mapping seismogenic faults in forested mountainous terrain,
639 southeastern Alps, Slovenia. *Geophysical Research Letters*, 33, L20308.
- 640 Dong, P., & Leblon, B. (2004). Rock unit discrimination on Landsat TM, SIR-C and Radarsat
641 images using spectral and textural information. *International Journal of Remote Sensing*,
642 25, 3745–3768.
- 643 Drury, S.A. (1987). *Image Interpretation in Geology*. (pp. 64–88). London: Allen & Unwin.
- 644 Ehsani, A.H., & Quiel, F. (2008a). Geomorphometric feature analysis using morphometric
645 parameterization and artificial neural networks. *Geomorphology*, 99, 1–12.
- 646 Ehsani, A.H., & Quiel, F. (2008b). Application of self organizing map and SRTM data to
647 characterize yardangs in the Lut desert, Iran. *Remote Sensing of Environment*, 112, 3284–
648 3294.
- 649 Frankel, K.L., & Dolan, J.F. (2007). Characterizing arid region alluvial fan surface roughness
650 with airborne laser swath mapping digital topographic data. *Journal of Geophysical*
651 *Research–Earth Surface*, 112, F02025.

- 652 Gad, S., & Kusky, T. (2007). ASTER spectral ratioing for lithological mapping in the arabian-
653 nubian shield, the neoproterozoic Wadi Kid area, Sinai, Egypt. *Gondwana Research*, *11*,
654 326–335.
- 655 Gass, I.G. (1960). *The geology and mineral resources of the Dhali area*: Cyprus Geological
656 Survey Department Memoir 4.
- 657 Gass, I.G. (1968). Is the Troodos massif of Cyprus a fragment of Mesozoic ocean crust? *Nature*,
658 *220*, 39–42.
- 659 Glenn, N.F., Streutker, D.R., Chadwick, D.J., Thackray, G.D., & Dorsch, S.J. (2006). Analysis
660 of LiDAR-derived topographic information for characterizing and differentiating
661 landslide morphology and activity. *Geomorphology*, *73*, 131–148.
- 662 Govett, G.J.S., & Pantazis, T.M. (1971). Distribution of Cu, Zn, Ni and Co in the Troodos Pillow
663 Lava Series, Cyprus. *Transactions of the Institution of Mining and Metallurgy, Section B:*
664 *Applied Earth Science*, *80*, 27–46.
- 665 Grunsky, E., Harris, J., & McMartin, I. (2009). Predictive mapping of surficial materials, Schultz
666 Lake Area (NTS 66A), Nunavut, Canada. *Reviews in Economic Geology*, *16*, 177–198.
- 667 Harding, D.J., & Berghoff, G.S. (2000). Fault scarp detection beneath dense vegetation cover:
668 airborne lidar mapping of the Seattle fault zone, Bainbridge Island, Washington State. In,
669 *Proceedings of the American Society of Photogrammetry and Remote Sensing Annual*
670 *Conference*. Washington, D.C.
- 671 Haugerud, R.A., & Harding, D.J. (2001). Some algorithms for virtual deforestation (VDF) of
672 lidar topographic survey data. *International Archives of Photogrammetry and Remote*
673 *Sensing*, *34* (3/W4), 211–217.
- 674 Haugerud, R.A., Harding, D.J., Johnson, S.Y., Harless, J.L., Weaver, C.S., & Sherrod, B.L.

- 675 (2003). High-resolution Lidar topography of the Puget Lowland, Washington — A
676 bonanza for earth science. *GSA Today*, 13, 4–10.
- 677 Ji, C.Y. (2000). Land-use classification of remotely sensed data using Kohonen Self-Organizing
678 Feature Map neural networks. *Photogrammetric Engineering and Remote Sensing*, 66,
679 1451–1460.
- 680 Jianwen, M., & Bagan, H. (2005). Land-use classification using ASTER data and self-organized
681 neural networks. *International Journal of Applied Earth Observation and*
682 *Geoinformation*, 7, 183–188.
- 683 Kähler, G., & Stow, D.A.V. (1998). Turbidites and contourites of the Palaeogene Lefkara
684 Formation, southern Cyprus. *Sedimentary Geology*, 115, 215–231.
- 685 Kavzoglu, T., & Mather, P.M. (2002). The role of feature selection in artificial neural network
686 applications. *International Journal of Remote Sensing*, 23, 2919–2937.
- 687 Kohonen, T. (1982). Self-organized formation of topologically correct feature maps. *Biological*
688 *Cybernetics*, 43, 59–69.
- 689 Kohonen, T. (1990). The self-organizing map. *Proceedings of the IEEE*, 78, 1464–1480.
- 690 Kohonen, T. (2001). *Self-Organizing Maps, Third edition*. Berlin: Springer.
- 691 Kraus, K., & Pfeifer, N. (1998). Determination of terrain models in wooded areas with airborne
692 laser scanner data. *ISPRS Journal of Photogrammetry and Remote Sensing*, 53, 193–203.
- 693 Kühni, A., & Pfiffner, O.A. (2001). The relief of the Swiss Alps and adjacent areas and its
694 relation to lithology and structure: topographic analysis from a 250-m DEM.
695 *Geomorphology*, 41, 285–307.
- 696 Latypov, D. (2005). Effects of laser beam alignment tolerance on lidar accuracy. *ISPRS Journal*
697 *of Photogrammetry and Remote Sensing*, 59, 361–368.

- 698 Li, L.F., Wang, J.F., & Leung, H. (2009). A knowledge-based similarity classifier to stratify
699 sample units to improve the estimation precision. *International Journal of Remote*
700 *Sensing*, 30, 1207–1234.
- 701 Li, Z., & Eastman, J.R. (2006). The nature and classification of unlabelled neurons in the use of
702 Kohonen's self-organizing map for supervised classification. *Transactions in GIS*, 10,
703 599–613.
- 704 Li, Z., & Eastman, J.R. (in press). Commitment and typicality measures for the Self-Organizing
705 Map. *International Journal of Remote Sensing*.
- 706 Liberti, M., Simoniello, T., Carone, M.T., Coppola, R., D'Emilio, M., & Macchiato, M. (2009).
707 Mapping badland areas using LANDSAT TM/ETM satellite imagery and morphological
708 data. *Geomorphology*, 106, 333–343.
- 709 Lohani, B., & Mason, D.C. (2005). A case study on error identification and minimisation in
710 airborne altimetric LiDAR data. *International Journal of Geoinformatics*, 1, 53–61.
- 711 Mather, P.M., Tso, B., & Koch, M. (1998). An evaluation of Landsat TM spectral data and SAR-
712 derived textural information for lithological discrimination in the Red Sea Hills, Sudan.
713 *International Journal of Remote Sensing*, 19, 587–604.
- 714 McKean, J., & Roering, J. (2004). Objective landslide detection and surface morphology
715 mapping using high-resolution airborne laser altimetry. *Geomorphology*, 57, 331–351.
- 716 Moores, E.M., & Vine, F.J. (1971). Troodos Massif, Cyprus and other ophiolites as oceanic
717 crust: evaluation and implications. *Philosophical Transactions of the Royal Society of*
718 *London A*, 268, 443–467.
- 719 Pacifici, F., Chini, M., & Emery, W.J. (2009). A neural network approach using multi-scale
720 textural metrics from very high-resolution panchromatic imagery for urban land-use

- 721 classification. *Remote Sensing of Environment*, 113, 1276–1292.
- 722 Palamara, D.R., Nicholson, M., Flentje, P., Baafi, E., & Brassington, G.M. (2007). An evaluation
723 of airborne laser scan data for coalmine subsidence mapping. *International Journal of*
724 *Remote Sensing*, 28, 3181–3203.
- 725 Pignatti, S., Cavalli, R.M., Cuomo, V., Fusilli, L., Pascucci, S., Poscolieri, M., & Santini, F.
726 (2009). Evaluating Hyperion capability for land cover mapping in a fragmented
727 ecosystem: Pollino National Park, Italy. *Remote Sensing of Environment*, 113, 622–634.
- 728 Pike, R.J., & Wilson, S.E. (1971). Elevation-relief ratio, hypsometric integral, and geomorphic
729 area-altitude analysis. *Geological Society of America Bulletin*, 82, 1079–1084.
- 730 Prentice, C.S., Crosby, C.J., Harding, D.J., Haugerud, R.A., Merritts, D.J., Gardner, T.W.,
731 Koehler, R.D., & Baldwin, J.N. (2003). Northern California LIDAR data: A tool for
732 mapping the San Andreas Fault and pleistocene marine terraces in heavily vegetated
733 terrain. In, *American Geophysical Union Fall Meeting*. Abstract G12A-06.
- 734 Priddy, K.L., & Keller, P.E. (2005). *Artificial neural networks: an introduction*. (pp. 15–17).
735 Bellingham, WA: SPIE Press.
- 736 Prima, O.D.A., Echigo, A., Yokoyama, R., & Yoshida, T. (2006). Supervised landform
737 classification of Northeast Honshu from DEM-derived thematic maps. *Geomorphology*,
738 78, 373–386.
- 739 Richards, J.A. (1994). *Remote sensing digital image analysis: an introduction*. Berlin: Springer-
740 Verlag.
- 741 Rogge, D.M., Rivard, B., Harris, J., & Zhang, J. (2009). Application of hyperspectral data for
742 remote predictive mapping, Baffin Island, Canada. *Reviews in Economic Geology*, 16,
743 209–222.

- 744 Rothery, D.A. (1987). Improved discrimination of rock units using Landsat Thematic Mapper
745 imagery of the Oman ophiolite. *Journal of the Geological Society*, 144, 587–597.
- 746 Rowan, L.C., & Mars, J.C. (2003). Lithologic mapping in the Mountain Pass, California area
747 using Advanced Spaceborne Thermal Emission and Reflection Radiometer (ASTER)
748 data. *Remote Sensing of Environment*, 84, 350–366.
- 749 Roy, R., Launeau, P., Carrere, V., Pinet, P., Ceuleneer, G., Clenet, H., Daydou, Y., Girardeau, J.,
750 & Amri, I. (2009). Geological mapping strategy using visible near-infrared-shortwave
751 infrared hyperspectral remote sensing: Application to the Oman ophiolite (Sumail
752 Massif). *Geochemistry Geophysics Geosystems*, 10, Q02004.
- 753 Smith, S.L., Holland, D.A., & Longley, P.A. (2005). Quantifying Interpolation Errors in Urban
754 Airborne Laser Scanning Models. *Geographical Analysis*, 37, 200–224.
- 755 Staley, D.M., Wasklewicz, T.A., & Blaszczynski, J.S. (2006). Surficial patterns of debris flow
756 deposition on alluvial fans in Death Valley, CA using airborne laser swath mapping data.
757 *Geomorphology*, 74, 152–163.
- 758 Strahler, A.N. (1952). Hypsometric (area-altitude) analysis of erosional topography. *Geological*
759 *Society of America Bulletin*, 63, 1117–1142.
- 760 Van der Meer, F., Vazquez-Torres, M., & Van Dijk, P.M. (1997). Spectral characterization of
761 ophiolite lithologies in the Troodos Ophiolite complex of Cyprus and its potential in
762 prospecting for massive sulphide deposits. *International Journal of Remote Sensing*, 18,
763 1245–1257.
- 764 Varga, R.J., & Moores, E.M. (1985). Spreading structure of the Troodos ophiolite, Cyprus.
765 *Geology*, 13, 846–850.
- 766 Verikas, A., & Bacauskiene, M. (2002). Feature selection with neural networks. *Pattern*

- 767 *Recognition Letters*, 23, 1323–1335.
- 768 Wallace, J., Morris, B., & Howarth, P. (2006). Identifying structural trend with fractal dimension
769 and topography. *Geology*, 34, 901–904.
- 770 Wallace, J.A. (2005). LiDAR altimetry and hyperspectral imaging: new technologies for
771 geological and mineralogical mapping. *PhD thesis*, University of Waterloo, Ontario,
772 Canada.
- 773 Webster, T.L., Murphy, J.B., & Gosse, J.C. (2006a). Mapping subtle structures with light
774 detection and ranging (LIDAR): flow units and phreatomagmatic rootless cones in the
775 North Mountain Basalt, Nova Scotia. *Canadian Journal of Earth Sciences*, 43, 157–176.
- 776 Webster, T.L., Murphy, J.B., Gosse, J.C., & Spooner, I. (2006b). The application of lidar-derived
777 digital elevation model analysis to geological mapping: an example from the Fundy
778 Basin, Nova Scotia, Canada. *Canadian Journal of Remote Sensing*, 32, 173–193.
- 779 Wehr, A., & Lohr, U. (1999). Airborne laser scanning — an introduction and overview. *ISPRS*
780 *Journal of Photogrammetry and Remote Sensing*, 54, 68–82.
- 781 Wilson, R.A.M. (1959). *The geology of the Xeros-Troodos area*: Cyprus Geological Survey
782 Department Memoir 1.
- 783 Wood, J.D. (1996). The geomorphological characterisation of digital elevation models. *PhD*
784 *thesis*, University of Leicester, Leicester, UK.
- 785

786 **Figure captions**

787 Fig. 1. Location of the study area (dashed box) and simplified geology of the Troodos ophiolite.
788 Digital geology was provided by the Geological Survey Department of Cyprus.

789

790 Fig. 2. Existing geological maps of the study area shown in Fig. 1. (a) 1:250,000 and (b)
791 1:31,680-scale maps adapted from the digital geology provided by Geological Survey
792 Department of Cyprus. M–Mathiati mine and A–Agia Varvara Lefkosias.

793

794 Fig. 3. Field photographs showing the four main lithological units: (a) Basal Group, (b) pillow
795 lavas, (c) quarry exposure of the Lefkara Formation overlying pillow lavas (LF and PL,
796 respectively) and (d) alluvium–colluvium (AC).

797

798 Fig. 4. Flow diagram presenting the methodological approach implemented to assess the efficacy
799 of airborne LiDAR for detailed lithological mapping.

800

801 Fig. 5. Shaded relief DTM of the study area displaying the distinct topographic characteristics of:
802 (a) alluvium–colluvium, (b) Basal Group, (c) Lefkara Formation and (d) pillow lavas.

803

804 Fig. 6. Minimum and average separability (JM distance) for combinations of the slope (s),
805 absolute profile curvature (pr), absolute plan curvature (pl), residual roughness (r) and
806 hypsometric integral (h) variables.

807

808 Fig. 7. Optimal set of (normalised) morphometric variables selected as inputs to the SOM
809 classification: (a) slope, (b) absolute profile curvature, (c) absolute plan curvature, (d) residual
810 roughness and (e) hypsometric integral.

811

812 Fig. 8. (a) Lithological map of the study area generated using LiDAR-derived topographic data.
813 The dashed black box indicates the spatial extent of Fig. 9. (b) SOM-C map depicting
814 classification uncertainty.

815

816 Fig. 9. Detailed view of the mapping performance for the area shown in Fig. 8. (a) QuickBird
817 image, (b) lithological map generated using LiDAR-derived topographic data and (c) SOM-C
818 map. The white dashed line represents the pillow lava–Lefkara Formation contact from the
819 1:250,000-scale geological map in Fig. 2a.

820

821

Table 1. Number of pixels, the equivalent area and the proportion of the study area (PS) selected for each lithological class for training and validation purposes.

Lithological class	Training			Validation		
	Pixels	Area (m ²)	PS (%)	Pixels	Area (m ²)	PS (%)
Alluvium–colluvium	1712	27,392	0.17	4087	65,392	0.40
Basal Group	1780	28,480	0.18	3200	51,200	0.32
Lefkara Formation	2769	44,304	0.27	2451	39,216	0.24
Pillow lavas	3095	49,520	0.31	3208	51,328	0.32

Table 2. Candidate morphometric variables for lithological discrimination.

Morphometric variable	Description	Optimal moving window size (pixels)
Slope (°)	Magnitude of the steepest gradient	15 × 15
Relief (m)	Elevation range within a given area	3 × 3
Profile curvature (1/m)	Absolute value of vertical curvature component in aspect direction	21 × 21
Plan curvature (1/m)	Absolute value of horizontal curvature component in aspect direction	31 × 31
Slope roughness (°)	Standard deviation of slope	31 × 31
Residual roughness (m)	Standard deviation of residual topography	3 × 3
Hypsometric integral	Elevation distribution within a given area	11 × 11

Table 3. The relative importance of variables to the separability of lithologies, determined by individually removing each variable from the pair-wise JM distance calculations.

Variable removed	JM distance						Min.	Mean
	LF vs. PL	LF vs. BG	PL vs. BG	LF vs. AC	PL vs. AC	BG vs. AC		
None	1.20	1.22	1.70	2.00	2.00	2.00	1.20	1.69
Slope	0.27	0.50	0.41	1.92	1.95	1.94	0.27	1.17
Profile curvature	1.17	1.14	1.67	2.00	1.99	2.00	1.14	1.66
Plan curvature	0.81	1.02	1.59	2.00	1.99	2.00	0.81	1.57
Residual roughness	1.09	1.10	1.67	2.00	1.97	2.00	1.09	1.64
Hypsometric integral	1.05	1.13	1.65	2.00	1.99	2.00	1.05	1.64

LF, Lefkara Formation; PL, pillow lavas; BG, Basal Group; AC, alluvium–colluvium.

Table 4. Confusion matrix for SOM classification using the optimal set of morphometric variables.

Mapped as	Validation data				Row total	User's accuracy (%)
	Alluvium–colluvium	Basal Group	Lefkara Formation	Pillow lavas		
Alluvium–colluvium	3594	1	30	11	3636	98.8
Basal Group	0	1614	299	383	2296	70.3
Lefkara Formation	2	816	1114	672	2604	42.8
Pillow lavas	491	769	1008	2142	4410	48.6
Column total	4087	3200	2451	3208		
Producer's accuracy (%)	87.9	50.4	45.4	66.8		
Overall accuracy = 65.4%						
K = 0.53						

Figure 1
[Click here to download high resolution image](#)

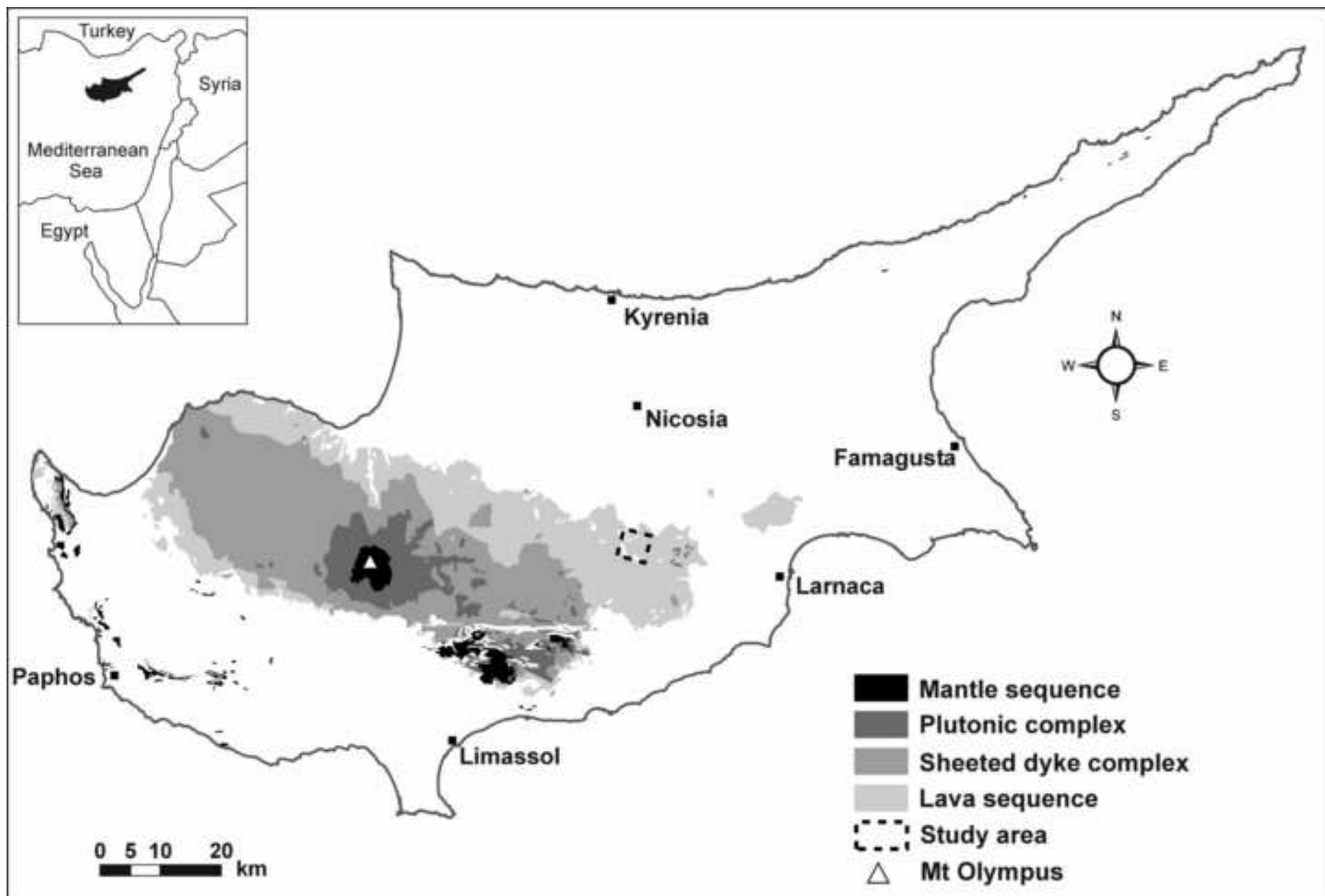


Figure 2
[Click here to download high resolution image](#)

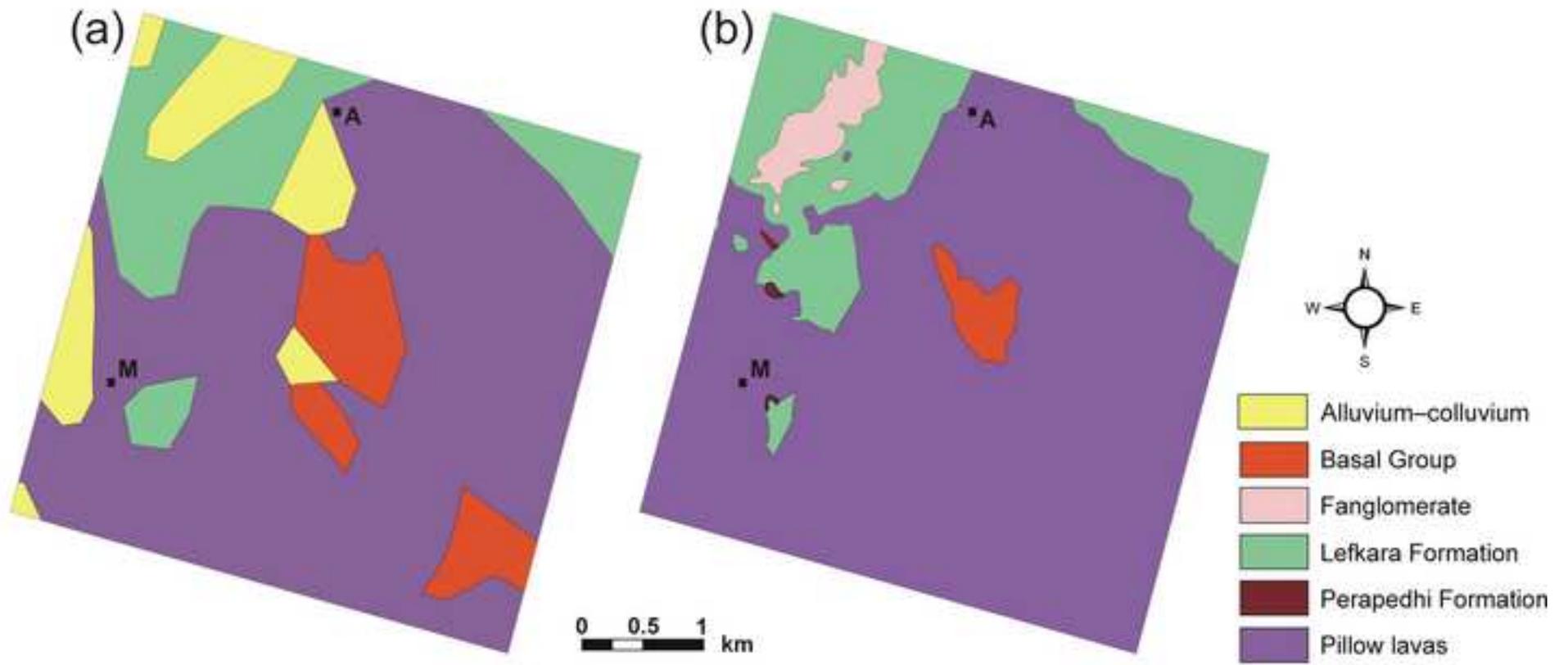


Figure 3
[Click here to download high resolution image](#)



Figure 4
[Click here to download high resolution image](#)

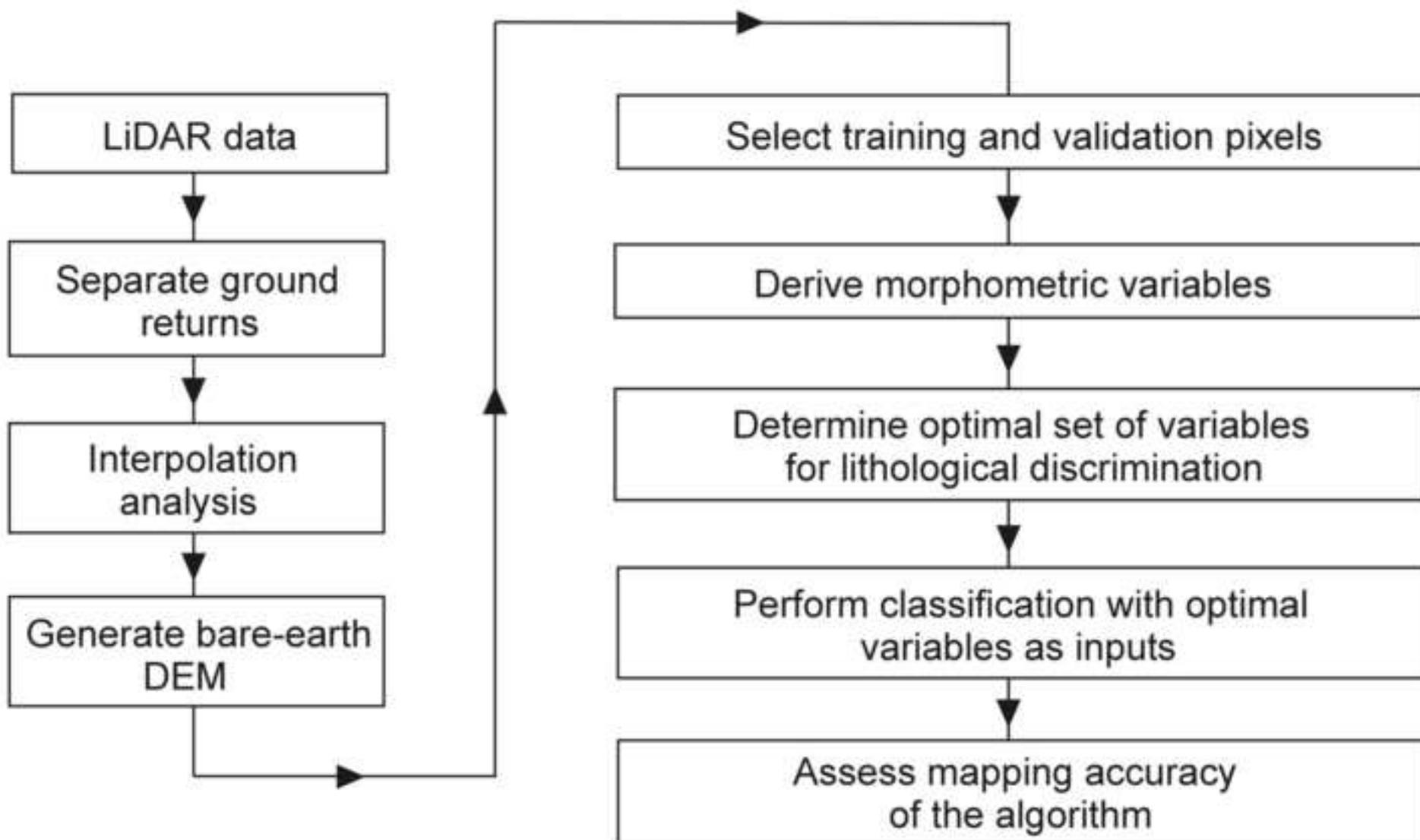


Figure 5
[Click here to download high resolution image](#)

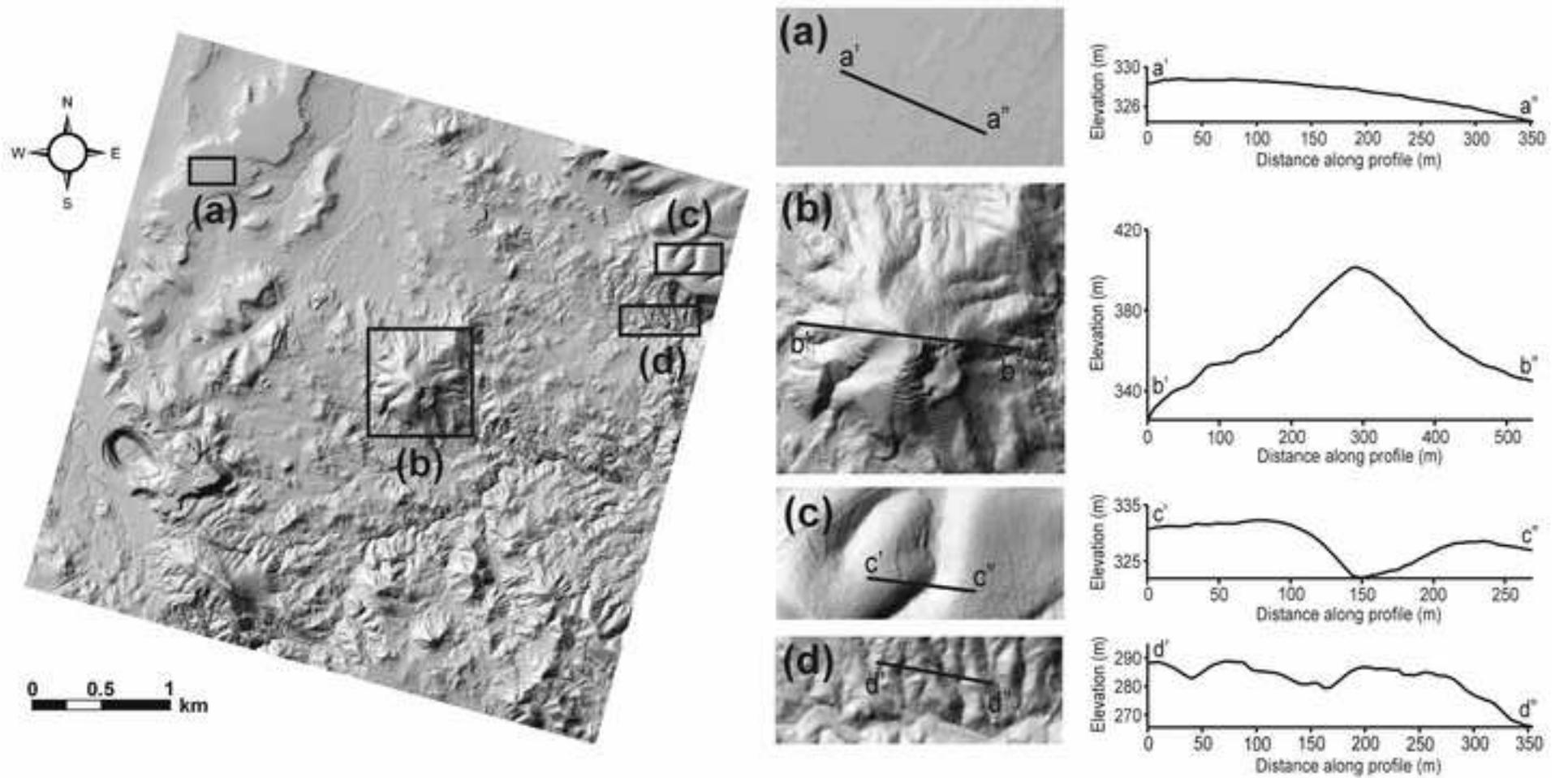


Figure 6
[Click here to download high resolution image](#)

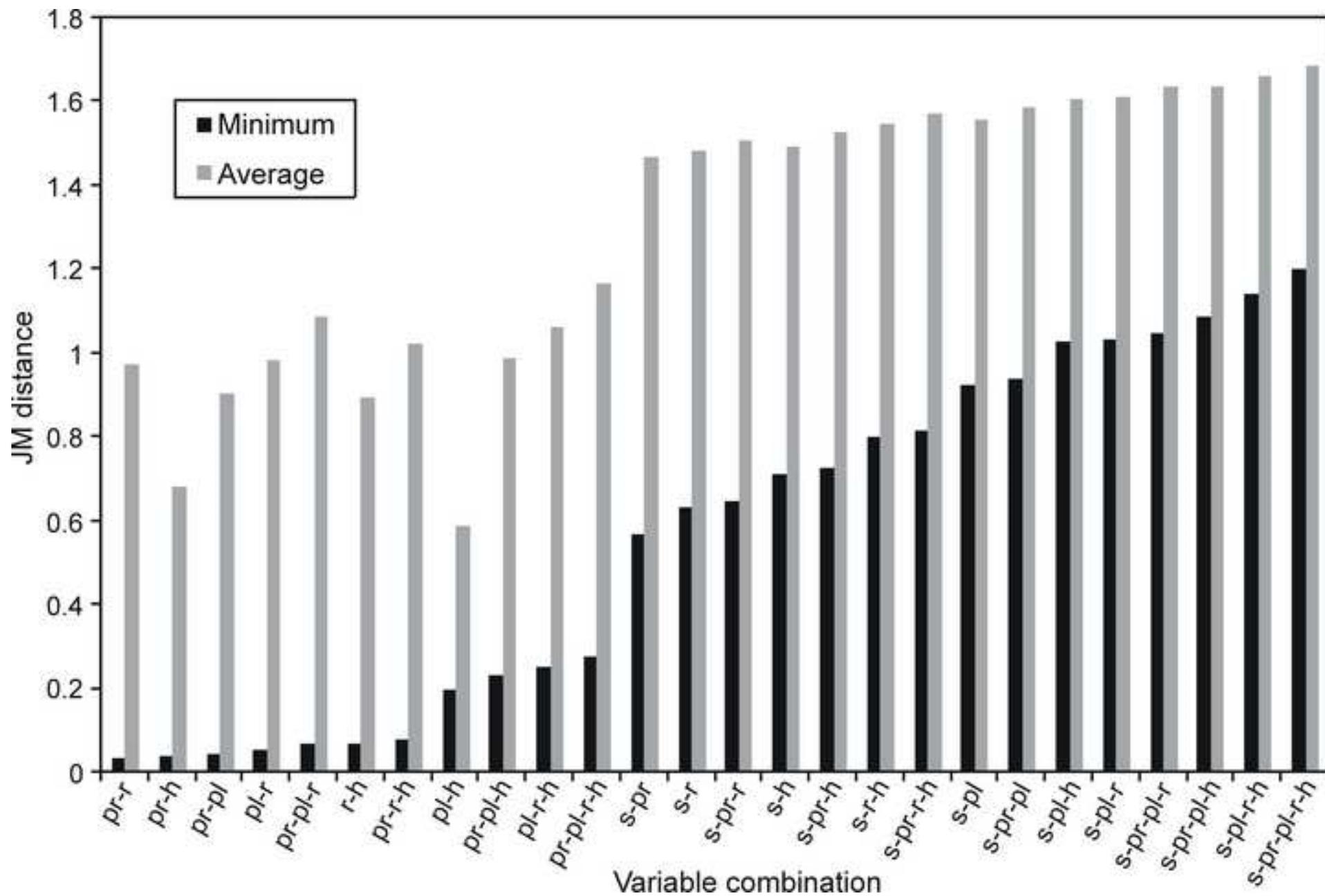


Figure 7
[Click here to download high resolution image](#)

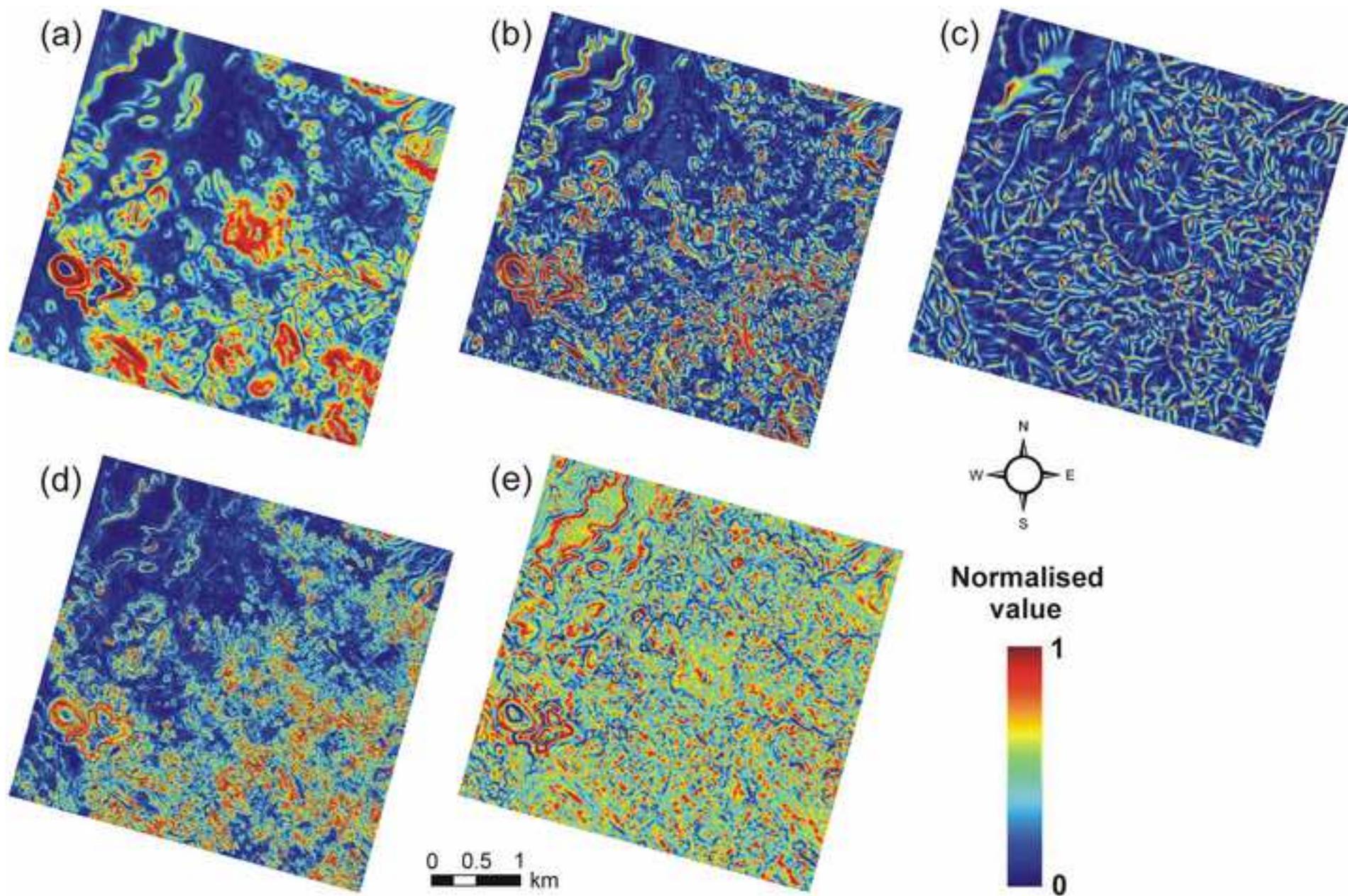


Figure 8
[Click here to download high resolution image](#)

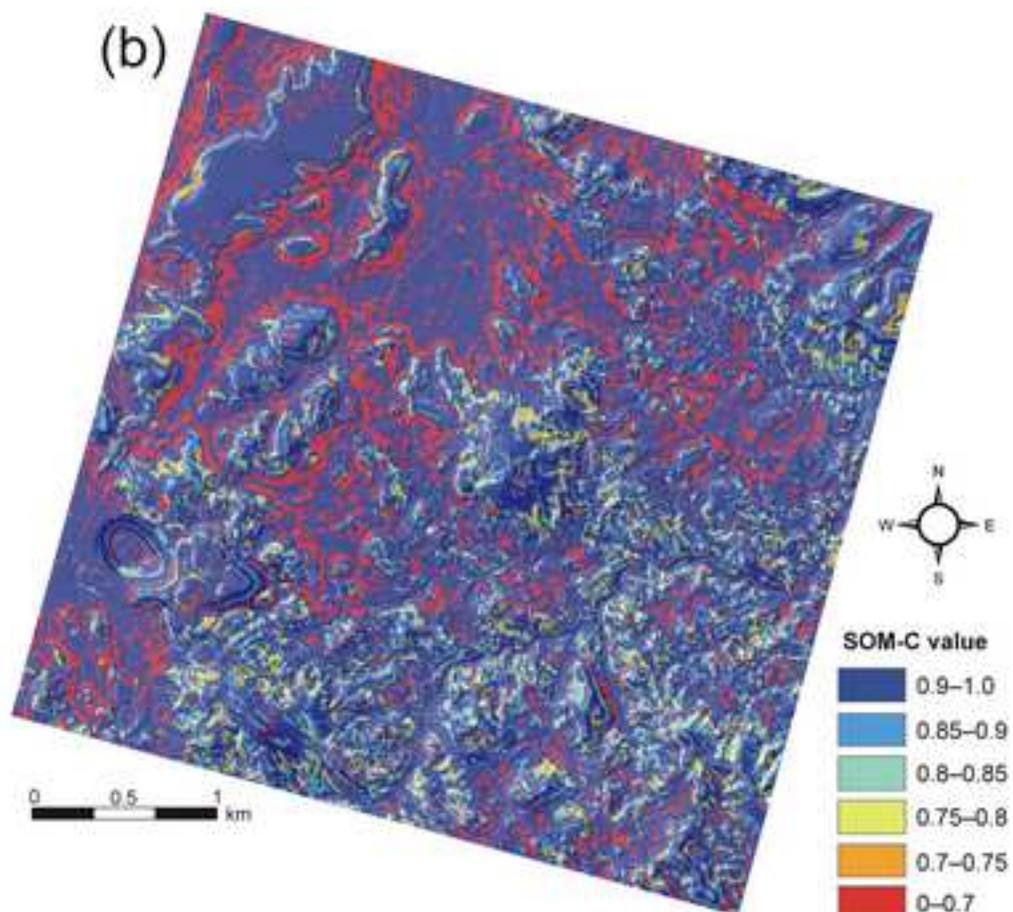
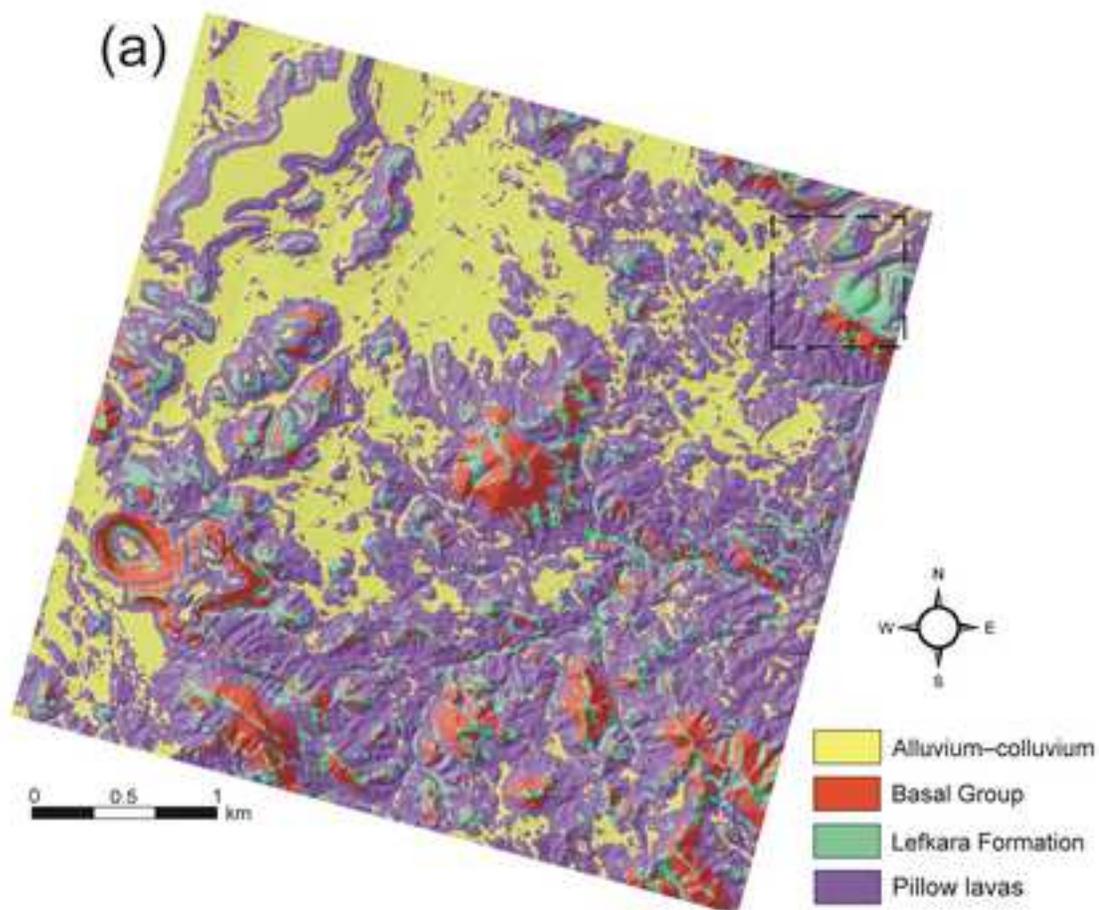


Figure 9
[Click here to download high resolution image](#)

